

**Ice Binding Structure and Mechanism of an  
Antifreeze Protein from Winter Flounder.**

By

Frank Sicheri, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

©Copyright by Frank Sicheri, July 1995

**Ice Binding Structure and Mechanism of an Antifreeze Protein from Winter Flounder.**

Doctorate of Philosophy (1995)

McMaster University

(Biochemistry)

Hamilton Ontario

Title: Ice Binding Structure and Mechanism of an Antifreeze  
Protein from Winter Flounder.

Author: Frank Sicheri B.Sc. (McMaster University)

Supervisor: Dr. Daniel S.C. Yang

Number of Pages: xii, 126

## ABSTRACT

The 1.5Å x-ray crystal structure of a lone  $\alpha$ -helical antifreeze protein (AFP) from winter flounder is presented, providing the first detailed look at an AFP and its ice binding features. The structure was determined using a non conventional multi-parameter molecular replacement scheme with two idealized  $\alpha$ -helices as the search models. The described method may prove useful for the determination of other  $\alpha$ -helical structures. The AFP's ice binding structure consists of four repeated ice binding motifs (IBMs), the side chains of which are inherently rigid or restrained by pair-wise side chain interactions to form a flat binding surface. Comparison shows that analogous IBM sequences are found highly conserved in all members of the  $\alpha$ -helical class of AFPs. Also revealed are the presence of elaborate N- and C-terminal cap structures which help to account for the AFP's unusual stability for a lone  $\alpha$ -helix in aqueous solution. Based on the crystal structure, a model amenable to testing is proposed which accounts for the proteins ice binding ability and specificity. Experiments to test the validity and generality of this ice binding model are proposed.

## **ACKNOWLEDGMENTS**

I would like to thank my supervisor Dr. Daniel Yang for providing me the environment to accomplish my goals and for his example as a scientist and friend. Thanks are also extended to the other members of my Research Committee, Dr. E. Pai, Dr. R.M. Epanand and Dr. W.W. Chan for their guidance. Great thanks to Susan Xue, Wai Ching Hon and Steve Bubanko for their help in the lab in my final year.

I would also like to thank my fellow members of the Comprehensives Support Group, Ernesto Assante, Jason Young and especially my long time friend Paul Ala for helpful scientific discussions.

To Sally, my mother and sister and all my other family members I give my love for all their spiritual support. In memory of my loving father, I dedicate this work.

# TABLE OF CONTENTS

---

	<b>PAGE</b>
<b>ABSTRACT</b>	iii
<b>ACKNOWLEDGMENTS</b>	iv
<b>TABLE OF CONTENTS</b>	v
<b>ABBREVIATIONS</b>	vii
<b>LIST OF FIGURES</b>	ix
<b>LIST OF TABLES</b>	xii
<b>CHAPTER 1 INTRODUCTION AND OBJECTIVES</b>	<b>1</b>
1.1 Introduction	1
1.2 Objectives	10
<b>CHAPTER 2 EXPERIMENTAL</b>	<b>11</b>
2.1 Protein Purification & Crystallization	11
2.2 Data Collection & Processing	12
2.3 Structure Solution	13
i. Phasing methods	13
ii. Refinement	35

<b>CHAPTER 3</b>	<b>STRUCTURE ANALYSIS</b>	<b>43</b>
3.1	Helix Stabilization	45
3.2	Ice Binding Structure	49
3.3	Solvent Structure	54
3.4	Crystal Packing Arrangement	59
<b>CHAPTER 4</b>	<b>ANTIFREEZE PROTEIN ICE BINDING MODEL</b>	<b>64</b>
4.1	Winter Flounder System	64
4.2	Implications for other Antifreeze Proteins	69
<b>CHAPTER 5</b>	<b>COMPARISON OF THE ANTIFREEZE PROTEIN STRUCTURE TO OTHER MODELS</b>	<b>72</b>
5.1	Previous X-Ray Crystal Structure	72
5.2	Simulation Studies	73
<b>CHAPTER 6</b>	<b>APPLICATION OF THE MOLECULAR REPLACEMENT STRATEGY TO OTHER <math>\alpha</math>-HELICAL STRUCTURES</b>	<b>77</b>
<b>CHAPTER 7</b>	<b>TESTING THE PROPOSED ICE BINDING MODEL</b>	<b>80</b>
7.1	NMR Studies on mini Antifreeze Protein Constructs	80
7.2	Site Specific Mutagenesis on a Recombinant AFP	83
7.3	Prediction of Ice Binding Specificities	85

<b>CHAPTER 8</b>	<b>CONCLUSIONS</b>	<b>91</b>
<b>GLOSSARY</b>		<b>92</b>
<b>APPENDIX A</b>	<b>An <i>XPLOR</i> shell script for a multi-parameter, two molecule translation search</b>	<b>105</b>
<b>APPENDIX B</b>	<b>An <i>XPLOR</i> shell script for refinement of clustered molecular replacement solutions</b>	<b>108</b>
<b>REFERENCES</b>		<b>115</b>

The x-ray crystal structure and the proposed ice binding model of the winter flounder AFP described herein has been accepted for publication in *Nature* (Sicheri & Yang, 1995).

The structure solution of the winter flounder AFP described herein has been submitted for publication in *Acta Crystallographica D section*.



## ABBREVIATIONS

---

AFP	=	antifreeze protein
CD	=	circular dichroism
FOM	=	figure of merit
GST	=	glutathione-S-transferase
HPLC	=	high performance liquid chromatography
MIRAS	=	multiple isomorphous replacement and anomalous scattering
NMR	=	nuclear magnetic resonance
MR	=	molecular replacement
PC	=	Patterson correlation coefficient
SA	=	simulated annealing

## LIST OF FIGURES

FIGURE	TITLE	PAGE
1.1	Typical antifreeze activity profiles	5
1.2	Mechanism of AFP ice crystal growth inhibition	7
1.3	Helical net representations of the alanine rich $\alpha$ -helical AFPs	8
2.1	Perspectives of the poly-alanine search model oriented in one of the 22 prominent MR rotation solutions	17
2.2	Rotational symmetry elements of the $\alpha$ -helical search model	19
2.3	Representative contour plots of the MR translation search results using observed structure factor amplitudes	20
2.4	Representative contour plots of the MR translation search results using calculated structure factor amplitudes	22
2.5	Possible AFP packing arrangements consistent with single molecule MR results	24
2.6	Representative contour plots of the multi-parameter PC and $R_{\text{factor}}$ search results	26
2.7	Refinement results of the top clustered subfamily solutions	28
2.8	Harker planes through anomalous and isomorphous difference Fourier maps	30
2.9	Cross sections through a MIRAS derived, solvent flattened, phase extended, electron density map	32

2.10	Stereo plot of the origin correlated SA refined, multi-parameter search solutions	34
2.11	Residue average B-factor plots for the 4°C and -180° winter flounder AFP crystal structures	38
2.12	Ramachandran plot for the four AFP molecules in the 4°C and -180° winter flounder AFP crystal structures	39
2.13	Stereo plot of the superimposed winter flounder AFP structures	40
2.14	Luzzati plots for the estimate of overall coordinate error	41
2.15	Representative simulated annealing 'omit' electron density maps	42
3.1	Ribbon diagrams of AFP molecules one and two @ 4°C	44
3.2	Stereo plots of AFP molecule 2 @ 4°C showing the helix stabilizing salt bridge interaction between lysine 18 and glutamate 22	46
3.3	Stereo plots of the N- and C- terminal cap structures	47
3.4	Stereo $2 F_o  -  F_c $ electron density map of a winter flounder ice binding motif	50
3.5	N-terminal axial view of the AFP's ice binding surface	51
3.6	Cross sections through representative IBMs	53
3.7	Schematic of the spatial arrangement of the winter flounder AFP's ice binding groups	55
3.8	Superimposed solvent structure around the winter flounder alanine residues	57
3.9	Superimposed solvent structure around the winter flounder IBMs	58

3.10	Stereo plot of the hydrophobic crossing sites between non symmetry related AFP molecules	61
3.11	Stereo view of N-terminal cap structures affected by crystal packing interactions	63
4.1	Ice binding model proposed by Wen and Laursen (1992b)	65
4.2	Winter flounder AFP ice binding model	67
4.3	Front and side perspectives of a winter flounder AFP molecule docked along the $\langle 01\bar{1}\bar{2} \rangle$ binding axis of the $\{20\bar{2}1\}$ ice plane	68
7.1	AFP model peptides for NMR and X-Ray crystallographic analysis	81
7.2	Wild type and site directed mutants of a recombinant winter flounder AFP	84
7.3	Ice binding specificity assay	86
7.4	Predicted binding specificity of the yellowtail flounder AFP	89
7.5	Predicted binding specificity of a novel winter flounder AFP	90

## LIST OF TABLES

TABLE	TITLE	PAGE
1.1	AFP classes	2
2.1	Heavy atom derivative data collection and phasing statistics	14
2.2	Native data collection and refinement statistics	15
2.3	Subfamily $R_{\text{factor}}$ statistics following various stages of crystallographic refinement	29
2.4	Identification of AFP rotations about their helix axes by the SA refinement of systematically placed threonine residues	36
3.1	Intermolecular hydrogen bonding interactions of the 4°C winter flounder crystal structure	60

# CHAPTER 1 INTRODUCTION AND OBJECTIVES

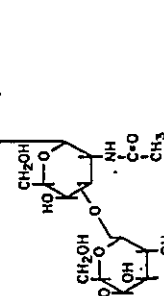
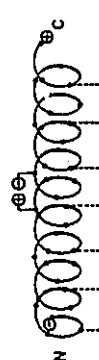

## 1.1 Introduction

Antifreeze proteins (AFPs) extend the habitat of many fish into polar regions that would normally put them at great risk of freezing. These proteins are found circulating in blood at high concentrations and provide a line of defense against the growth of ice crystals in the fish's body. A number of excellent reviews on AFPs have been published (Feeney and Yeh, 1978; DeVries, 1983; DeVries, 1984; Ananthanarayanan, 1989; Hew & Yang 1992).

Structurally diverse, AFPs can be loosely grouped into three classes (Table 1.1). The first class of AFPs, found in species of cod, are made up of 4 to 40 repeats of the tripeptide Ala-Thr-Ala. Onto each threonine residue is attached the disaccharide  $\beta$ -D-galactosyl-(1 $\rightarrow$ 3)- $\alpha$ -N-acetyl-D-galactosamine which has been shown by chemical modification to be critical for the protein's antifreeze function (Ahmed, Osuga & Feeney, 1973). The variation of tripeptide repeats gives rise to AFP isoforms which range in size from 2,600 to 33,700 Da. CD (Ananthanarayanan, 1989) and NMR (Rao & Bush, 1987) studies show that these AFPs adopt a very flexible extended structure at 0° to 4°C . NMR and modeling studies suggest that these AFPs sample a short range three fold helical conformation in solution (Rao & Bush, 1987).

The second class of proteins are the  $\alpha$ -helical AFPs from species of flounder and

**Table 1.1 AFP classes.**

	Fish type	Distinguishing class features
<p><b>Class I: Carbohydrate containing AFPs.</b></p> <p><math>[-Ala-Ala-Thr-]_{n-1-40}</math></p> 	<p>Antarctic cod</p> <p>Greenland cod</p> <p>Atlantic cod</p> <p>Atlantic tomcod</p> <p>Arctic polar cod</p> <p>Saffron cod</p>	<p>-4 to 40 repeats of the glycosylated tripeptide Ala Thr Ala glyco = <math>\beta</math>-D-galactosyl-(1 <math>\rightarrow</math>3)-<math>\alpha</math>-N-acetyl-D-galactosamine. -extended and highly flexible backbone structure. -Mr = 2,600-33,700</p>
	<p>Winter flounder</p> <p>Yellowtail flounder</p> <p>Shorthorn sculpin</p> <p>Grubby sculpin</p> <p>Arctic sculpin</p>	<p>-alanine rich, 75 to 100% <math>\alpha</math>-helical at 0-4°C</p> <p>-extended but ordered structures.</p> <p>-winter flounder AFP believed to adopt a lone <math>\alpha</math>-helical conformation.</p> <p>-Mr = 3,300 to 6,000</p>
<p><b>Class II: <math>\alpha</math>-helical AFPs.</b></p> 	<p>Sea raven</p> <p>Smelt</p> <p>Ocean pout</p> <p>Arctic eel pout</p> <p>Arctic eel pout</p>	<p>-a grouping of non-homologous globular AFPs</p> <p>-sea raven, smelt and herring AFPs predicted to adopt a C-type lectin tertiary fold.</p> <p>-ocean pout AFP shown to adopt a novel B-sheet protein fold.</p> <p>-Mr = 5,000 to 24,000</p>
<p><b>Class III: Globular AFPs.</b></p> 		

sculpin. These AFPs range in size from 3,300 to 6,000 Da and are rich in alanine. CD studies show them to be very ordered at 0 to 4°C (Ananthanarayanan & Hew, 1977; Chakrabarty, Ananthanarayanan and Hew, 1989; Wen & Laursen, 1992). Similar in one respect to the first class of AFPs, the  $\alpha$ -helical AFPs form elongated structures; Patterson maps of x-ray diffraction data collected from winter flounder AFP crystals are consistent with a lone  $\alpha$ -helical structure (Yang, Sax, Chakrabarty & Hew, 1988).

The third class of AFPs comprises of a number of non-homologous AFPs that adopt compact globular structures which range in size from 5,000 to 24,000 Da. The best characterized member of this class is an ocean pout AFP (isoform 12), that was shown by NMR methods to adopt a novel  $\beta$ -sheet protein fold (Sonnichsen, Sykes, Chao & Davies, 1993). AFPs from sea raven, smelt and herring also adopt globular like protein folds. However, these AFPs show primary sequence homology (23-33% identity) to members of the carbohydrate recognition domain super-family (Ewart, Rubinsky, & Fletcher, 1992; Ng & Hew, 1992; Ewart, K.V., & Fletcher, G.L., 1993) suggesting an entirely different fold from that possessed by ocean pout AFPs. Based on the primary sequence homology, the tertiary structure of the sea raven AFP has recently been modeled (Sonnichsen, Sykes & Davies, 1995).

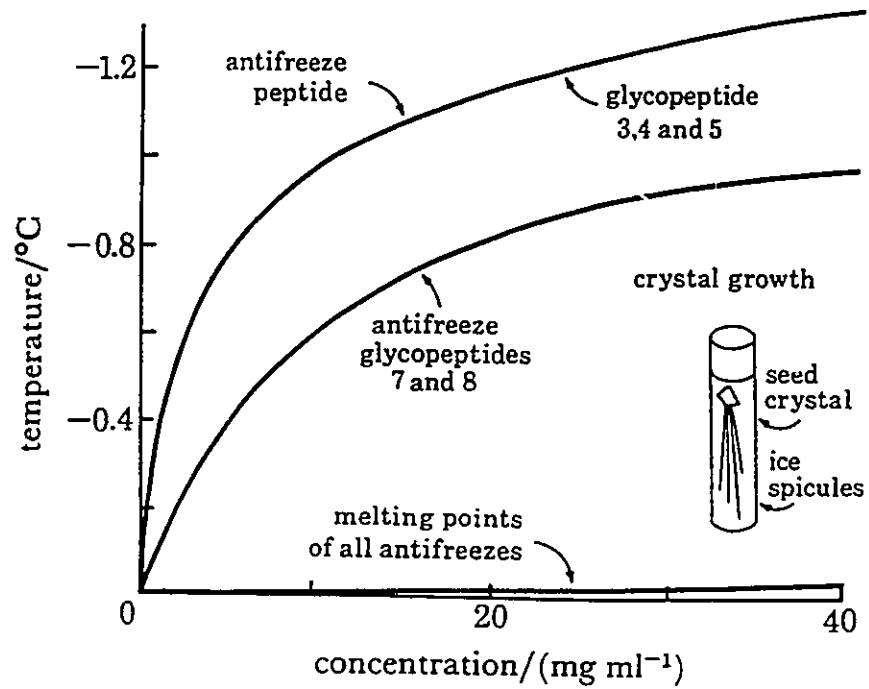
Although structurally diverse, all AFPs are believed to function similarly. AFPs have the ability to completely inhibit the growth of ice crystals thus reducing the apparent freezing temperature of a solution. The melting temperature, however, is left largely unaffected. The result is the paradoxical separation of freezing and melting



temperatures. Typical profiles of the effect of AFPs on a solution's melting and apparent freezing temperatures are shown in figure 1.1. The difference between melting and apparent freezing temperatures is formally defined as antifreeze activity. At physiological concentration of 20 to 40 mg/ml, the more effective AFPs provide a saturating level of approximately 1°C antifreeze protection. This is sufficient to protect fish in polar salt waters where temperatures can fall to as low as -1.9°C, approximately 1°C below the freezing temperature of most organisms body fluids (DeVries, 1984).

In addition to causing a thermal hysteresis effect, antifreeze proteins also inhibit the recrystallization of ice from the solid phase. After a freezing event, a large number of ice crystals is produced. With time, the smaller ice crystals shrink while the larger crystals grow. This recrystallization process is believed to further increase the physical damage to an organisms cells and tissues (Knight & Duman, 1986). A number of 'antifreeze like' proteins which are believed to function mainly to inhibit ice recrystallization have been discovered in insects, plants, bacteria and fungi (Urrutia, Duman & Knight, 1992; Duman & Olser, 1993). Unlike fish which employ AFPs to prevent freezing altogether, these organisms can endure freeze thaw cycles.

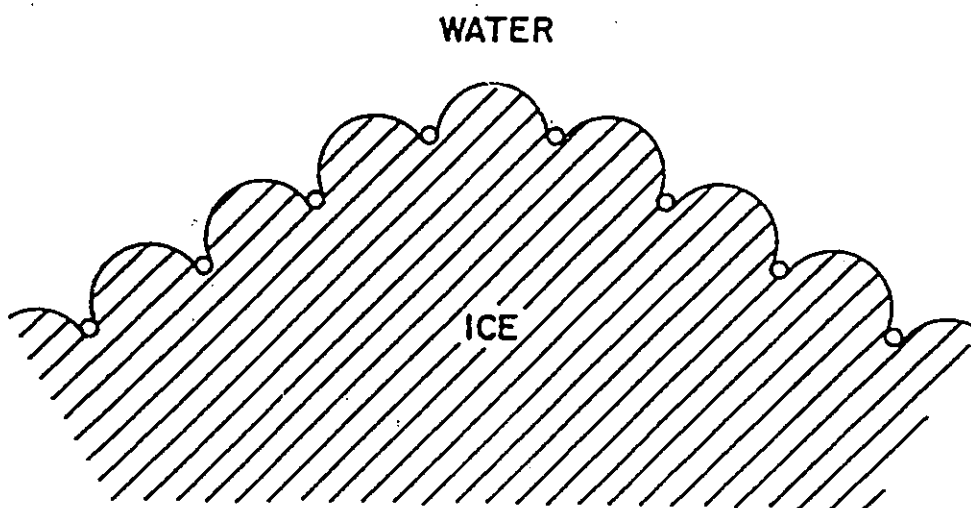
Antifreeze activity has its origins in the AFP's ability to bind ice. At 0°C and 1 atmosphere pressure, water crystallizes in the hexagonal space group  $P6_3/mmc$  (Megaw 1934). This form of ice is referred to as ice Ih. While most proteins are excluded from growing ice, AFPs readily adsorb onto ice surfaces and become incorporated into the ice itself (Knight, Cheng & DeVries, 1991; Knight, Driggers & DeVries, 1993). The



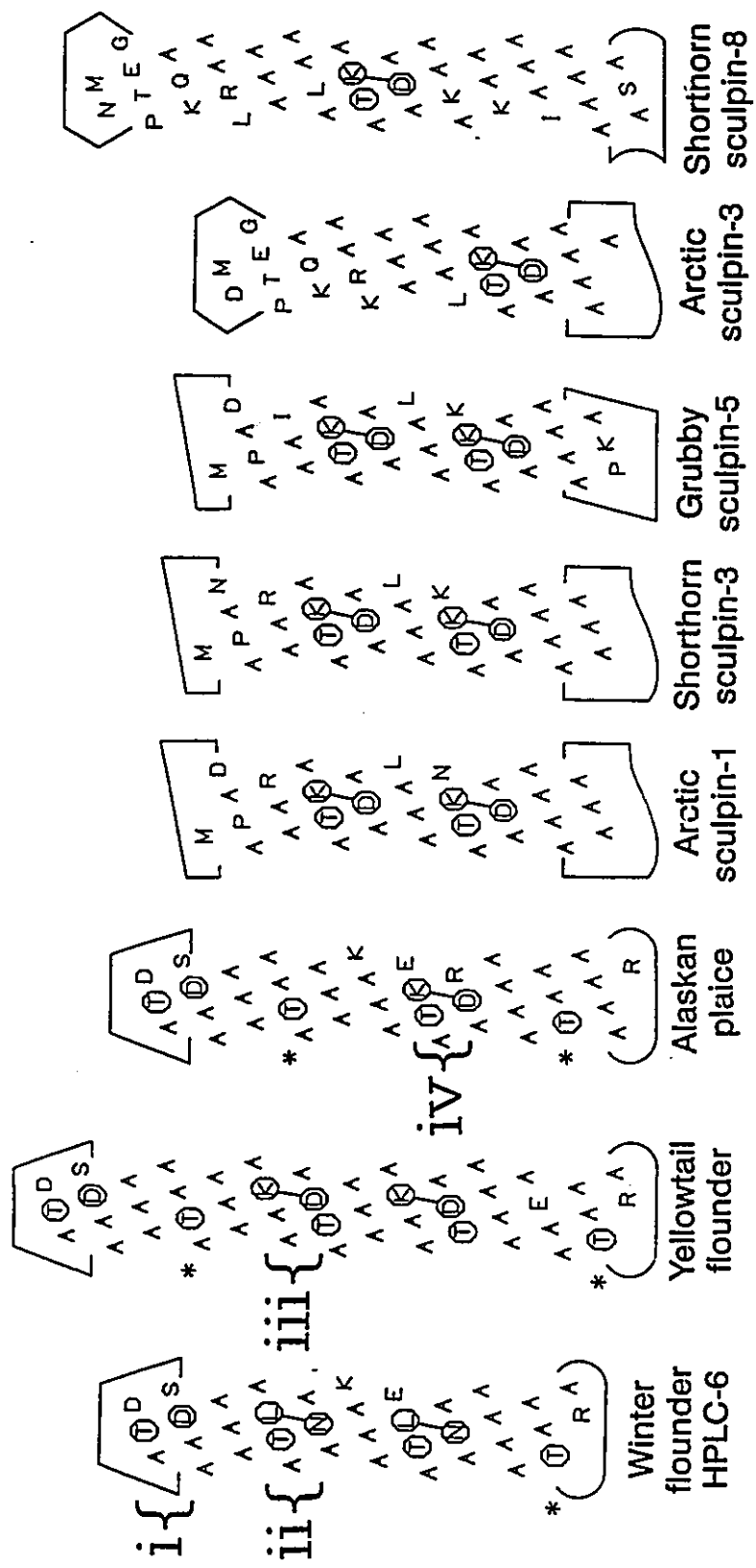
**Figure 1.1** Typical antifreeze activity profiles. AFPs depress the apparent freezing temperature of a solution but leave the melting temperature largely unaffected. Reproduced from DeVries, 1984.

amount of AFP bound to the surface of ice has been shown by spectroscopic studies to correlate directly with the magnitude of the antifreeze effect (Brown, Yin, Burcham & Feeney, 1985). Furthermore, ice binding appears highly specific with each AFP displaying a unique preference for a plane(s) and orientation(s) on ice (Knight, Cheng & DeVries, 1991; Knight, Driggers & DeVries, 1993). The method used to characterize the binding specificity of AFPs represents a milestone in AFP research and will be described further in chapter 5.3. Inhibition of ice crystal growth is believed to occur by way of the Kelvin effect (fig 1.2), whereby the surface area of ice increases unfavorably as ice is forced to grow around adsorbed protein molecules (Raymond & DeVries, 1977; Knight & DeVries, 1989). This kinetic versus thermodynamic origin accounts for the paradox of the thermal hysteresis effect. While the 3D structures of class 1 AFPs from cod (Rao & Bush, 1987) a class 2 AFP from winter flounder (Yang, Sax, Chakrabarty, & Hew, 1988) and a class 3 AFP from ocean pout (Sonnichsen, Sykes, Chao & Davies, 1993) have been analyzed by NMR and x-ray crystallographic techniques to demonstrate distinct tertiary folds, very little detailed information regarding the ice binding structure of each has been revealed due to limits in resolution. As a result, our understanding of the AFP's ice binding mechanism is based largely on inferred protein structure (Yang, Sax, Chakrabarty & Hew, 1988; Wen & Laursen, 1992; Chou, 1992; Knight, Driggers & DeVries, 1993; Lal, Clark, Lips, Ruddock & White, 1993; Madura et al., 1994).

The extensively characterized AFPs from winter flounder belong to the alanine



**Figure 1.2** Mechanism of AFP ice crystal growth inhibition: the Kelvin Effect. The surface area of ice increases unfavorably as ice is forced to grow around adsorbed protein molecules (shown as open circles). The increase in the surface curvature of ice results in an increase in the free energy of the surface. Reproduced from Knight, Cheng, & DeVries, 1991.



**Figure 1.3** Helical net representations of the alanine rich  $\alpha$ -helical AFPs. Representative IBMs are labeled *i* to *iv*. End-cap sequences which are predicted to form similar 3-D structures are shown with common cartoon shapes. IBMs are shown with circled residues. Residues demonstrated in the winter flounder x-ray structure or predicted in the other  $\alpha$ -helical AFPs to interact to stabilize ice binding conformations are shown joined by lines. Incomplete IBMs are shown with asterisks. AFP sequences are taken from Knight, Cheng & DeVries, 1991; Hew, Chakrabarty & Yang, 1987.

rich,  $\alpha$ -helical class of AFPs (fig. 1.3). Their small size and periodic nature, based loosely on the repeat of eleven amino acids, makes them ideally suited for the study of their ice binding mechanism. The two major isoforms found, HPLC6 and HPLC8, are distinct gene products as revealed by their cDNA sequences (Davies, Roach & Hew, 1982; Pickett, Scott, Davies, Wang, Joshi & Hew, 1984). In mature form, both AFPs are composed of 37 amino acids but differ in the positioning and composition of a helix stabilizing (Chakrabarty & Hew 1991) salt bridge (HPLC6: Lys 18/Glu 22; HPLC8: Lys 22/Asp 26). Synthesized in the liver as pro-AFP forms, the AFPs are processed in the blood to mature forms by the removal of a 21 N-terminal poly-peptide sequence and a C-terminal glycine residue (Hew, Wang, Yan, Cai, Sclater & Fletcher, 1986). The C-terminus of the mature AFP is amidated (Hew, Wang, Yan, Cai, Sclater & Fletcher, 1986).

Binding studies have shown that the winter flounder AFPs bind specifically along the  $\langle 01\bar{1}2 \rangle$  axis of the  $\{20\bar{2}1\}$  pyramidal ice planes (Knight, Cheng & DeVries, 1991). In addition, mutation studies have demonstrated that the spatial arrangement of the AFP's ice binding residues (Thr, Asn, Asp and Leu residues shown circled in figure 1.3) are critical for maintaining antifreeze function (Wen & Laursen, 1992). A previous x-ray crystal structure demonstrated the HPLC6 isoform to be a single  $\alpha$ -helix. This would locate polar ice binding residues on helix face and the helix stabilizing salt bridge to a relatively apolar face composed mainly of alanine (Yang, Sax, Chakrabarty, & Hew, 1988). In that structure, side chains and the helix termini of both AFP

molecules in the asymmetric unit appeared disordered, making it difficult to account for the AFPs's then poorly understood ice binding specificity. While solvent exposed side chains (all side chains are solvent exposed in a lone  $\alpha$ -helical structure) are commonly found disordered in protein structures and the termini of lone  $\alpha$ -helices are inherently unstable (Chakrabarty & Baldwin, 1995), refinement statistics ( $R_{\text{factor}}$  which would not improve beyond 27.5% for 2.5-5Å data) suggested the possibility of error in the x-ray model. With the binding specificity of the winter flounder AFP now well characterized (Knight, Cheng & DeVries, 1991), a detailed understanding of the winter flounder structure remains the greatest impediment to unraveling the origin of the AFP's ice binding mechanism. Presently, the current understanding of how the winter flounder AFPs bind to ice is based largely on side chain conformations inferred from molecular modeling studies (Knight Driggers & DeVries, 1991; Wen & Laursen, 1992; Chou, 1992; Lal, Clark, Lips, Ruddock & White, 1993; Madura *et al.*, 1994).

## 1.2 Objectives

To investigate the origin of the winter flounder AFP's ice binding ability and specificity by determining to high resolution its structure by x-ray crystallographic methods

## CHAPTER 2 EXPERIMENTAL

### 2.1 Protein Purification & Crystallization

The presence of greater than seven very similar AFP isoforms in the serum of the winter flounder fish presents a problem for protein purification (Fourney, Joshi, Kao & Hew, 1984). Through a modification of an established purification protocol (Davies, Roach & Hew, 1982), protein of increased purity was obtained from which many crystals diffracting to greater than  $1.5\text{\AA}$  were produced. Crystals from previous protein preparations diffracted to approximately  $2.5\text{\AA}$ . The previous purification procedure involved the use of size exclusion, ion exchange, and reverse phase chromatography. In the modified procedure, ion exchange chromatography was replaced with an isoelectric focusing technique as described below.

Lyophilised semi-purified AFP, prepared from fish sera by exclusion chromatography with Sephadex G75, was kindly provided by Dr. Choy Hew and Dr. Garth Fletcher. Isoelectric focusing using a Biorad Rotofor unit was employed to fractionate AFP isoforms. 50 to 150 mg of semi-purified protein was applied in 6M urea, 5% wt./vol. ampholyte (pH 3-10), and focused for 5.5 hours @12W/5°C. Fractions from pH 5.2 to 8.5 which were shown by analytical reverse phase HPLC to contain the two main AFP isoforms, HPLC6 and HPLC8, were then pooled. The two main isoforms were resolved by preparative reverse phase HPLC and lyophilised



(analytical column=Waters  $\mu$ Bondapak C<sub>18</sub>, 30cm x 7.1mm I.D.; preparative column=Waters DeltaPak C<sub>18</sub>, 30cm X 19 mm I.D.; gradient=15 to 40% acetonitrile against 0.5% aqueous trifluoroacetic acid). A typical yield of the HPLC6 isoform from semi-purified AFP was 5-10%. The HPLC6 AFP isoform @ 4mg/ml was crystallized from a 5:1 mixture of acetone and 20mM Tris buffer (pH 7.2) at 4°C as reported (Yang, Chung, Chen, Rose & Hew 1986).

## 2.2 Data Collection & Processing

Diffraction data sets were collected at 4°C on crystals mounted in capillary tubes using a Rigaku RAXIS IIc area detector. Cu-K $\alpha$  radiation was produced using a Rigaku RU200 generator operating at 3kW (60mA x 50kV), with 0.3 mm cathode, SUPPER double focusing mirror optics and nickel filter. 5° oscillations were employed with exposure times of 30 to 40 minutes. Crystals were very stable to x-ray irradiation. Data was processed using software provided by Rigaku. Isomorphous LaNO<sub>3</sub> and TbNO<sub>3</sub> heavy atom derivatives were prepared by soaking crystals in 50 mM acetone solutions for 4 to 12 hours. Scaling of heavy atom derivative to native data sets was performed with the *PHASES* package of crystallographic programs (Furey, 1990). A -180°C native data set diffracting to greater than 1.5Å was collected using a rayon loop mounted crystal with butanol as the cryoprotectant. Crystal temperature was maintained using a Molecular Structure Corp. cooling device. Data

collection parameters and statistics are provided in tables 2.1 and 2.2. Descriptions of crystallographic terms and methods shown in the body of this work in bold italicised text are provided in the Glossary section.

## 2.3 Structure Solution

### i. Phasing Methods

Structure determination was approached using both *molecular replacement* and *isomorphous replacement* techniques. Due to problems with heavy atom solubility, non-isomorphism and the limited number of targetable groups on the AFP, good heavy atom derivatives proved difficult to obtain. Two suitable derivatives were found late in the structure determination by which time significant progress had been made by molecular replacement. In the end, the structure @ 4°C was solved by molecular replacement using a non-conventional multi-parameter two molecule search with heavy atom derivatives providing an independent confirmation of the correctness of the molecular replacement solution. Subsequently, the 4°C structure was used to solve the -180°C structure. All described molecular replacement approaches were performed with the program *XPLOR* (Brunger, 1992).

Conventional molecular replacement approaches to solving the AFP crystal structure proved unsuccessful. This involved using the *Patterson function*, native structure factor amplitudes and a single search molecule to first determine the rotation and then translation parameters of each molecule in the asymmetric unit. Up to three

**Table 2.1 Heavy atom derivative data collection and phasing statistics.**

Derivative/crystal #	LaNO <sub>3</sub> /1	LaNO <sub>3</sub> /2	LaNO <sub>3</sub> /3	TbNO <sub>3</sub> /1	TbNO <sub>3</sub> /2
Resolution (Å)	3.0	3.0	3.0	3.0	3.0
R <sub>sym</sub> (%)	10.2	8.2	9.2	10.0	6.5
R <sub>iso</sub> (%)	18.0	17.8	17.8	14.0	13.2
# observed ref.	4448	2316	2782	4869	4479
# unique ref.	1222	889	978	1200	1253
completeness (%)	91.0	68.2	75.0	90.6	93.8
R <sub>cullis</sub> (%)	57	70	59	60	64
Phasing Power	1.82	1.5	1.55	1.54	1.52
Figure of Merit	0.38	0.33	0.34	0.41	0.41
# phased ref.	1087	845	922	1075	1158

**The overall mean figure of merit for 1184 phased reflections is 0.76**

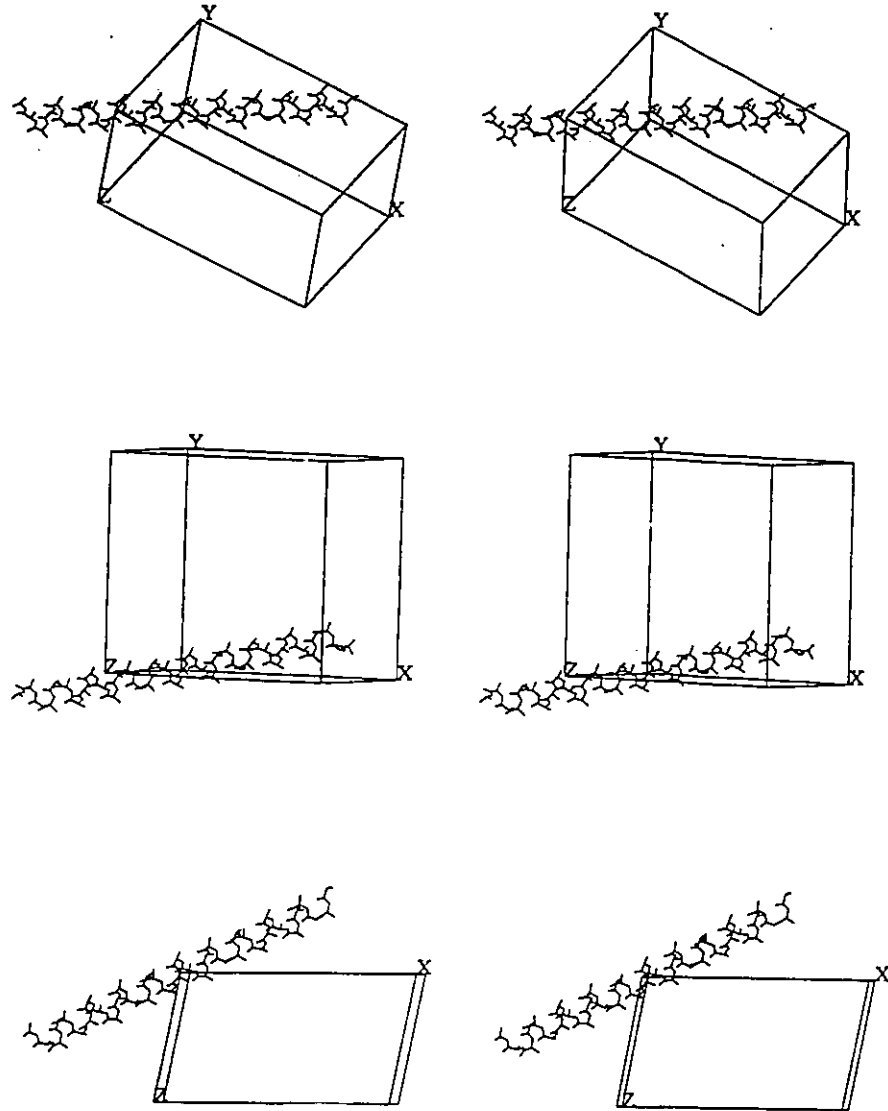
**Table 2.2 Native data collection and refinement statistics.****Space Group P2<sub>1</sub>,****Cell parameters**

<b>4°C</b>	a=38.17(7)Å	b=37.01(6)Å	c=21.84(2)Å	β=101.1(1)°
<b>-180°C</b>	a=37.41(3)Å	b=36.56(4)Å	c=21.35(2)Å	β=103.34(6)°

<b>Data collection statistics</b>	<b>4°C</b>	<b>-180°C</b>
Resolution (Å)	1.7-40	1.5-40
No. of observed reflections (> 1σ)	16614	28251
No. of unique reflections (> 1σ)	5505	8329
Completeness (%)		
overall	81.7	89.2
1.8-1.7Å	60.5	----
1.6-1.5Å	----	79.3
R <sub>sym</sub>	4.4%	5.4%
<b>Refinement parameters</b>		
No. of molecules per asymmetric unit	2	2
No. of protein atoms	454	454
No. of water molecules	39	130
Resolution	1.7-8.0Å	1.5-8.0Å
No. of reflections (> 1σ)	5429	8218
R <sub>factor</sub> /R <sub>free</sub>	17.9/22.6%	17.7/19.6%
Deviation from ideal bond distances	0.014 Å	0.013 Å
Deviation from ideal bond angles	2.2°	1.8°
Disorder Regions	Asn 27 molecule 1	None
Est. Coord. Error (Luzzati plot)	0.18Å	0.15Å

translation parameters (depending on the symmetry of the crystallographic space group) and three rotation parameters define the position of a molecule in a unit cell. A combined search of all 9 orientation parameters is not usually feasible with today's computer technology but with the use of the Patterson function, rotation and translation parameters can be searched separately. Rather than a single 9 parameter search, two significantly smaller 3 parameter searches are performed.

In the winter flounder  $P2_1$  crystal system, an estimated solvent content of 47% from *Mathews coefficient* calculations suggested that two AFP molecules were present in the asymmetric unit. Since CD and a previous x-ray structure analysis strongly suggested that the winter flounder AFP adopts a lone  $\alpha$ -helical structure, a 37mer of alanine in an idealized  $\alpha$ -helical conformation ( $\phi = -57^\circ$ ,  $\psi = 47^\circ$ ) was employed as the search model. The use of this model failed to identify unique rotation and translation solutions. In addition, *self rotation* studies failed to identify a non crystallographic symmetry axis relating the two AFP molecules in the asymmetric unit. In the rotation searches, *Patterson correlation coefficient* (PC) refinement of the top rotation function peaks revealed 22 strong orientations ( $> 7\delta$ ), all of which possessed the same helix axis tilt with respect to the crystallographic axes. Perspectives of one of the rotation solutions relative to the crystal axes are shown in figure 2.1. The helix axis lies on the diagonal plane  $Z = -X$ , approximately  $13^\circ$  from normal to the unique crystal axis  $Y$ . Two groups of 11 solutions differed in the direction of the AFP molecules (referred to

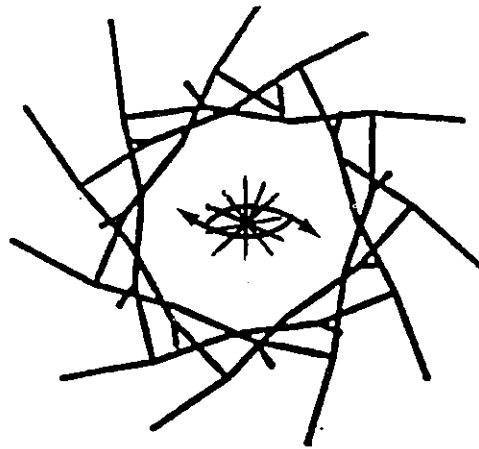
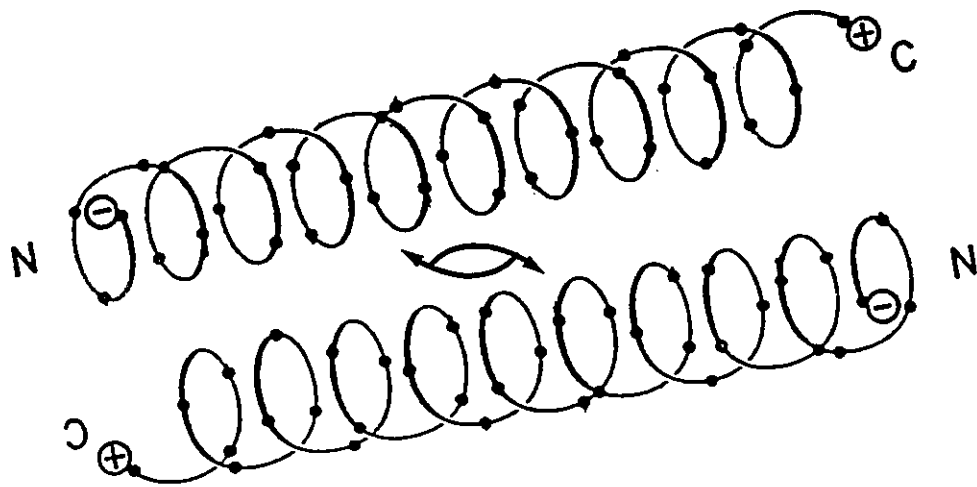


**Figure 2.1** Perspectives of the poly-alanine search model oriented in one of the 22 prominent ( $>7\sigma$ ) molecular replacement rotation solutions. All 22 solutions possess the same helix axis tilt.

as amino terminus up or down with respect to the Y axis) and the 11 solutions in each group were related by a  $31^\circ$  periodicity about the helix axis. This rotational ambiguity reflects the symmetry of the  $\alpha$ -helical structure with 11 fold and two fold rotational symmetry elements parallel and perpendicular to the helix axis respectively (figure 2.2). Unfortunately, none of the 22 rotation solutions were discernibly better at any of the resolution shells tested and this rotational ambiguity could not be broken by adding side chains in favoured conformations to the search model.

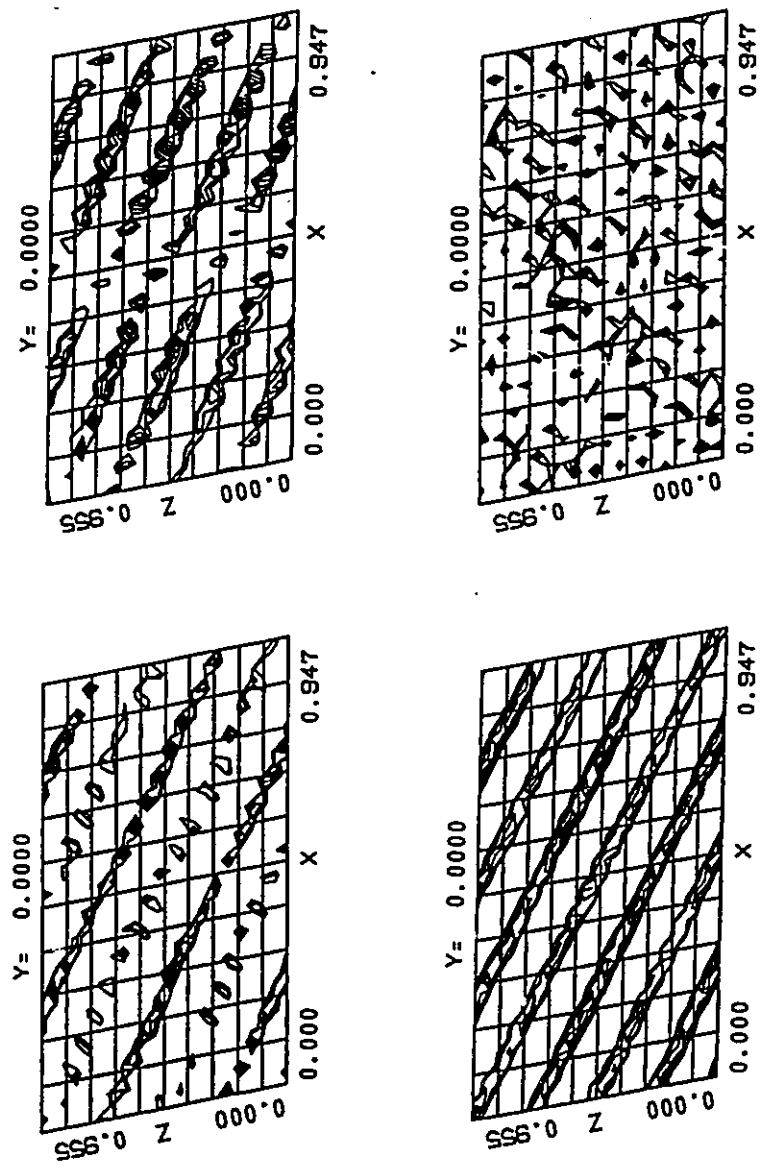
Using many of the rotation solutions, translation searches were carried out. For the  $P2_1$  space group, translation searches are restricted to the XZ plane. Rather than two solutions corresponding to the positions of the two molecules in the asymmetric unit, families of solutions appearing as two distinct ridges ( $Z=-X+1/2$  and  $Z=-X+1/4$ ) parallel to the projection of the helix axis on the XZ plane were observed (fig 2.3). The relative height of the two ridges and the position of peaks on each ridge varied greatly with the resolution limits used. This translational ambiguity appears to be due to the tilt of the AFP molecules in the unit cell with helix axes aligned nearly perpendicular ( $13^\circ$  off) to the unique crystal axis Y.

In order to explore the possibility of circumventing the rotational and translational ambiguities that were encountered, molecular replacement searches were performed on error free (calculated) diffraction data. It was hoped that this would allow us to optimize the molecular replacement methods for subsequent use with the observed structure factor amplitudes. Structure factor amplitudes were calculated for



**Figure 2.2** Rotational symmetry elements of the  $\alpha$ -helical search model which give rise to ambiguity in the molecular replacement rotation searches. Top, two fold rotation axes perpendicular to the helix axes; Bottom, 11 fold screw axis parallel to the helix axis.

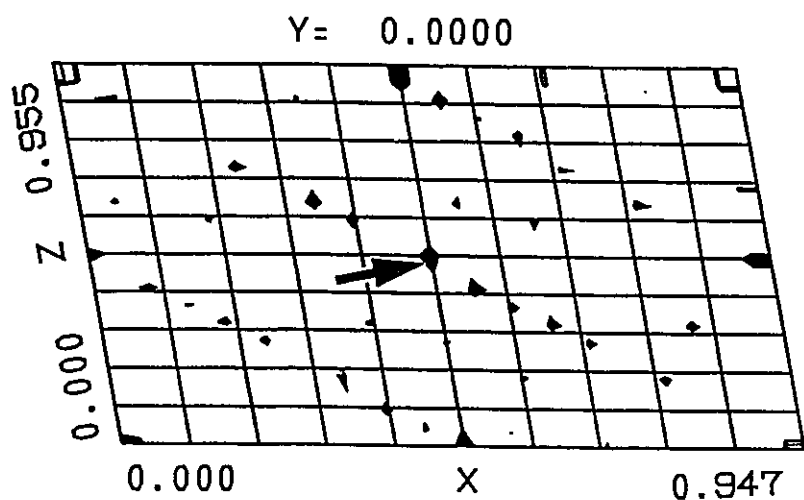
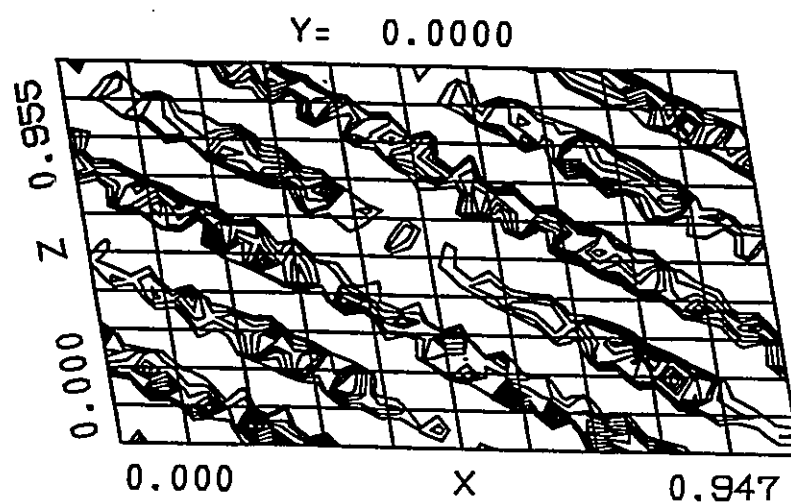




**Figure 2.3** Representative contour plots of molecular replacement translation search results using a poly-alanine search model and the observed structure factor amplitudes. Resolution ranges as read from top left to bottom right; 3-15Å, 3-10Å, 2-15Å, 3-6Å. Plots are contoured at 70 to 100% of the total PC variation with increments of 5%.

two AFP molecules, built with random side chain and idealized  $\alpha$ -helical backbone conformations, positioned in the experimental unit cell with rotations and translations selected from our single molecule search results. Rotation and translation searches were then performed using the poly-alanine search model and the same rotational and translational ambiguities found with the experimentally derived structure factor amplitudes were observed. Only by including all side chains in the search model, in proper conformations, were correct rotation and translation distinguishable. Still, the related rotation and translation solutions were still apparent. Representative plots of the translation search results are shown in figure 2.4. Since the winter flounder AFP side chain conformations were not known beforehand, it appeared that a conventional molecular replacement approach would not be effective in surmounting the inherent rotational and translational difficulties of the winter flounder AFP crystal system.

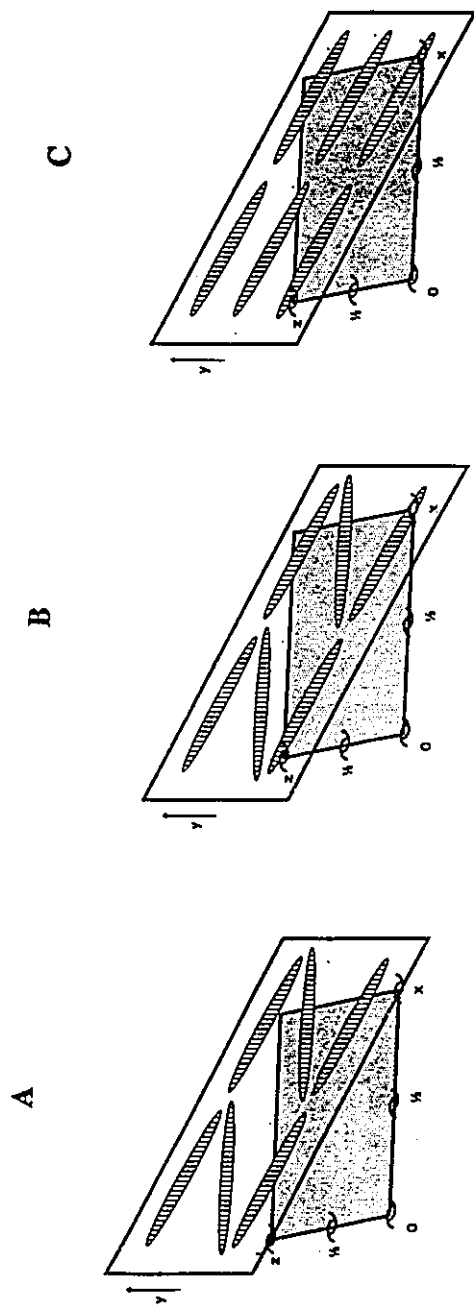
A combined two molecule multi-parameter search was then attempted. It was hoped that by searching with both molecules of the asymmetric unit simultaneously that the rotational and translational ambiguities of the AFP crystal system could be resolved. As noted, multi-parameter searches are not usually tractable for the determination of protein structures because computational requirements increase prohibitively with the number of parameters searched at one time. By including greater than one molecule in a search at one time, the number of total parameters becomes more excessive. However, in the winter flounder AFP crystal system, the



**Figure 2.4** Representative contour plots of molecular replacement translation search results using calculated factor amplitudes and a poly-alanine search model (top) and a model with all side chains included (bottom). The single arrow shows the expected position of the AFP molecules ( $0,0,0$  and  $0.5\ 0.5\ 0.5$  are equivalent positions in the Patterson map). Resolution ranges =  $3\text{-}15\text{\AA}$ . Plots contoured at 70 to 100% of the total PC variation with increments of 5%.

conventional MR studies reduced the number of orientation parameters to be searched significantly, making a combined multi-parameter search feasible. After the conventional MR search, the  $\alpha$ -helical tilts of the AFP molecules were confidently known but ambiguity existed in the rotation of each AFP molecule about its helical axis, the direction of each molecule (up or down), the translation of each molecule on the XZ plane (confidently restricted to the diagonal ridges) and the relative positioning of the two AFP molecules along the Y axis. To reduce the number of parameters further, we decided to hold fixed the rotation of each AFP molecule about its helical axis, assuming that errors in these parameters would be least critical. Since side chains were absent from the search model and translation and rotation are coupled in an  $\alpha$ -helical structure, we reasoned that errors in the rotation parameters would result in small errors in translations that could be corrected later during structure refinement.

Considering the rotation and translation results of the one molecule searches (i.e. the helix tilts and the translation diagonal ridges), three families of packing arrangements are possible for two rod like molecules (length = 55Å, diameter=8Å) in the experimental unit cell. These are shown in figure 2.5. In one case, the two AFP molecules in the asymmetric unit lie on the ridges containing the crystallographic two fold axes and in two cases the AFP molecules lie on the ridges between the axes. A parallel versus criss-crossing alignment of AFP molecules on a sheet differentiates the latter two packing families. The possible helix direction combinations for the two

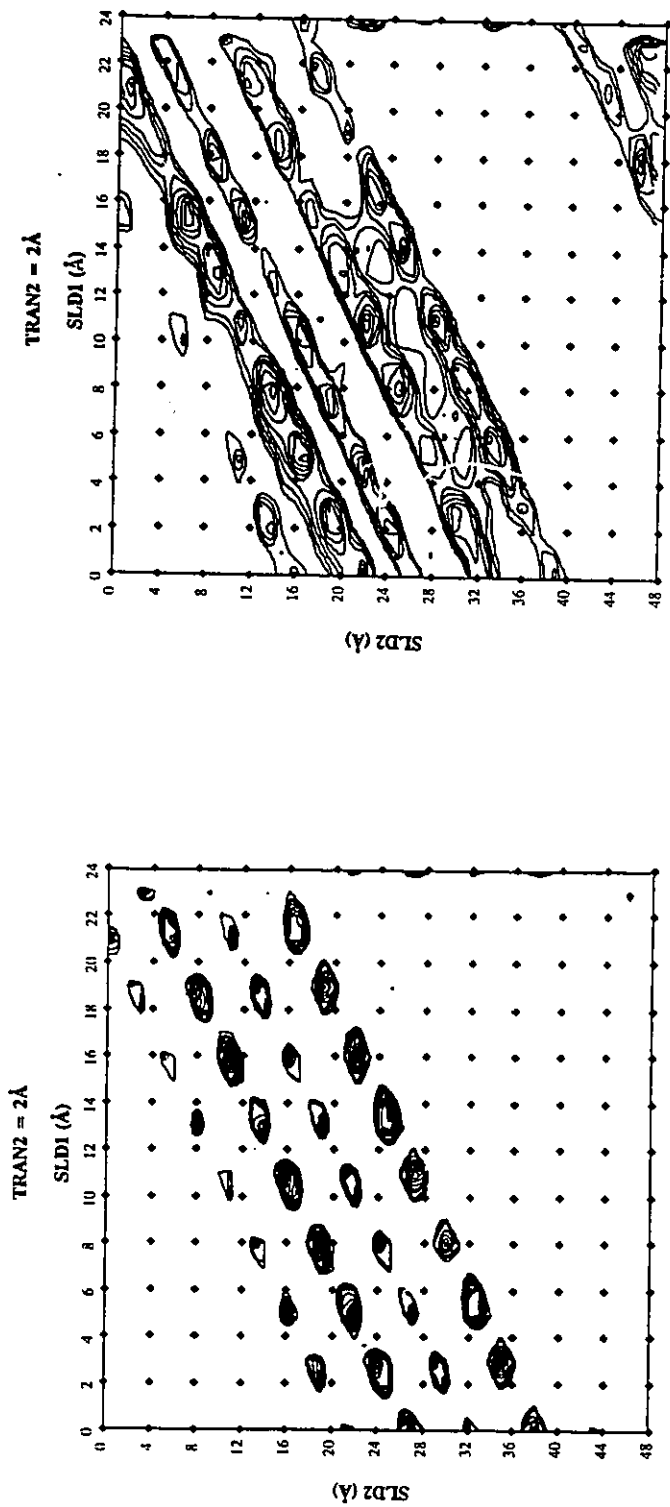


**Figure 2.5** Possible AFP packing arrangements consistent with the single molecule molecular replacement results. In arrangement A, AFP molecules are positioned on the diagonal ridges containing the crystallographic two fold axes. In arrangement B and C, AFP molecules are positioned on the diagonal ridges between the crystallographic two fold axes. As shown, a parallel versus a criss-crossing alignment of AFP molecules differentiates the latter two packing families

molecules in the asymmetric unit are up up, up down, and down down. Combination of the three packing arrangements and the three helix direction combinations leads to nine distinct subfamilies.

For each of the nine subfamilies, a search was performed against the *R-factor* and PC target functions (15-3Å). A representative example of the *XPLOR* shell script employed is given in Appendix A. The two poly-alanine search models were first rotated to orient them with desired tilts and directions. The parameters then searched included the diagonal slide of each search model along the XZ plane and the translation of one molecule along the Y axis. 1Å translation increments were used. The result of each PC search was a featurefull three dimensional map (i.e. a search map) from which a clear solution was not apparent (fig. 2.6). Using the  $R_{\text{factor}}$  target function rather than the PC function gave qualitatively similar maps but map features were much less pronounced (fig 2.6). Because a clear solution was still not apparent, a large number of the multi-parameter solutions were taken and refined against  $R_{\text{factor}}$ . It was hoped that this would help to differentiate which solution(s) was correct. Since the 1Å translation increment gave rise to obviously related solutions, the 3D map of each subfamily was *clustered* using a peak search algorithm modified from the *PHASES* package.

The top 180 clustered solutions from each of the 9 subfamilies were taken and committed to *refinement*. Refinement included 140 cycles of *rigid body* minimization, 40 cycles of *positional refinement* and 1 cycle of *simulated annealing* (SA). A

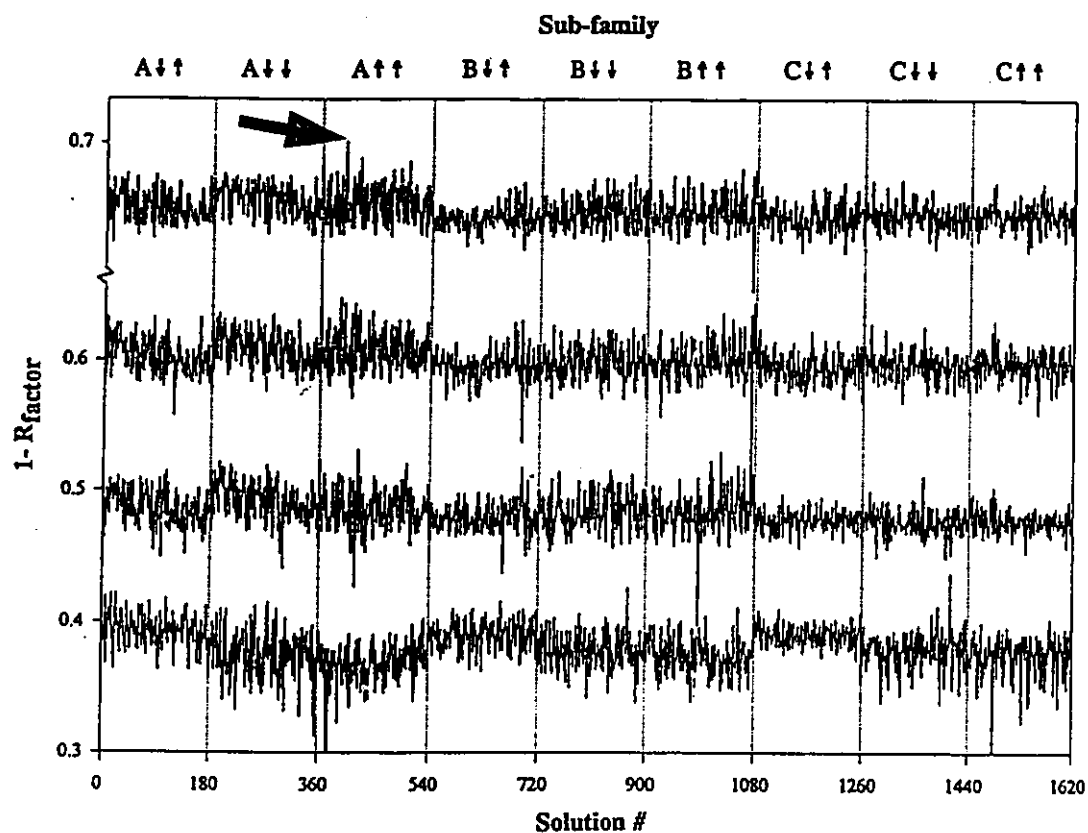


**Figure 2.6** Representative contour plot of the multi-parameter PC (left) and  $R_{\text{factor}}$  (right) search results. Parameters searched include the diagonal slide of AFP molecule one (SLD1), diagonal slide of molecule two (SLD2), and the translation of molecule 2 along the Y axis (TRAN2). Distances searched were 24Å, 48Å and 37Å respectively, using 1Å increments. Plots contoured at 70 to 100% of total PC or  $R_{\text{factor}}$  variation with increments of 5%.

representative example of the *XPLOR* shell script employed for refinement purposes is given in Appendix B. Results and statistics after each stage are shown in figure 2.7 and table 2.3. After the SA stage, one solution from subfamily 3, was clearly distinguishable with an  $R_{\text{factor}}$  of 30.2%. In comparison, the best solution from all other subfamilies possessed an  $R_{\text{factor}}$  of 32.3%. Initially and after rigid body and positional refinement stages, the top SA solution appeared 139th, 11th and second overall respectively. At all these stages, however, the solution was not significantly better than a large number of other solutions. Visual inspection of the top solution before and after SA refinement showed that one of the AFP molecules remained straight while the other became significantly curved. We were not sure if the difference in curvature between the two AFP molecules was real or represented an artefact of the SA protocol. Visual inspection of other MR solutions, also with reasonable refinement statistics, showed that in a number of cases, the AFP molecules also became curved after SA refinement. However, significant differences in the center of mass positions of the AFP molecules and the positions of the curvature in the AFP molecules existed.

At this time, two related isomorphous derivatives were found (table 2.1). Representative *Harker planes* through difference Patterson maps calculated for the  $\text{TbNO}_3$ , crystal #2 derivative are shown in figure 2.8. The presence of two heavy atom sites (major site = 0.37, 0.17, 0.55; minor site = 0.06, 0.00, .18) are clearly evident in the isomorphous difference Patterson map while only the major site is





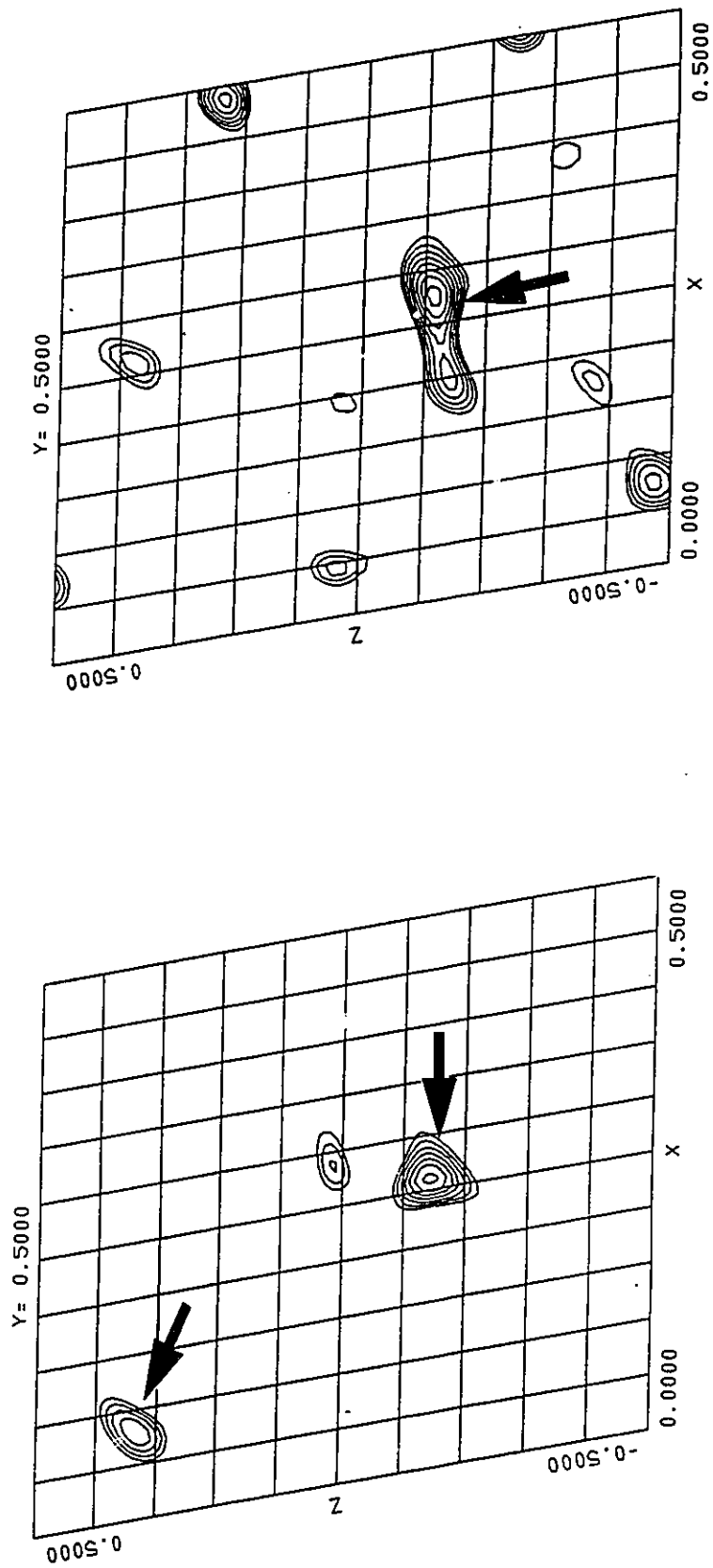
**Figure 2.7** Refinement results of the top clustered subfamily solutions. From bottom to top respectively, plots of  $1-R_{\text{factor}}$  versus solution number at stages before refinement (START), after 140 cycles of rigid body refinement (RIGID), after 40 cycles of positional refinement (POS) and after simulated annealing ((SA) 3000°C to 300°C followed by 40 cycles positional refinement). A resolution range of 3.0-15Å with a  $2\sigma$  cut off was employed. Subfamilies are denoted by their packing family A, B or C (see figure 2.5) and the helix directions  $\uparrow\uparrow$ ,  $\uparrow\downarrow$  or  $\downarrow\downarrow$ .  $R_{\text{factor}}$  statistics are given in table 2.3.

**Table 2.3**  $R_{\text{factor}}$  statistics following various stages of subfamily solution refinement.

---

		SUBFAMILY								
		A↓↑	A↓↓	A↑↑	B↓↑	B↓↓	B↑↑	C↓↑	C↓↓	C↑↑
<b>START</b>	<b>AVE</b>	0.607	0.626	0.631	0.609	0.622	0.625	0.609	0.62	0.624
	<b>SD</b>	0.011	0.017	0.016	0.008	0.013	0.013	0.01	0.013	0.015
	<b>BEST</b>	0.578	0.584	0.597	0.59	0.573	0.588	0.584	0.562	0.597
<b>RIGID</b>	<b>AVE</b>	0.514	0.508	0.515	0.522	0.516	0.516	0.521	0.523	0.523
	<b>SD</b>	0.013	0.014	0.015	0.007	0.013	0.015	0.01	0.009	0.008
	<b>BEST</b>	0.485	0.478	0.468	0.502	0.484	0.47	0.482	0.488	0.497
<b>POS</b>	<b>AVE</b>	0.397	0.391	0.392	0.405	0.402	0.4	0.404	0.403	0.402
	<b>SD</b>	0.013	0.013	0.014	0.01	0.012	0.015	0.011	0.009	0.01
	<b>BEST</b>	0.366	0.363	0.351	0.375	0.373	0.356	0.369	0.372	0.369
<b>SA</b>	<b>AVE</b>	0.353	0.347	0.349	0.361	0.358	0.357	0.361	0.359	0.359
	<b>SD</b>	0.011	0.011	0.014	0.01	0.011	0.013	0.008	0.008	0.008
	<b>BEST</b>	0.327	0.327	0.302	0.333	0.328	0.323	0.334	0.332	0.333

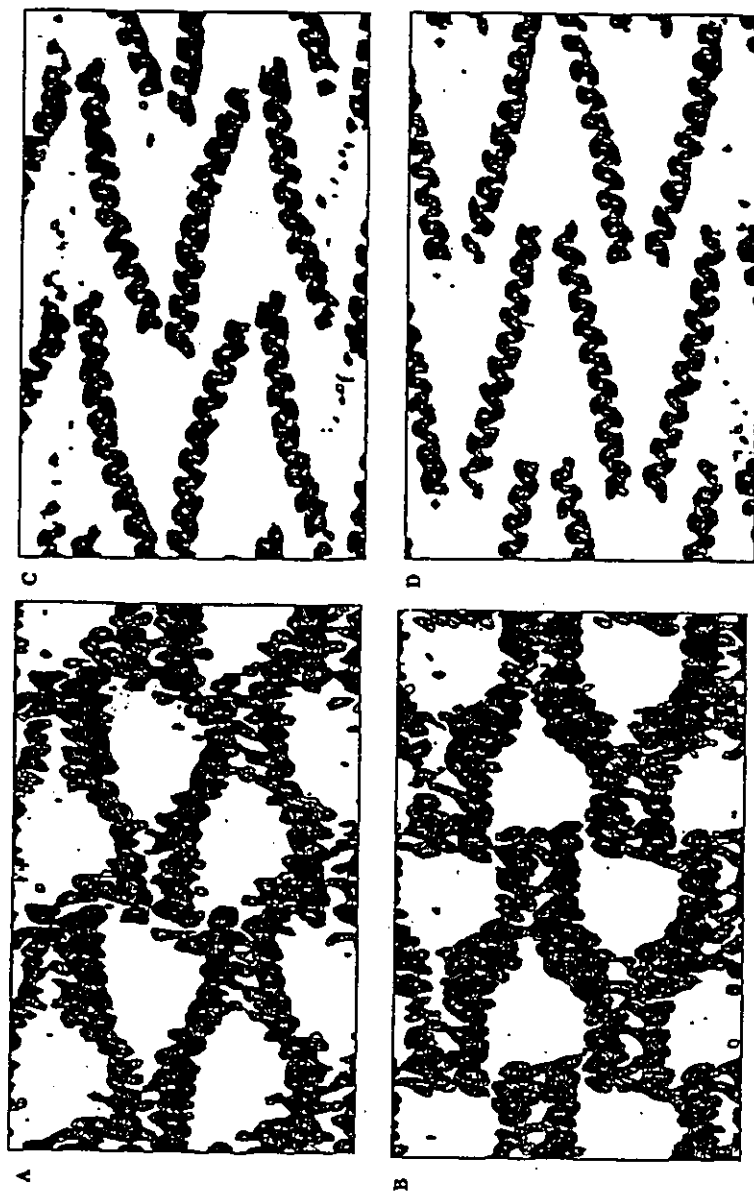
---



**Figure 2.8** Harker planes through anomalous (right) and isomorphous (left) difference Patterson maps derived from the  $\text{TbNO}_3/2$  heavy atom derivative. Resolution = 2.5-15Å. Heavy atom sites are labeled with arrows.

evident in the noisier anomalous difference Patterson map. The noisy character of the latter map is due in part to a relatively weak anomalous signal and also to limitations in the diffraction measurement conditions. Because of a plate-like protein crystal morphology (the unique crystal axis is perpendicular to the large plate surface) it was impossible to orient crystals within a capillary tube to allow concurrent collection of *Friedel mates*. Heavy atom positions consistent with Harker plane features were also identified for each derivative using the program *HASSP* (Terwilliger, Kim, & Eisenberg, 1987). MIRAS phases were calculated, refined and combined using the *PHASES* package. This was followed by 16 cycles of *solvent flattening* and then *phase extension* as described by Wang (1985) and implemented in the *PHASES* package. An underestimate of 40% solvent was found to be most effective for solvent flattening purposes. An *electron density* map ( $FOM * F_o$ ) which clearly demonstrated the gross packing arrangement of the AFP molecules was produced (fig. 2.9). The approximate centre of mass positions of the two AFP molecules was in agreement with the top molecular replacement solution but side chain features and helix directions could not be resolved. Before solvent flattening and phase extension, MIRAS electron density maps were not interpretable.

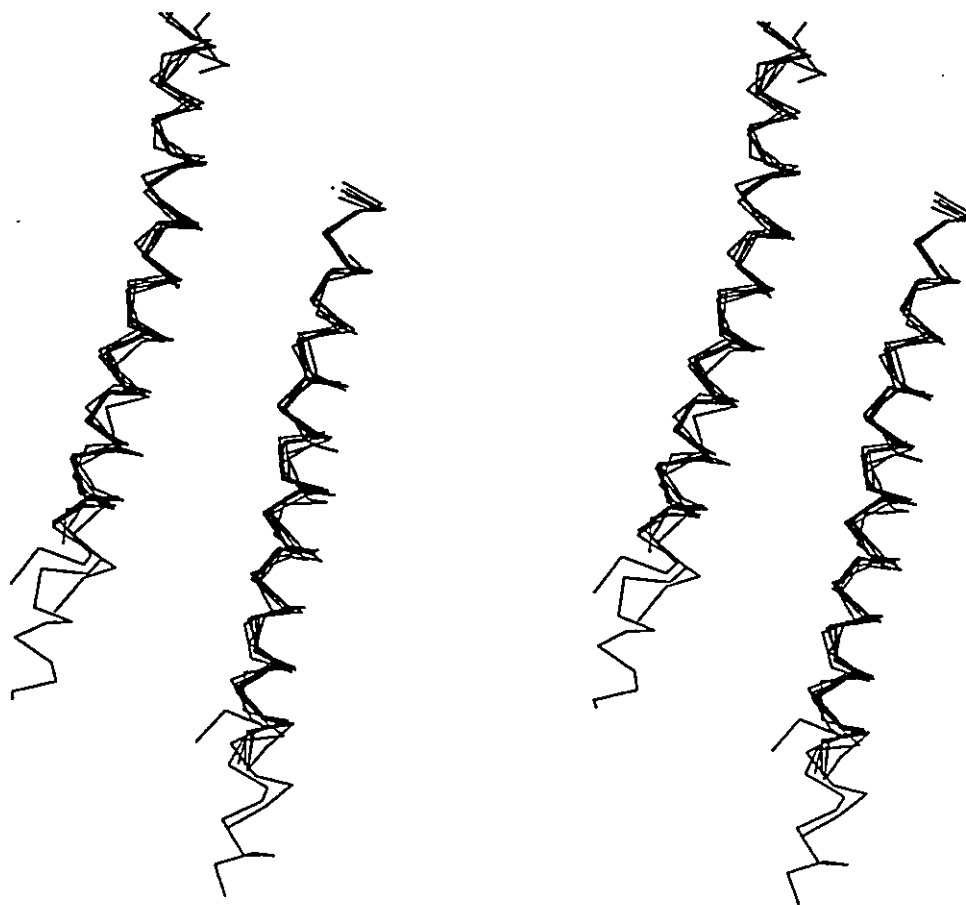
In addition to confirming the gross packing arrangements of the  $\alpha$ -helical molecules in the unit cell, heavy atom derivatives proved useful in providing an alternate means of filtering the large number of MR solutions that possessed reasonable refinement statistics after SA refinement. The top 10 solutions of each subfamily after



**Figure 2.9** Cross sections through a MIRAS derived, solvent flattened, phase extended electron density map ( $FOM \cdot F_o$ ) showing AFP molecule one and two (panels A & B respectively). For comparison, panels C & D show the same cross sections through an electron density map ( $F_{calc}$ ) of the top MR solution after SA refinement. The unique crystal axis Y is vertical and the diagonal axis Z=-X lies horizontal. The origins of MIRAS and MR maps are correlated for direct comparisons. Curvature distortions in the MIRAS map arise from errors in the automatic boundary determination during solvent flattening and do not accurately reflect the curvature of the AFP molecules. Using the top MR solution for boundary calculation produced much improved maps. Resolution for all maps is 3-15Å.

SA refinement were used to phase the structure factor amplitudes of one of the  $\text{LaNO}_3$  derivatives. The solutions that revealed the heavy atom positions as major peaks in 2  $|F_o| - |F_c|$  maps (15-3Å) were *origin correlated* and viewed. Five of ten solutions from subfamily 3 identified the heavy atom positions and superimposed very well, including the position of the backbone curvature in one of the two AFP molecules, as did one solution from each of subfamilies 1, 6 and 9 (fig. 2.10). The latter two cases were surprising since the poly-alanine search molecules originated on different translation ridges (search models migrated away from their starting ridges during the rigid body refinement stage). In all but the subfamily 1 solution, the poly-alanine search models were oriented with the same directions (up up).

The multi-parameter search/refine results, the MIRAS derived electron density map and the filtering of top multi-parameter search/refine solutions using the heavy atom derivatives gave us confidence in the approximate positioning of the two AFP molecules, the synchronisation of back bone atoms, the backbone curvature in one AFP molecule, and the direction of the AFP molecules. The precise translation/rotation (the exact register) of each AFP molecule was still in doubt since the rotations of the molecules about their helical axes were not sampled in a search. Threonine side chains are known to strongly favour a single conformation on  $\alpha$ -helical structures (Piela, Nemethy & Scheraga, 1987; McGregor, Islam & Sternberg, 1987) and since four are present in the flounder sequence (fig. 1.3), we were able to use this to resolve the



**Figure 2.10** Stereo plot of the origin correlated molecular replacement solutions. Only the displayed solutions provided sufficiently accurate phase information to identify heavy atom positions in  $(\text{TbNO}_3/2)$  difference Fourier maps.

remaining ambiguity. The four Thr side chains were added to seven different positions on each AFP molecule in the top search/refine molecular replacement solution (Thr placements corresponded to different helix registers). After a cycle of SA refinement (10-2Å, 3000-300°C), the correct Thr placement on each AFP molecule was identified by  $R_{factor}$  (table 2.4). Confirming the correctness of the Thr placements, most of the remaining side chains of the AFP molecules were clearly apparent in  $2|Fo|-|Fc|$  electron density maps.

## ii. Refinement

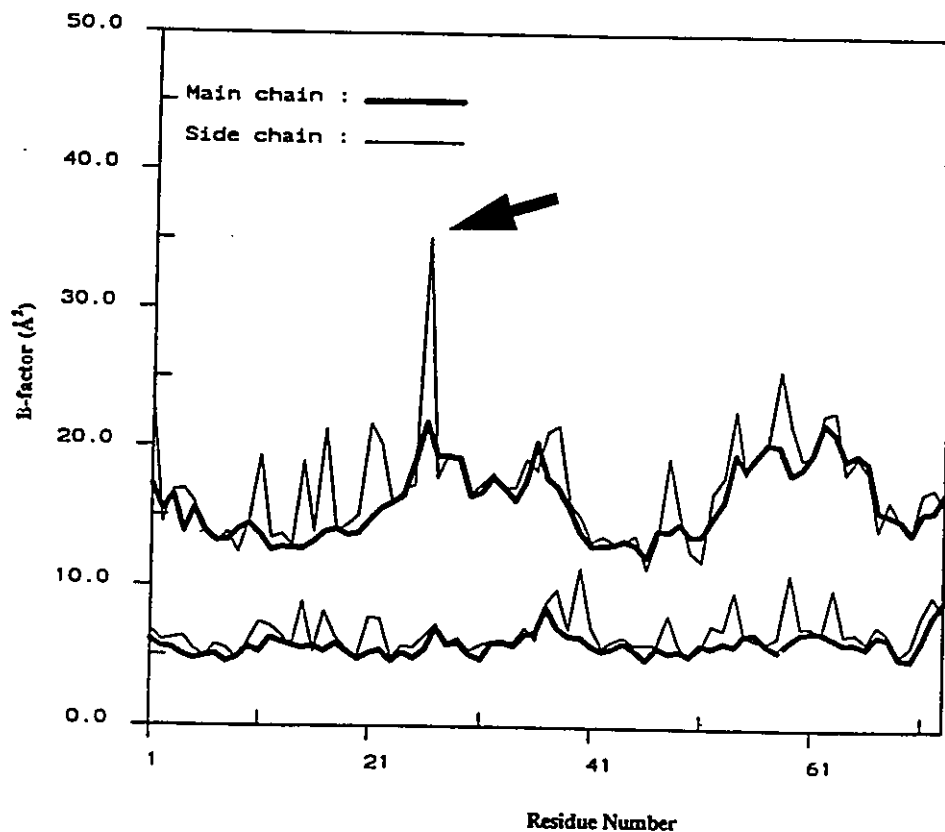
At this stage, 10% of the structure factor amplitudes were partitioned for the analysis of  $R_{free}$  (Brunger, 1992). After the addition of all side chains to the two polyalanine+threonine models, cycles of SA refinement and model rebuilding using O (Jones, Zou, Cowan & Kjeldgaard, 1991) were performed. SA omit maps were used to confirm all regions of the structure and all but the side chain of Asn 27 in one AFP molecule appeared ordered. Resolution was then increased from 2.0Å to 1.7Å using cycles of SA refinement. Solvent was iteratively picked using the automated water selection program *ASIR* (Tong, Berghuis & Luo, 1994) modified for use with *XPLOR*. Only water molecules forming more than one hydrogen bond were selected and a B-factor cut-off of 55 was employed. 39 water molecules were included before  $R_{free}$  started to increase. A *restrained isotropic B-factor refinement* was performed on the final structure. The protein structure at 4°C without solvent was then used as the



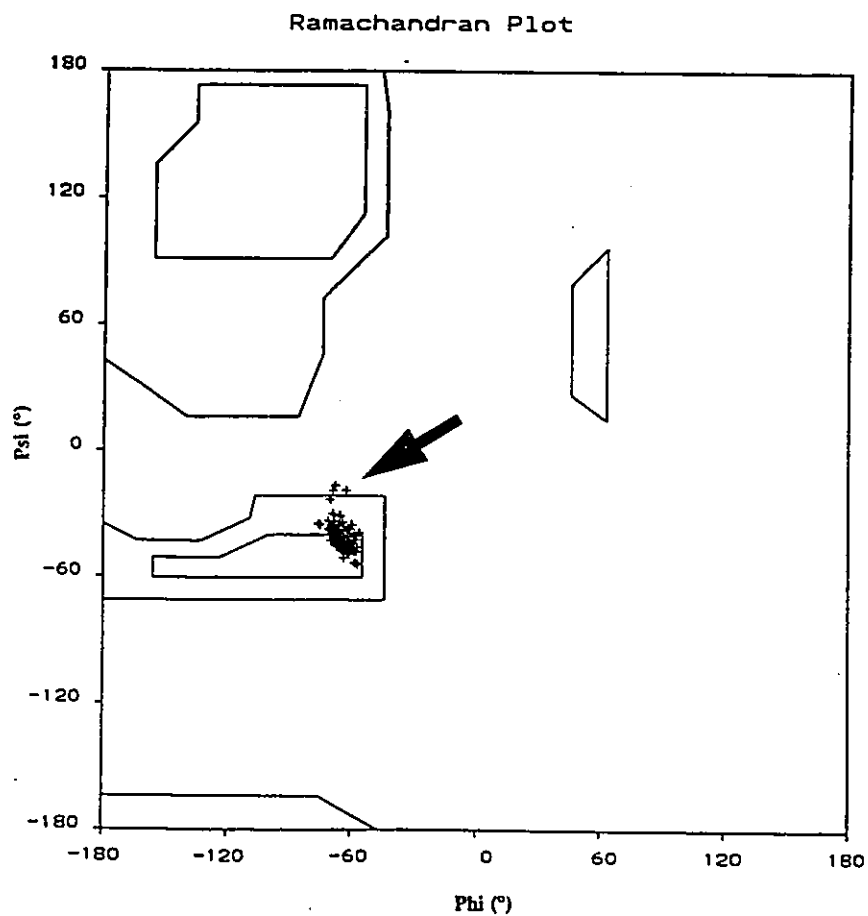
**Table 2.4 Identification of AFP rotations about their helix axis by the crystallographic refinement of systematically placed Thr residues.**

<b>Thr placement #</b>	<b><math>R_{\text{factor}}</math></b>
	mol #1
1	<b>30.10</b>
2	30.24
3	30.43
4	30.55
5	30.46
6	30.81
7	30.90
	mol #2
1	<b>29.41</b>
2	30.43
3	30.60
4	30.37
5	30.59
6	31.43
7	31.02

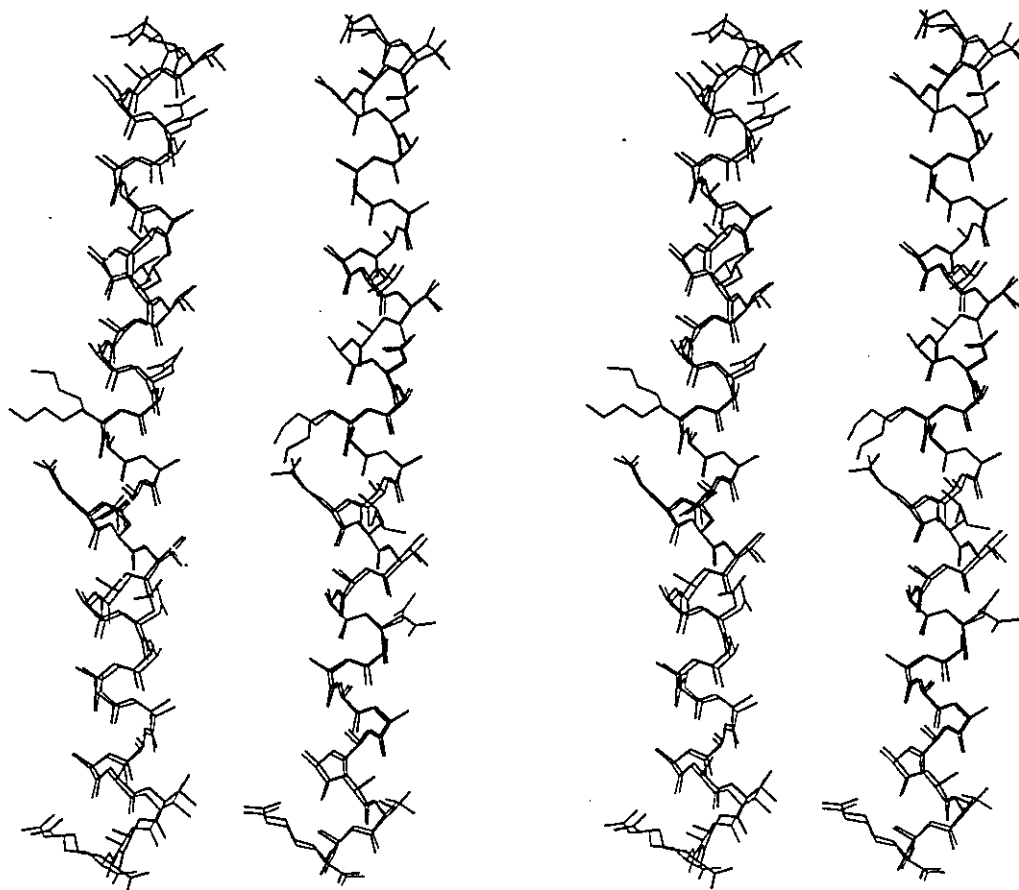
starting model for the -180°C structure. After cycles of rigid body and SA refinement, SA omit maps were used to correct problem regions of the structure. Resolution was extended to 1.5Å and solvent was picked as described for the 4°C structure. 130 water molecules were included before  $R_{\text{free}}$  started to increase. Well defined solvent molecules other than water were not apparent. Both AFP molecules in the -180°C crystal structure were completely ordered. A summary of the refinement statistics are given in Table 2.2. B-factor and Ramachandran plots for the 4°C and -180°C structures are shown in figure 2.11 and 2.12 respectively. All four AFP molecules are completely  $\alpha$ -helical with the exception of the last peptide unit in each which adopts a  $3_{10}$  conformation. The structure of the four AFP molecules at 4°C and -180°C are shown superimposed in figure 2.13. *Luzzati plots* (Luzzati, 1952) providing estimates of coordinate errors for the two crystal structures are shown in figure 2.14. Representative *SA 'omit' electron density maps* showing the only disordered side chain in the 4°C structure and the same residue in the -180°C structure are shown in figure 2.15. Additional electron density maps of an entire AFP molecule and part of the AFP's ice binding structure are shown in figures 3.2 and 3.4 respectively.



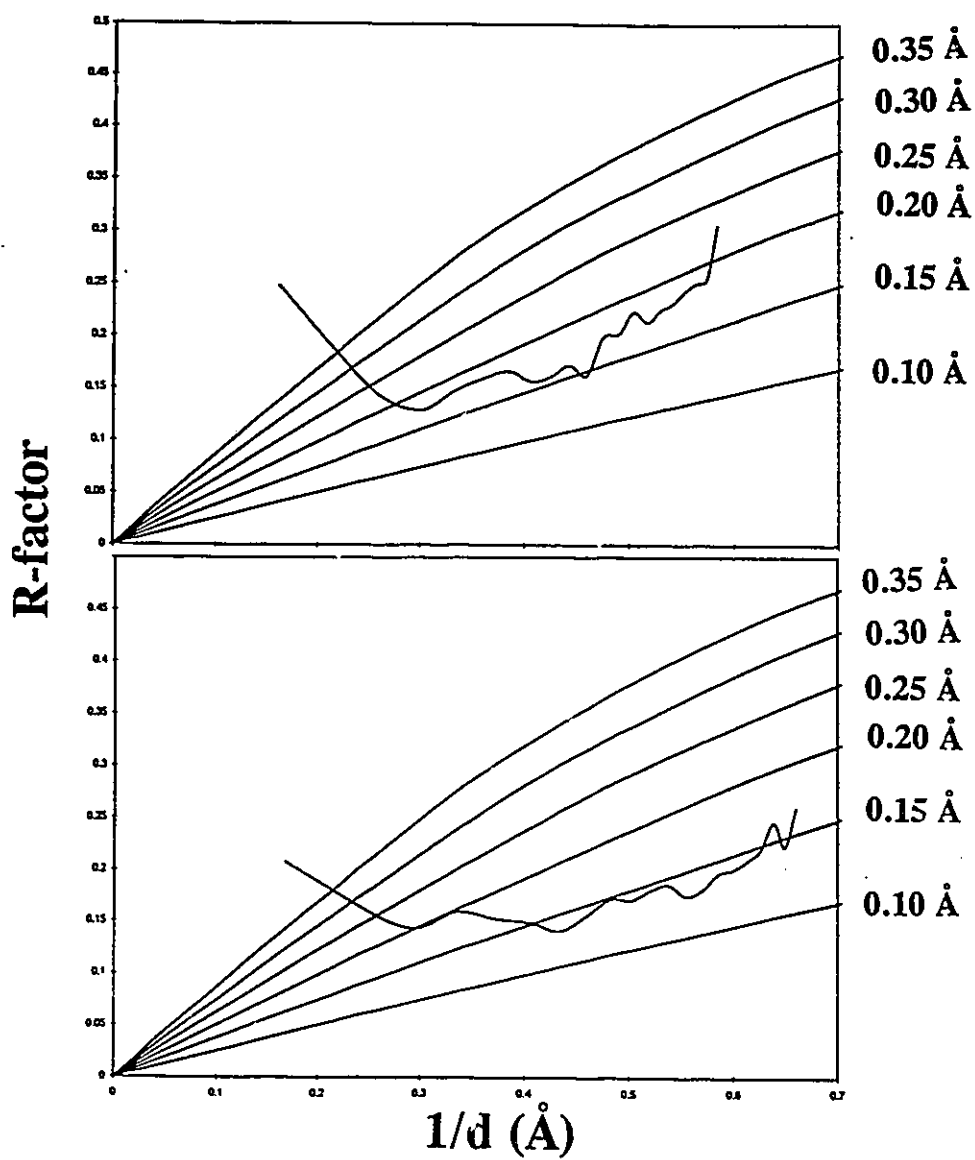
**Figure 2.11** Residue averaged B-factor plots for the 4°C (upper traces) and -180°C (lower traces) winter flounder AFP crystal structures. Molecule one and two correspond to residues 1-37 and 38-72, respectively. Average solvent B-factors for the 4°C and -180°C structures are 43.0 Å<sup>2</sup> and 33.0 Å<sup>2</sup>, respectively. Arrow denotes Asn 27 which is found disordered.



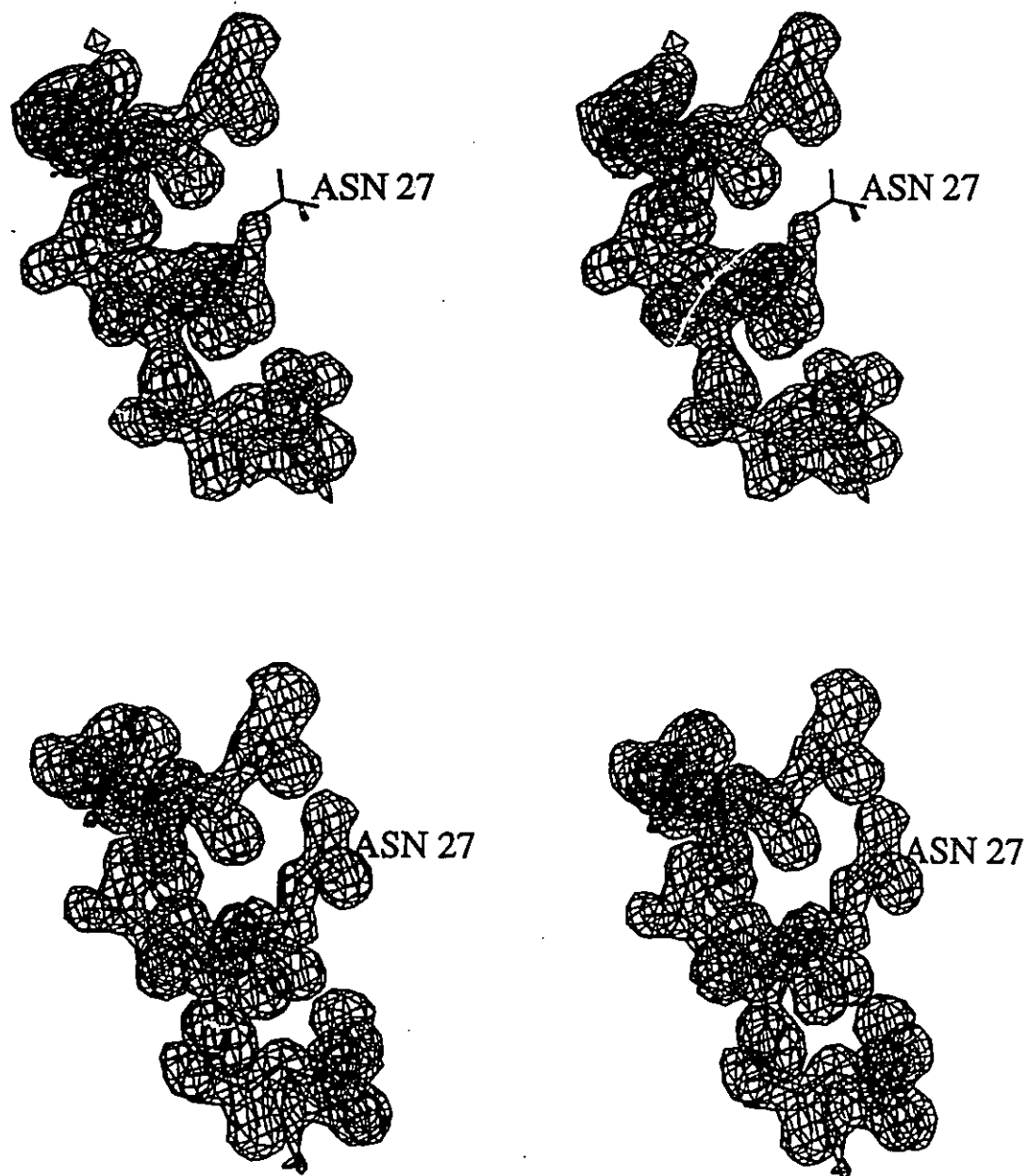
**Figure 2.12** Ramachandran plot for the four AFP molecules of the 4° and -180°C crystal structures. The four outlying points shown by an arrow correspond to the 3/10 conformations which terminate each AFP molecule.



**Figure 2.13** Stereo plot of the superimposed winter flounder AFP structures. Red and blue colouring denote the 4° and -180°C structures, respectively. AFP molecules one are shown on the left. Molecules are aligned with amino termini up. Average R.M.S. difference between superimposed and opposing backbone structures are 0.26Å and 1.52Å respectively.



**Figure 2.14** Luzzati plots for the estimate of overall coordinate error; top, 4°C crystal structure (0.18Å error); bottom -180°C crystal structure (0.15Å error).



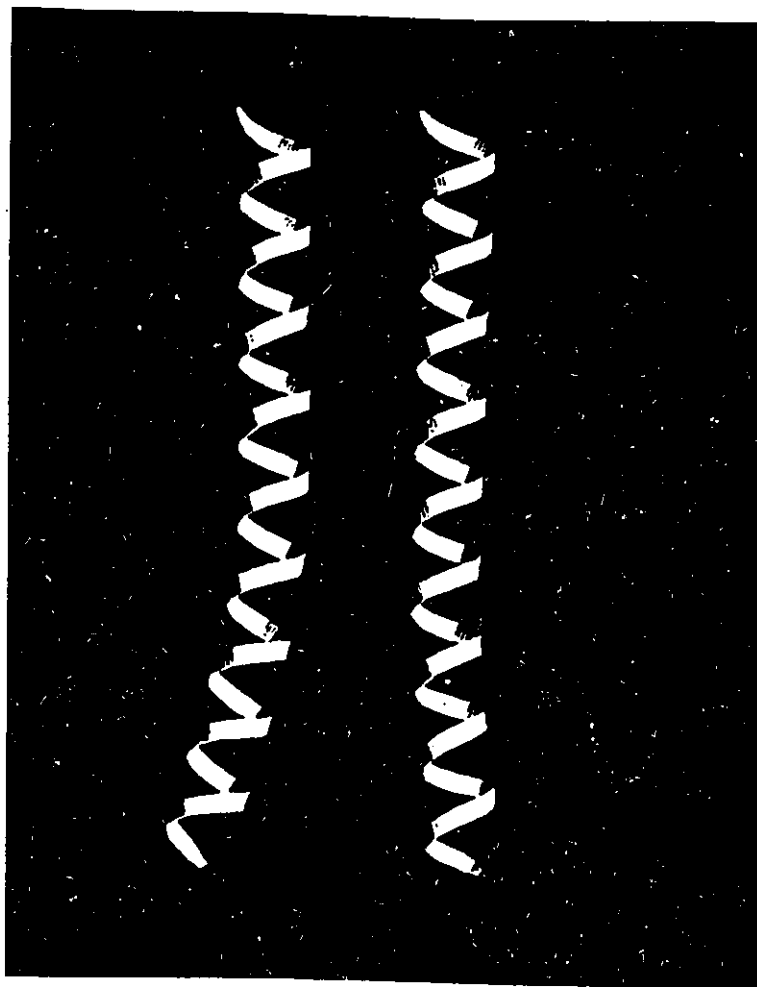
**Figure 2.15** Representative simulated annealing 'omit' electron density maps showing Asn 27 of AFP molecule 1 @ 4°C (top) and the same residue at @ -180°C (bottom). Asn 27 of AFP molecule 1 @ 4°C is the only disordered side chain.

## CHAPTER 3 STRUCTURE ANALYSIS

As reflected by the difference in unit cell dimensions (table 2.2) and the superimposed structures in figure 2.13, significant non-isomorphism is observed between the 4°C and -180°C crystal structures. This provides four distinct AFP molecules with which to discern biologically relevant structure from crystal imposed artifacts. This is useful since all of the AFP's side chain structure is solvent exposed and thus can potentially be affected by crystal packing interactions. Although differences between the 4°C and -180°C crystal structures exist, the structures are overall very similar. For this reason, we focus mainly on the -180°C structure.

Both AFP molecules in the asymmetric unit at both temperatures are completely  $\alpha$ -helical with the exception of the last peptide unit in each which adopts a  $3_{10}$  conformation. Figure 3.1 shows that AFP molecule 2 is straight while molecule 1 is gently bowed. This difference in curvature has no discernible effect on backbone hydrogen bond lengths (molecule 1, ave= $2.92 \pm 0.07 \text{ \AA}$ , max= $3.09 \text{ \AA}$  min= $2.78 \text{ \AA}$ ; molecule 2, ave= $2.94 \pm 0.07 \text{ \AA}$ , max= $3.12 \text{ \AA}$ , min= $2.76 \text{ \AA}$ ) and likely reflects the flexibility of the protein rather than a disposition towards bending. Reflecting a form that follows function, all 14 non-alanine side chains present in each AFP molecule appear to serve roles in either ice binding or helix stabilization.

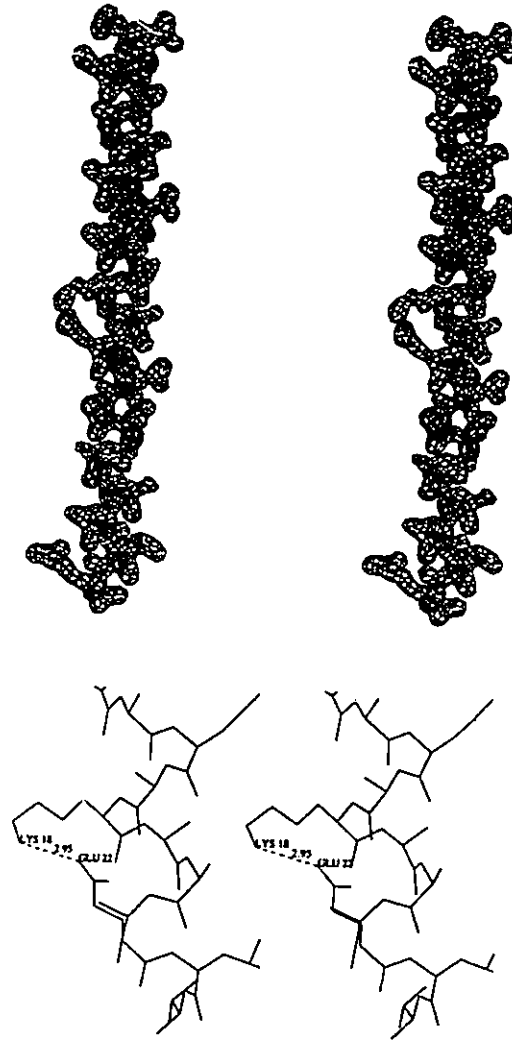




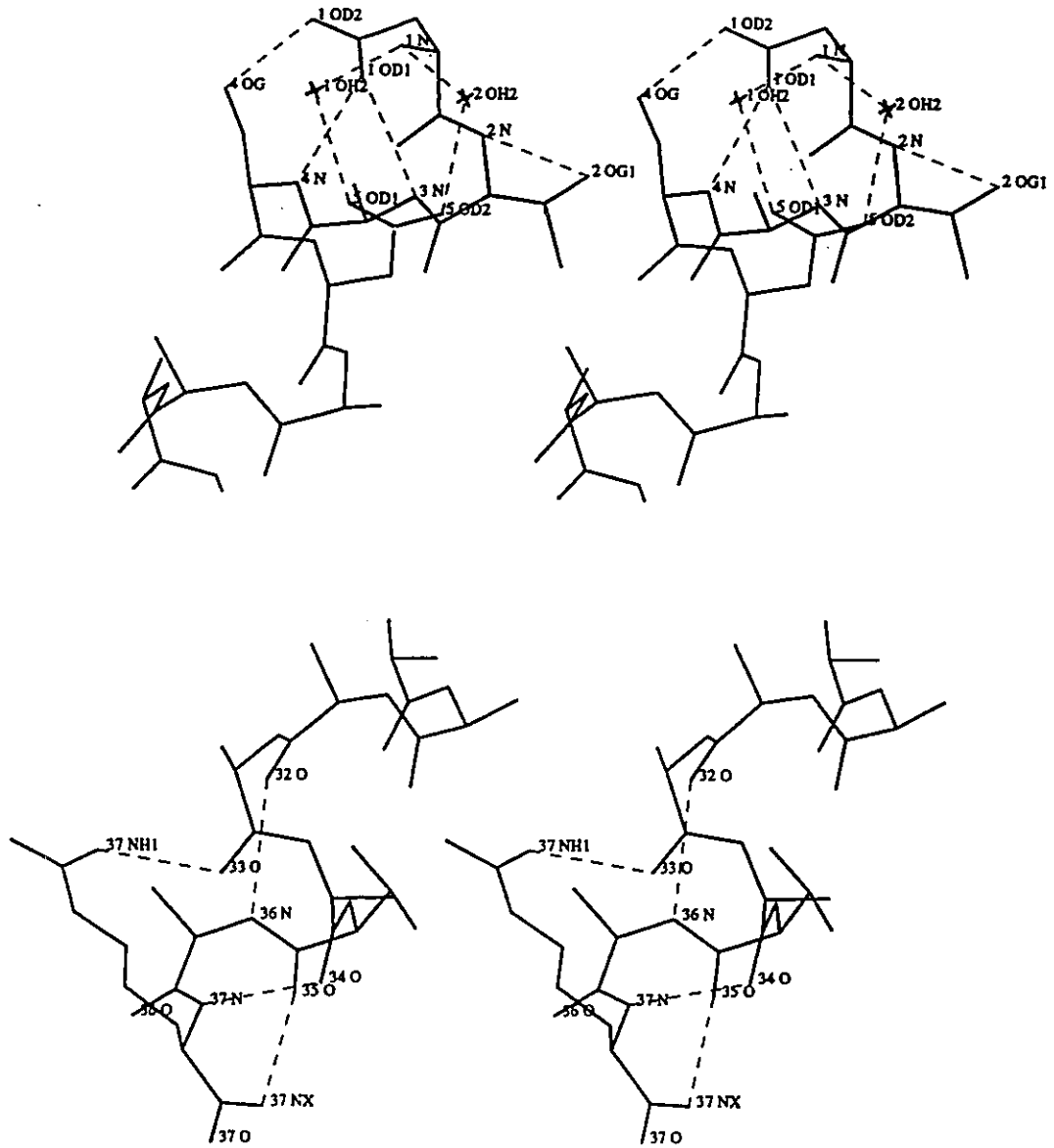
**Figure 3.1** Ribbon diagrams of AFP molecules 1 and 2 @ 4°C, from left to right respectively, demonstrate the flexibility of the protein's backbone structure. AFP molecules are aligned with amino termini up and ice binding surfaces facing left.

### 3.1 Helix stabilization

Helix stability has been shown by mutagenesis experiments to have a direct effect on the AFP's ability to inhibit ice crystal growth (Chakrabartty & Hew, 1991; Wen & Laursen, 1992; Jorgenson; Mori, Matsui, Kanaoka, Yanagi, Yabusaki & Kikuzono, 1993). As lone  $\alpha$ -helices are inherently unstable in solution (Chakrabartty & Baldwin, 1995), and the AFP employs ice binding residues (Thr and Asn) which are themselves destabilizing (Chou & Fasman, 1974) it is curious that the winter flounder AFPs are remarkably stable with estimated helical contents approaching unity at 4°C (Ananthanarayanan & Hew, 1977; Chakrabartty, Ananthanarayanan and Hew, 1989; Wen & Laursen, 1992). Part of this stability has been attributed to the protein's rich alanine content (62.8 mole%) (Ananthanarayanan & Hew, 1977; Chakrabartty, Ananthanarayanan and Hew, 1989) and an intra-helical salt bridge between Lys 18 and Glu 22 (Chakrabartty, Ananthanarayanan & Hew, 1989; Chakrabartty & Hew, 1991). In the crystal structure, the intra-helical salt bridge is confirmed (fig 3.2), and in addition, the presence of elaborate N- and C- terminal cap structures (Richardson & Richardson, 1988; Presta & Rose, 1988) are revealed (fig. 3.3). Since the fraying of helix termini is one of the greatest sources of instability for a lone  $\alpha$ -helical structure (Chakrabartty & Baldwin, 1995), and the AFP's ice binding structure extends over all of its 57Å length, the cap structures are likely to be the AFP's most strategic helix stabilizing elements.



**Figure 3.2** Stereo plots of AFP molecule 2 @ 4°C showing the helix stabilizing salt bridge between Lys 18 and Glu 22. In the  $2|F_o| - |F_c|$  electron density map (top), the interacting side chains are visible mid way along the left side of the AFP molecule. A close up of the salt interaction in the bottom figure.



**Figure 3.3** Stereo plots of the N- (top) and C- (bottom) terminal cap structures. The helix axis, N- to C-terminus, lies on the plane of the page from the top right to bottom left. Hydrogen bonds are represented by thin dashed lines. The two N-terminal cap bound water molecules are labeled 1 OH<sub>2</sub> and 2 OH<sub>2</sub>.

The N-terminal cap structure consists of an ordered network of eight hydrogen bonds involving the side chains of residues Asp 1, Thr 2, Ser 4, Asp 5 and two tightly bound water molecules. The hydrogen bonding network effectively caps all four local backbone groups that lack intra-helical hydrogen bonding partners. The apical positioning of the Asp 1 side chain within the cap structure also allows for a stabilizing interaction with the helix dipole, while the close proximity of Asp 5 to the free amino terminus may offset the latter's destabilizing charge.

The C-terminal cap structure, characterized by a C $\alpha$ 1 helix distortion (Baker & Hubbard, 1984) makes use of the side chain of Arg 37 and the amidated carboxy-terminus to form three hydrogen bonds. The structure leaves uncapped only one of four local backbone sites. The Arg side chain forms one capping hydrogen bond and is well positioned to interact favorably with the helix dipole. Amidation of the C-terminus eliminates a destabilizing charge interaction with the helix dipole and replaces a hydrogen bonding acceptor group with a donor group that serves to cap the carbonyl group of Ala 35.

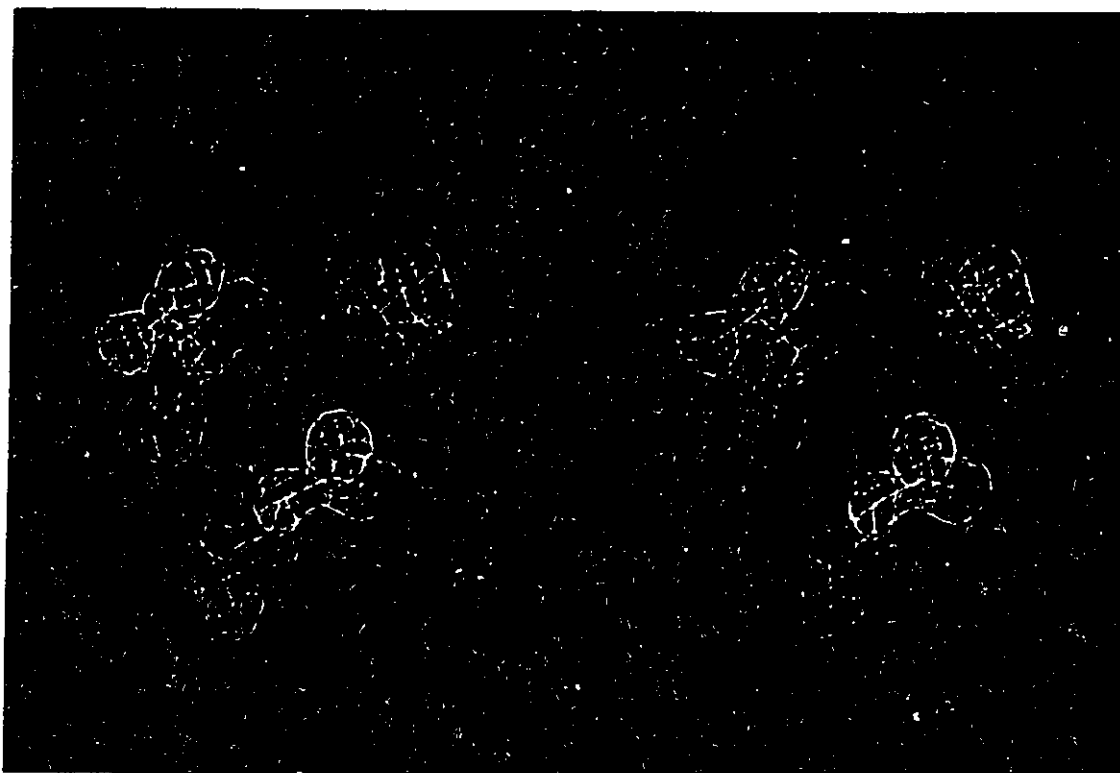
A comparison of primary sequences reveals that the N-cap sequence observed in the winter flounder AFP (Asp Thr Ala Ser Asp) is identical to those observed in a number of other  $\alpha$ -helical AFPs (figure 1.3). In addition, two other conserved N-terminal sequences are apparent in the sculpin AFPs which may represent other interesting helix stabilizing cap structures. X-ray and NMR structural studies on these

N-terminal sequences have been initiated in Dr. Daniel Yang's laboratory to determine if they do in fact form alternative capping structures (chapter 5.1).

### 3.2 Ice binding structure

The AFP's ice binding structure is composed of the repeat of four similar ice binding motifs, each of which is made up of the residues Thr/Asp or Thr/Asn/Leu (Fig. 1.3). The last IBM is incomplete, having only a Thr residue. Figures 3.4 and 3.5 show a front view of an ice binding motif and an axial view of the entire ice binding surface, respectively. Unexpectedly, we find that IBM residues are rigidly constrained to conformations that produce an exceptionally flat ice binding surface.

The  $\beta$ -branched Thr side chains, are sterically constrained ( $\chi_1 = -61 \pm 4^\circ$ ) by the  $\alpha$ -helical conformation of the protein's backbone structure (Piela, Nemethy & Scheraga 1987; McGregor, Islam, & Sternberg, 1987). The side chains of Asp 5, Asn 16 and 27 adopt the same conformation ( $\chi_1 = -75 \pm 3^\circ$   $\chi_2 = 13 \pm 13^\circ$ ) yet are restrained by different means. Asp 5 is held fixed by the aforementioned capping interactions, while Asn 16 and 27 form weak hydrogen bonds to the helix backbone (Fig. 3.4). Close van der Waals contacts between Leu and Asn side chains suggest that the Leu residues may play a role in stabilizing the Asn conformations. This may account for the finding that substitution of both Leu residues of the central IBMs with Ala results in a 33% loss in protein function (Wen & Laursen, 1992a). Previously, modeling studies have led to the proposal that the Leu side chains function independently to contact ice



**Figure 3.4** Stereo  $2|F_o|-|F_c|$  electron density map of a winter flounder IBM. The helix axis (amino terminus up) lies vertical on the plane of the page. The map is contoured at  $1\sigma$ . A close contact between the Asn and Leu side chains ( $3.7\text{\AA}$ ) is evident. In the four unique crystal environments sampled by the AFP, 14/16, 10/12, and 3/8 instances of Thr, Asx and Leu IBM side chains, respectively, adopt the displayed conformations. In all other instances, the IBM side chains are affected by crystal packing interactions.

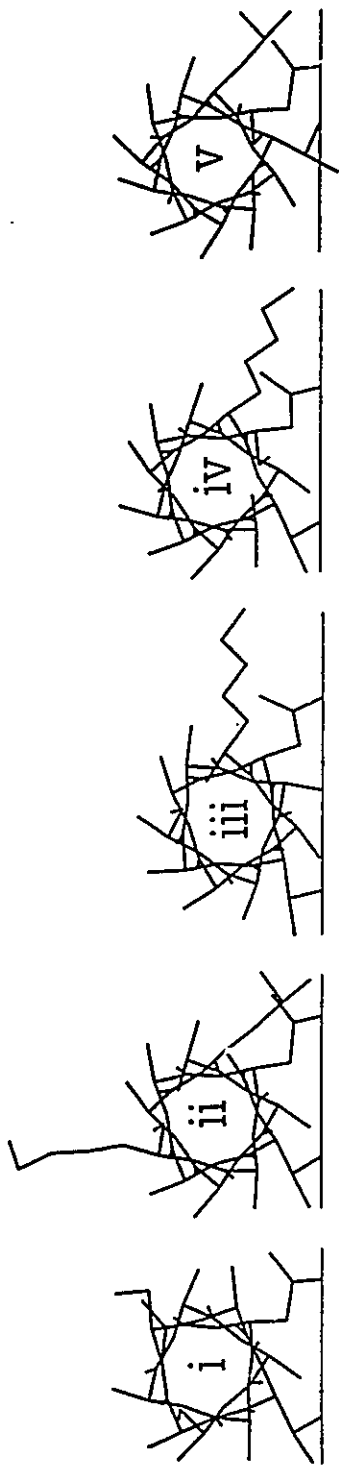


**Figure 3.5** N-terminal axial view of the AFP's ice binding surface highlighting its flat and accessible nature. The amino terminus, pointing towards the reader, is tilted down  $5^\circ$  from the horizon with the ice binding surface facing upwards. Carbon, oxygen and nitrogen atoms are coloured yellow, red and blue, respectively.



directly (Wen & Laursen 1992b). Comparison of the primary sequences of the other  $\alpha$ -helical AFPs shows that the pairing of IBM residues with  $i,i+4$  spacing is a highly conserved feature (Fig. 1.3). In most of the recognizable IBM sequences, Asp is found paired with Lys and unlike the interaction between Leu and Asn residues, the potential for interaction between Asp and Lys residues when spaced appropriately on an  $\alpha$ -helical structure is well characterized (Merutka & Stellwagen 1991; Marqusee & Baldwin, 1987). Further suggesting an inter-dependent function, a correlated absence of the paired residues is always observed in the incomplete IBMs of the  $\alpha$ -helical AFPs (Fig. 1.3).

The restrained conformation of IBM side chains has a great influence on the accessibility of the AFP's ice binding groups. As shown in the axial projection of ice binding surface in figure 3.5, the OH, NH<sub>2</sub>, and C=O groups of the Thr, Asn, and Asp side chains, respectively, protrude moderately from a flat local surface composed with neighboring apolar groups. An inspection of cross sections through the winter flounder IBMs and 'modeled' IBMs from the other  $\alpha$ -helical AFPs shows that all may form surfaces with similar characteristics (fig 3.6). While the radial separation between ice binding groups does vary from one residue in most of the AFPs to two residues in only the yellowtail flounder AFP, the accessibility of ice binding groups is not destroyed. Also influencing the accessibility of polar groups greatly is the conserved use of Ala in the positions located radially between the ice binding residues. In IBMs modeled with radial separations larger or smaller than those observed in the  $\alpha$ -helical



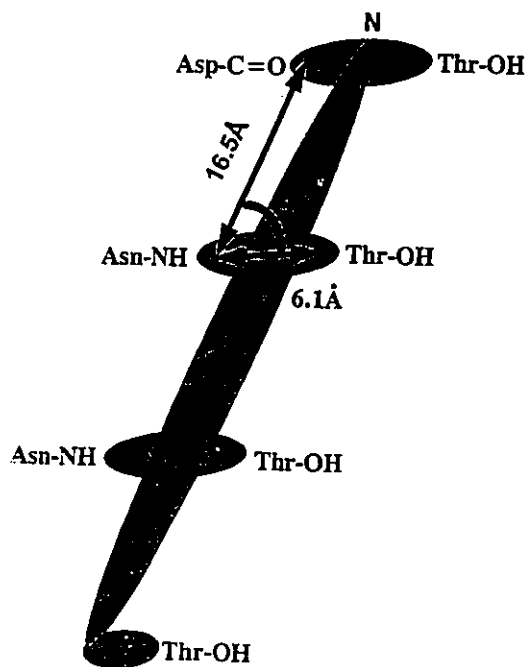
**Figure 3.6** Cross sections through the representative IBMs labeled in figure 1.3. IBMs *i* and *ii* are taken from the winter flounder crystal structure. In IBMs *iii* and *iv*, Asp side chains are modeled after the Asp side chain conformations observed in *i* and *ii* while the Lys side chains are modeled in an extended conformation. In *i-iv*, the hydrogen bonding groups of the ice binding residues are accessible to an external surface as indicated by the line drawn through the Thr OH and the Asn NH<sub>2</sub> or Asp C=O groups. In model IBMs in which the radial separation between Thr and Asx residues is made smaller as in *v* or larger than those observed in the alanine rich AFPs, or when the conserved alanine residue(s) located radially between Thr and Asx residues is substituted with a larger sized residue(s), the accessibility of the ice binding groups is greatly impaired.

AFPs (Fig. 3.6 v) or with larger sized residues in the intervening positions, the ice binding surface is no longer flat and ice binding groups become eclipsed by apolar groups. This effect may be the primary cause for the observed loss of antifreeze activity when ice binding residues are rearranged in the winter flounder AFP (Wen & Laursen, 1992a).

As a result of the restrained conformation of IBM residues, ice binding groups adopt a well defined, periodic spatial arrangement (Fig. 3.7). Equivalent ice binding groups within adjacent IBMs are separated by  $16.5 \pm 0.5 \text{ \AA}$  while the ice binding groups within a single IBM are separated by  $6.1 \text{ \AA} \pm 0.3 \text{ \AA}$  and form a  $51 \pm 3^\circ$  crossing angle with the helix axis.

### 3.3 Solvent Structure

Since the winter flounder AFPs are generally believed to bind to ice through hydrogen bonding interactions (Knight, Driggers & DeVries; Yang, Sax, Chakrabarty, & Hew, 1988; Wen & Laursen 1992; Chou, 1992; Lal, Clark, Lips, Ruddock & White 1993; Madura *et al.*, 1994), the ordering of solvent around the protein's ice binding structure may provide clues as to its binding mechanism. In addition, analysis of ordered solvent helps to assess whether the AFP possesses an enhanced ability to bind solvent that may contribute to its antifreeze mechanism (Haschemeyer, Guschlbauer & DeVries 1977; Duman, Patterson, Kozak and DeVries, 1980). We have analyzed the

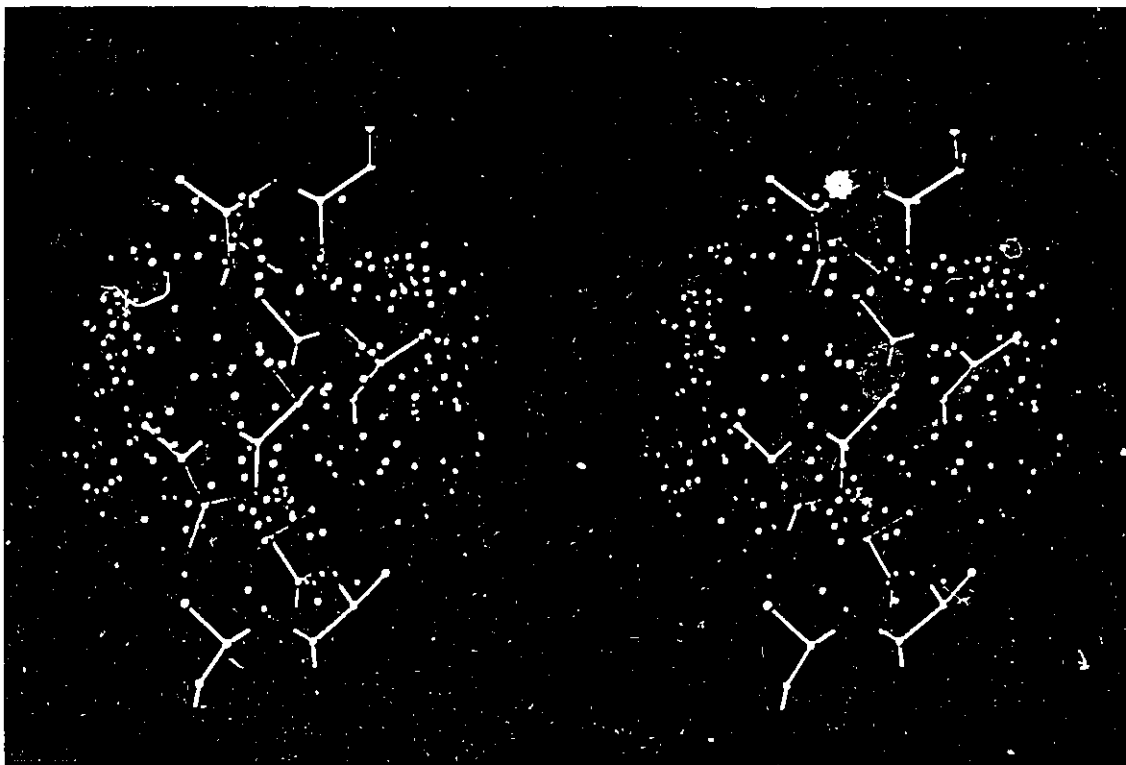


**Figure 3.7** Schematic of the spatial arrangement of the winter flounder AFP ice binding groups relative to the helix axis. The amino terminus and the ice binding groups are labeled.

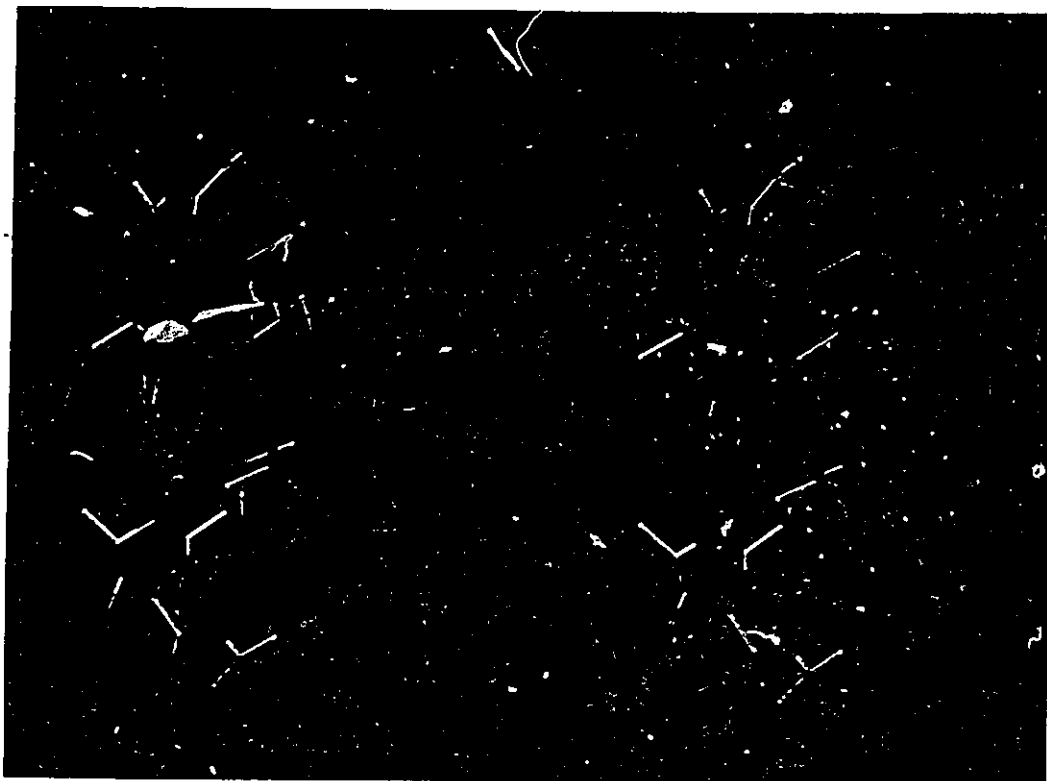
solvent structure around the winter flounder IBMs and the alanine residues which form the majority of the AFP's structure. Together these account for most of the bound solvent in the  $-180^{\circ}\text{C}$  crystal structure. The periodic nature of the AFP structure provides multiple instances of both groups for the study of solvent order by superposition.

In figure 3.8, we have superimposed the alanine residues of the two AFP molecules with their surrounding solvent. A cluster of water molecules is evident in the vicinity of the backbone  $\text{C}=\text{O}$  groups but no significant ordering in the form of recurring clathrate structure is evident. In contrast, water forms a random shell around alanine  $\text{C}\beta$  atoms. Since the binding of water to backbone  $\text{C}=\text{O}$  sites is a characteristic feature of solvent exposed  $\alpha$ -helices (Baker & Hubbard, 1984) the majority of the AFP structure that is represented by these alanine residues shows no evidence of an enhanced ability to interact with solvent.

Figure 3.9 shows the hydration pattern around the IBMs of the two AFP molecules. No specific hydration pattern is evident around the critical ice binding groups. Instead, water forms an elongated surface around the polar groups that is limited only by accessibility. These results are consistent with the observed variation from ideal geometry seen for the hydrogen bonding of water to the side chains of other proteins (Thanki, Thornton & Goodfellow 1988). The major implication for ice



**Figure 3.8** Superimposed solvent structure around the winter flounder AFP alanine residues. Alanine residues compose most of the protein's solvent exposed surface ( $\cong 60$  mol%). Alanine residues from the  $-180^{\circ}\text{C}$  AFP structures were superimposed with surrounding solvent by a least square fitting of backbone atoms. Characteristic clustering of solvent in the vicinity of the backbone oxygen atoms is evident. No unusual solvent ordering is apparent. Protein atoms of the displayed  $\alpha$ -helical segment (amino terminus up) are coloured as in figure 3.5. Blue to red colouring of water molecules represents isotropic B-factor variations ranging from 7 to  $55 \text{ \AA}^2$ .



**Figure 3.9** Superimposed solvent structure around the winter flounder IBMs. IBMs from the  $-180^{\circ}\text{C}$  structures with surrounding solvent were superimposed by a least square fitting of backbone atoms. Protein atoms of the displayed IBM containing  $\alpha$ -helical segment (amino terminus up) are coloured as in figure 3.5. Water molecules are shown in red.

binding is that hydrogen bonding requirements are unlikely to impose severe constraints on the AFP's selection of a binding site on ice.

### 3.4 Crystal packing arrangement

Crystal packing can be visualized as the repeat of two distinct parallel sheets, each composed of one of the AFP molecules in the asymmetric unit (see fig. 2.2, packing family A). Within a sheet, an AFP molecule makes head to tail contacts with its symmetry related partners while across the sheets, each AFP molecule makes contacts with four non-symmetry related molecules. Head to tail contacts involve electrostatic and hydrogen bonding interactions listed in table 3.1. AFP molecules on adjacent sheets form wide crossing angles that vary from  $26^\circ$  to  $32^\circ$  due to differences in the local curvature of each AFP molecule. Intermolecular contacts between AFP molecules on different sheets involve electrostatic, hydrogen bonding and hydrophobic interactions. Electrostatic and hydrogen bonding interactions are listed in table 3.1 while the four relatively hydrophobic crossing sites between adjacent helices are shown in figure 3.10.

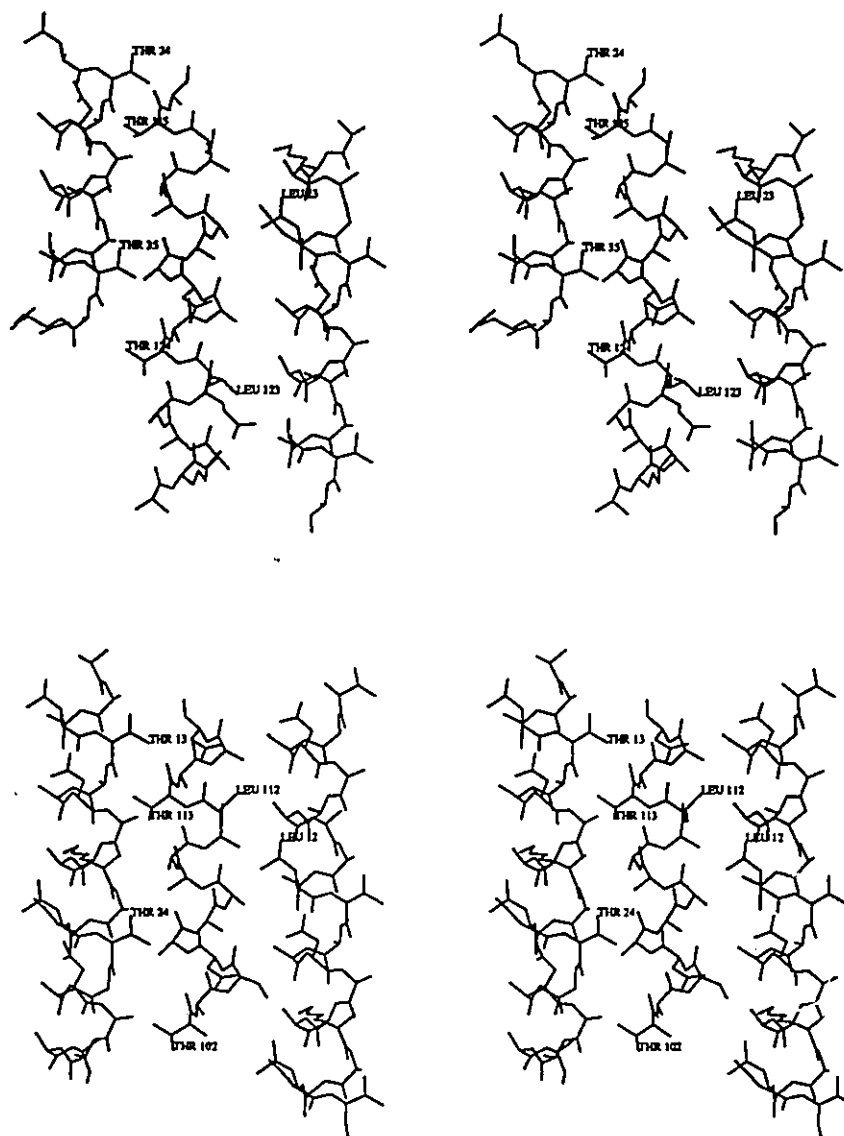
Crystal contacts have an effect on a number of side chain conformations which are clearly discernible in comparing the four unique AFP structures. Regarding the protein's helix stabilizing structure, most of the N- and C-terminal cap structures are not affected. However, crystal packing interactions involving the side chain of Asp 1



**Table 3.1 Intermolecular hydrogen bonding interactions of the 4°C crystal structure.**

Residue	Atom	Residue	Atom	Distance (Å)
Asp 1	N	Glu 122	OE2	3.0
Asp 1	OD2	Lys 118	NZ	3.1
Ala 3	O	Arg 37	NT1*	2.9
Asp 5	OD1	Arg 37	NH1	2.9
Asn 16	OD1	Arg 137	NE	3.0
Lys 18	NZ	Thr 135	OG1	3.3
Glu 22	OE1	Asp 101	N	2.7
Arg 37	NE	Asn 116	OD1	3.1
Asp 101	OD1	Arg 137	OT2*	3.1
Asp 105	OD2	Arg 137	NH1	2.8

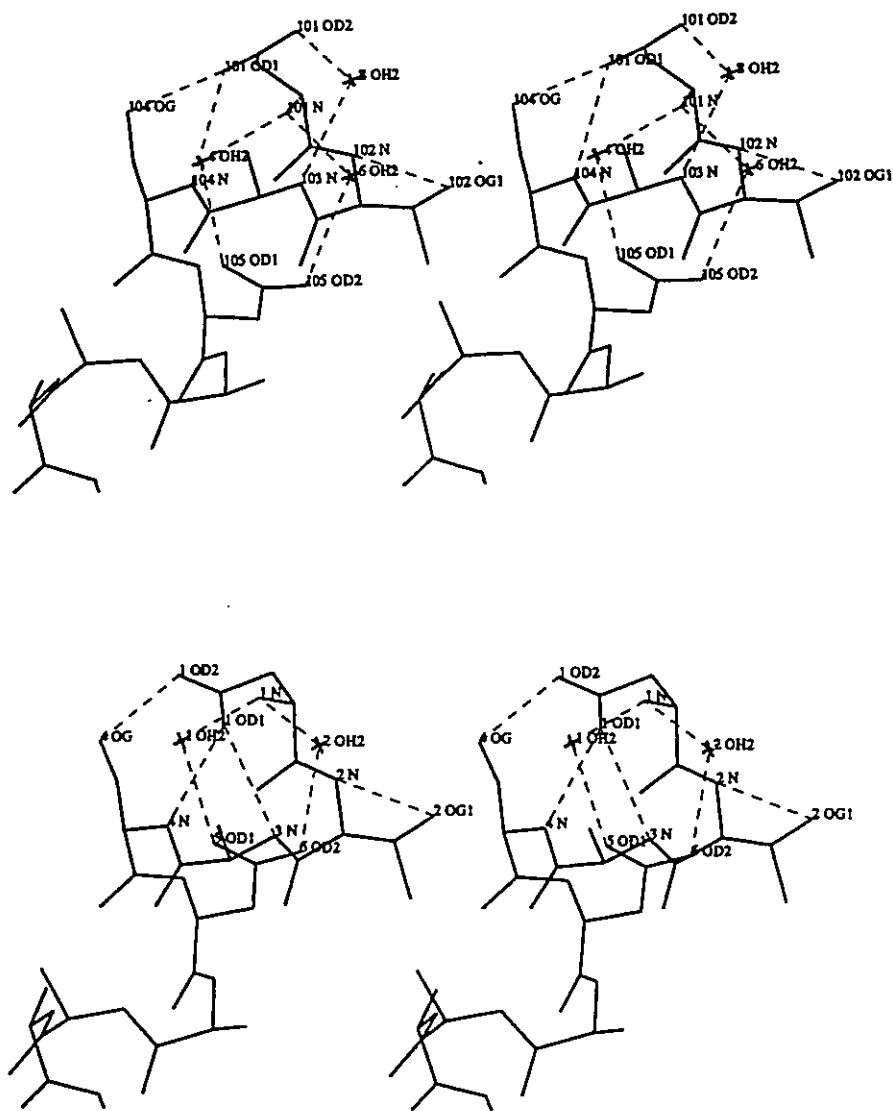
\* Nonstandard nomenclature refers to the C=O and NH<sub>2</sub> groups of the amidated carboxy termini. AFP molecule 1 = residues 1-37; AFP molecule 2 = residues 101-137.



**Figure 3.10** Stereo plot of the hydrophobic crossing sites between non symmetry related AFP molecules. 621, 777, 941 and 796 Å<sup>2</sup> of surface area are buried in the four contact points as read from top left to bottom right. Intermolecular hydrogen bonding contacts are listed in table 3.1.

in three out of four AFP molecules, gives rise to a slightly different N-cap structure shown in figure 3.11. The result of the packing interaction is that one of the C=O groups of Asp 1 is displaced requiring a water molecule to bridge a capping interaction. In three out of four AFP molecules, the Lys 18 and Glu 22 side chains are involved in crystal packing contacts which prevent intramolecular association. With respect to the proteins ice binding structure, in five out of eight instances (i.e. there are two Leu residues per AFP molecule) crystal packing interactions involving Leu side chains prevent intramolecular interactions with adjacent Asn side chains. Furthermore, the conformation of Thr 35 in two of the four AFP molecules and Asn 27 in one are distorted by crystal packing interactions.

The packing arrangement of the AFP molecules with relatively small contact patches and wide crossing angles between any two AFP molecules does not suggest an intermolecular association that may be relevant in solution or once the AFP is bound to the surface of ice. This agrees with solution studies which shows that the AFP has little tendency to aggregate (Chakrabartty, Ananthanarayanan & Hew, 1989) and with estimates of AFP densities on the surface of ice, based on elipsometry measurements, that suggest that AFP molecules are sparsely and independently bound ( Wilson, Beaglehole & DeVries, 1993).

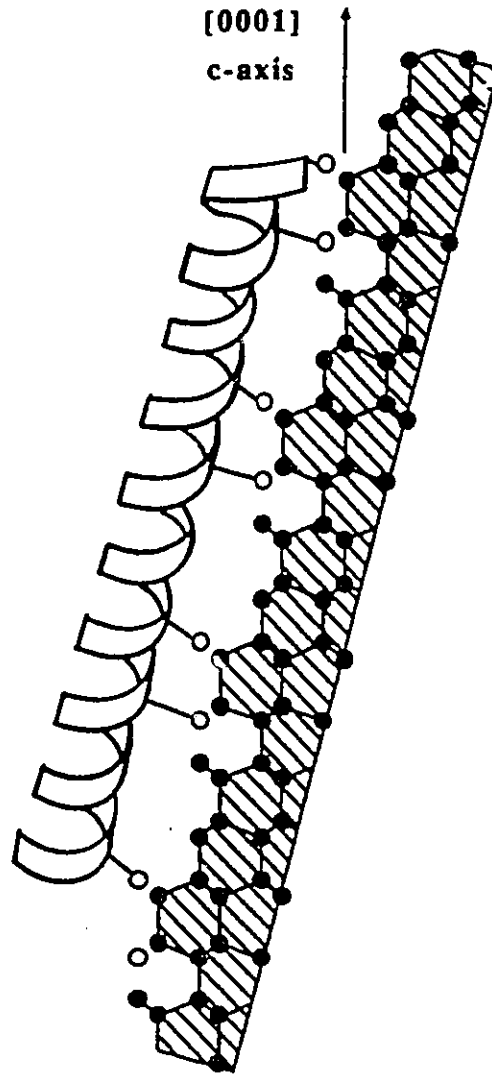


**Figure 3.11** Stereo view of the effect of crystal packing interactions on an N-terminal cap structure. The top structure, observed in three out of four AFP molecules, differs slightly from the bottom structure due to crystal packing interactions which distort the conformation of the Asp 1 side chain. Water molecules are labeled as OH2.

## CHAPTER 4 AFP ICE BINDING MODEL

### 4.1 Winter flounder system

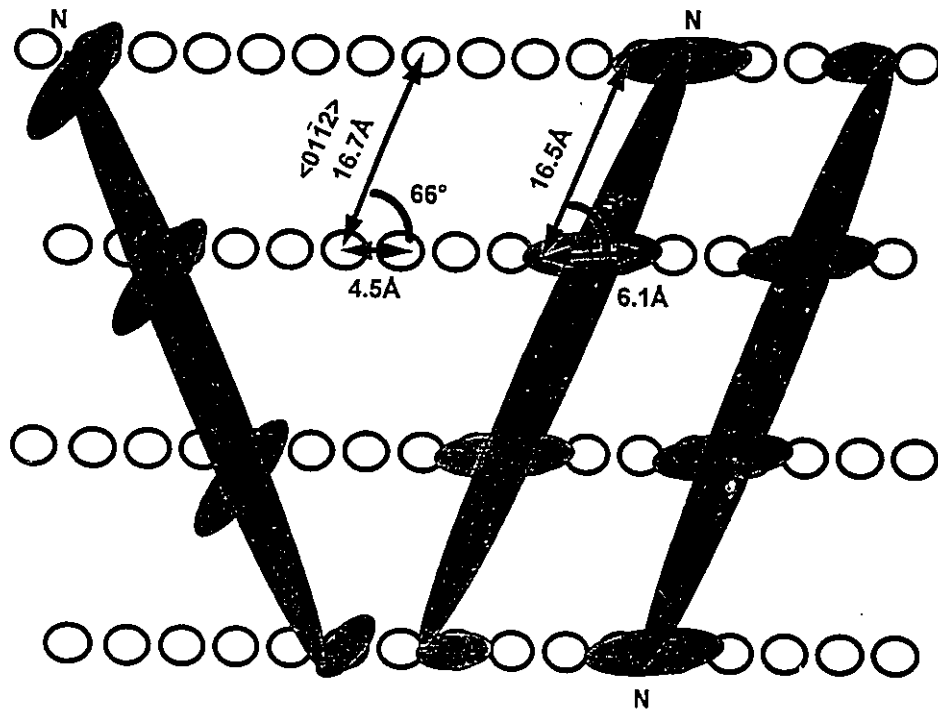
The flat and rigidly constrained properties of the AFP's ice binding structure lead us to hypothesize that the crystal structure of the winter flounder AFP closely reflects its ice binding conformation. Current ice binding models envision ice binding groups to extend from the AFP's structure to specifically occupy ice lattice positions. A schematic of one such model is shown in figure 4.1 with the AFP's ice binding groups occupying ice lattice positions in the first layer of ice (Wen & Laursen 1992b). A similar model envisions the AFP's ice binding groups to anchor more deeply into ice by occupying second or third layer sites (Knight, Driggers & DeVries 1993). Supporting these binding models was the identification of a potential spatial match between the ice binding groups of the winter flounder AFP and ice by molecular modeling (Wen & Laursen 1992b). However, several lines of evidence revealed by our crystal structure appear to contradict this view. Firstly, a sufficiently close spatial match between the water molecules on the {2021} ice face and the ice binding groups of the AFP is not observed; while the average separation between Thr or Asx residues in adjacent IBMs closely matches the  $16.7\text{\AA}$   $\langle 0112 \rangle$  axial repeat distance of water molecules on the binding plane, the closest match on ice found to fit the spacing of binding groups within an IBM is  $4.5\text{\AA}$ . Secondly, the ice binding groups do not protrude sufficiently from



**Figure 4.1** Ice binding model proposed by Wen and Laursen, 1992b. Polar ice binding atoms show as open circles are envisioned to rigorously occupy lattice positions in the top layer of ice. Water occupied lattice positions are shown as solid circles. A similar model (Knight, Driggers and DeVries, 1993) envisions protein atoms to occupy deeper lattice sites.

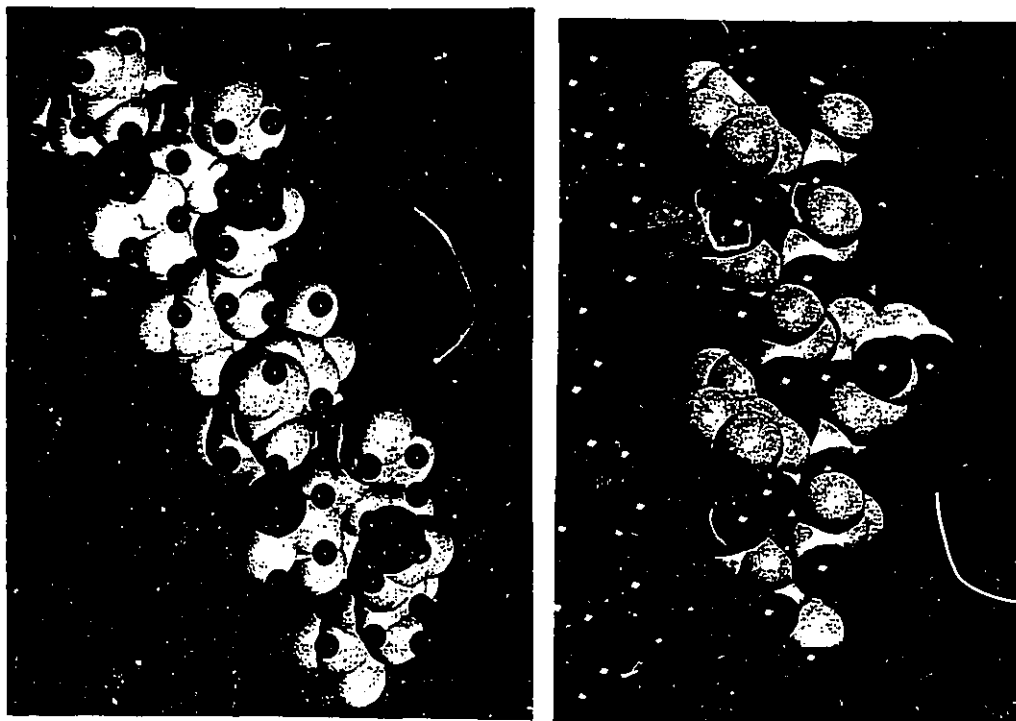
the AFP's ice binding surface to clear sterically hindering groups; the C $\gamma$  atoms of the Thr residues and the C $\beta$  atoms of Ala 9, 20 and 31 would prevent the ice binding groups from entering into ice (i.e. first layer sites or deeper) even if a coincidental lattice match was available. Thirdly, the trigonal planar (SP $^2$ ) coordination of the Asn and Asp hydrogen bonding groups differ from the tetrahedral (SP $^3$ ) coordination of water molecules in ice.

In contrast, we envision that the binding of AFP to ice is defined by a less stringent hydrogen bonding criteria and this forms the basis of an ice binding model shown in figure 4.2 (schematic) and figure 4.3 (space filling model). A characteristic feature of the  $\{20\bar{2}1\}$  adsorption plane is a ridge and valley topology that is subtended by the  $\langle 01\bar{1}2 \rangle$  binding axis by  $66^\circ$ . This topology complements the AFP structure when the AFP is aligned along the binding axis. Considering that the hydrogen bonding groups extend minimally from the AFP's flat ice binding surface, water molecules on the ridges of the  $\{20\bar{2}1\}$  ice plane are the most probable sites for binding. The  $4.5\text{\AA}$  spacing of water molecules along the ridges provides accessible hydrogen bonding targets for both ice binding groups within an IBM. As a result of the flatness exhibited by the AFP's ice binding surface, our modeling studies suggest that the protein may bind along either direction of the  $\langle 01\bar{1}2 \rangle$  axis. This contrasts previous predictions that the AFP orients itself in a single direction (Wen & Laursen, 1992b; Madura et al., 1994).



**Figure 4.2** Winter flounder AFP ice binding model. Cartoon representation showing the complementarity between the AFP structure and the topology of the  $\{20\bar{1}2\}$  ice plane. AFP molecules, shown in dark shading with amino termini labeled, display the relative positioning and geometry of ice binding groups relative to the helix axis. Water molecules that form the prominent ridges on the ice plane are shown as circles. The two AFP molecules on the right are aligned along the  $\langle 01\bar{1}2 \rangle$  binding axis while the molecule on the left is aligned along the mirror related axis. The loss of complementarity may explain why this AFP does not bind along the mirror related axis.





**Figure 4.3** Front and side views (left and right respectively) of an AFP molecule (residues 9 to 28) docked along the  $\langle 01\bar{1}2 \rangle$  binding axis of the  $\{20\bar{2}1\}$  ice plane. The side view shows the punctuated flatness of the protein's ice binding surface and the ridge and valley topology of the ice plane. In the front view, protein atoms with the exception of ice binding groups are shown in yellow, and target molecules on the ice plane are shown in green.

The precise side chain conformations provided by the newly determined crystal structure reveals, for the first time, detailed characteristics of an AFP's ice binding surface. We believe that the relative flatness of this surface and the rigidity of the ice binding residues are critical to the AFP's ice binding mechanism. Rigidly constrained ice binding residues would help to maintain the AFP in its ice binding conformation while the flatness of the ice binding surface would maximize the accessibility the AFP's binding groups to an ice surface. Together, these properties would allow for the concerted formation of many hydrogen bonds between the AFP and ice to give permanence to binding. The spatial arrangement of ice binding residues likely accounts for the binding ability and the specificity of the AFP, but the underlying hydrogen bonding interactions are likely to be more liberally defined than previously proposed. The finding that all of the AFP's structure appears dedicated to either ice binding or helix stabilization strongly suggests that ice binding is the only essential characteristic required for antifreeze function.

#### **4.2 Implication for other AFPs.**

While the use of a flat and rigid surface for ice binding may be a general feature of most of the  $\alpha$ -helical AFPs it is not clear whether this is the case for all AFPs. This question will best be answered as the high resolution structures of other AFPs become available. However, a number of AFP systems have been sufficiently characterized to

comment on the potential generality of the winter flounder AFP's ice binding characteristics.

In the  $\alpha$ -helical class of AFPs (class 2), a number of sculpin AFPs, namely Arctic sculpin-3 and shorthorn sculpin-8, possess primary sequences that differ significantly from most members of the class in lacking a periodicity to its primary sequence (fig 1.3). These AFPs possess only one recognizable IBM and have been shown to bind to a unique family of ice planes (Knight, Cheng & DeVries, 1991). Since truncation studies of the winter flounder AFP suggest that it is unlikely that an  $\alpha$ -helical AFP can make use of only a single IBM to bind ice (Chakrabarty, Ananthanarayanan & Hew, 1989), other residues must be involved in ice binding. A number of polar residues that may function in ice binding are found localized in a unique N-terminal cap sequence (discussed in chapter 3.1). Since in the winter flounder structure, an elaborate N-terminal cap structure appears to help maintain the ice binding conformations of Asp 5 & Thr 2, it is possible that a novel N-cap structure in the sculpin AFPs may perform a similar function. However, a search for homologous primary sequences in the Brookhaven protein data bank did not reveal a conserved 3D conformation for either of the two sculpin sequences. NMR and x-ray crystallographic studies have now been initiated in Dr. Yang's laboratory to further probe these non-periodic  $\alpha$ -helical AFPs (chapter 7.1).

At first glance, the high flexibility of the carbohydrate containing AFPs from cod (class 1) suggest that structural rigidity is not an important characteristic for ice binding. However, the inherent rigidity of the cyclical carbohydrate moieties may compensate for the AFP's backbone flexibility. With regard to flatness of the AFPs ice binding surface, NMR and modeling studies suggest that the AFPs may adopt a three fold helical conformation which would orient carbohydrate moieties along one face to form a flat ice binding surface (Rao & Bush, 1987).

NMR studies have revealed that an AFP from ocean pout (class 3), isoform HPLC 12, possesses two flat surfaces (Sonnichsen, Sykes, Chao & Davies, 1993; Chao, Davies, Sykes, & Sonnichsen, 1993) and recent mutagenesis studies have localized the ice binding residues to one of the faces (Chao, Sonnichsen, DeLuca, Sykes, & Davies, 1994). As side chain conformations were not accurately determined in the NMR study (structure is based largely on backbone atom constraints), it is not yet clear if the AFP conforms to the winter flounder model with restrained side chain conformations. High resolution ( $<2\text{\AA}$ ) structural studies on single crystals of an ocean pout AFP (Xue, Sicheri, Ala, Hew, & Yang, 1994) are currently underway in Dr. Daniel Yang's laboratory to try to resolve this question.

## CHAPTER 5 COMPARISON OF THE AFP CRYSTAL STRUCTURE WITH OTHER MODELS

### 5.1 Comparison to the previous x-ray crystal structure

In the previous x-ray crystallographic analysis of the winter flounder AFP structure (Yang, Sax, Chakrabarty & Hew, 1988), Patterson map features were visually interpreted to identify the tilt of the protein molecules and the relative positioning between the two AFP molecules in the asymmetric unit. In order to avoid a bias towards the previous structural model which failed to refine, the visual interpretations were not taken into account in the present structure determination. Our crystal structure now confirms that both interpretations of Patterson map features were in fact correct. It appears that the previous model failed to refine due to the difference in curvature between the search model and the structure of one of the two AFP molecules in the asymmetric unit. The difference in curvature was not surmountable by the refinement routines previously employed but was resolved by the *XPLOR SA* refinement routine employed in the present study. While, the packing arrangement observed in the previous determination is similar to that observed in the present structure determination, the synchronisation of the backbone atoms, the backbone curvature and direction of one AFP molecule, and the precise rotation/translation of both molecules differed significantly. Reflecting these differences, we were unable to

use the previous x-ray model to locate heavy atom positions in  $\text{LaNO}_3$  and  $\text{TbNO}_3$  difference Fourier maps.

## 5.2 Comparisons to simulation studies

The low resolution image of the winter flounder AFP provided by the first crystallographic analysis spawned numerous modeling studies to elucidate side chain conformations that gave the AFP its ice binding ability and exceptional  $\alpha$ -helical stability. These included simpler studies of the AFP in vacuum or in solution (Chou, 1991; McDonald, Brady & Clancy 1993; Jorgenson *et al.*, 1993) and more complex studies of the AFP on the surface of ice (Wen & Laursen, 1992; Lal, Clark, Lips, Ruddock & White, 1993; Madura, Wierzbicki, Harrington, Maughon, Raymond & Sykes, 1993). Since these studies have had a significant influence on the current thinking of how AFPs bind to ice, comparisons with the high resolution AFP crystal structure are warranted. What follows is a brief look at the salient features of the simpler and then the more complex theoretical studies.

With regard to the AFP's helix stabilizing features, the salt bridge interaction between the side chains of Lys 18 and Glu 22 was reproduced in two of the three simpler studies (Jorgenson *et al.*, 1993; McDonald, Brady & Clancy, 1993). These two studies were also able to reproduce a small proportion of the protein's N-terminal cap structure. All three studies failed to reproduce the C-terminal cap structure and

this may be attributed in part to how the carboxy terminus of the AFP was treated. Jorgenson *et al.* (1993) analyzed an immature form of the AFP which contained an additional glycine residue rather than the amidated carboxy terminus of the mature AFP (Hew, Wang, Yan, Cai, Sclater & Fletcher, 1986). In Chou's (1991) and McDonald, Brady & Clancy's (1993) studies, the amide modification of the carboxy terminus was not included. As noted in the x-ray structure analysis, the amidated carboxy terminus plays a critical role in defining the C-cap structure. With regard to the AFP's ice binding structure, all three studies accurately reproduced the conformation of the Thr side chains but in most instances Asn and Asp side chains were found to adopt unrestrained extended conformations which differed from the contracted and restrained conformations observed in the crystal structure. With regard to the AFP's backbone structure, the study of McDonald, Brady & Clancy (1993) was unique in observing that the AFP had a predisposition towards bending. The curvature of AFP molecule 1 in the 4° and -180°C crystal structures shows that the AFP molecule is in fact flexible but no evidence for a predisposition towards a particular bend is apparent.

Since the study of AFP molecules in vacuum or in solution were not able to reproduce much of the protein's key structural features, we question the accuracy of results and interpretations derived from more complex systems which explicitly treated both AFP and ice. In three separate studies, which were able to account for the AFP's orientational binding preferences on the {2021} ice plane (Knight, Cheng & DeVries,

1991), the resultant protein structures and the physical basis for the AFP's binding preferences differed greatly. In Wen & Laursen's (1992) study, binding specificity was attributed to optimal hydrogen bonding interactions arising from a spatial match between ice binding groups on the AFP and the water molecules on the ice surface (Fig 4.1). In Lal, Clark, Lips, Ruddock and White's (1993) study, the binding specificity of the AFP was attributed to a van der Waals complementarity between the AFP's ice binding surface and a ridge and valley topology of the ice plane. In the study of Madura, Wierzbicki, Harrington, Maughon, Raymond & Sykes (1993), Asp and Asn ice binding side chains were found to integrate into the ice lattice in an unusual manner. Considering the complexity of the simulation systems, it is not surprising that different results were obtained. The AFP has many degrees of conformational freedom which are unlikely to be fully sampled in a simulation study and hence, results will be highly dependent on the starting conformation of the AFP and ice models employed. For this reason, the abilities of the studies to account for the winter flounder AFP's ice binding specificity are likely to be artefactual.

The high resolution crystal structure of the winter flounder AFP should facilitate future simulation studies by providing a more accurate starting model. By reducing the complexity of the protein structure, it should be possible to improve the accuracy of simulation results further. Side chains involved solely in helix stabilization can be removed and implicitly treated with  $\alpha$ -helical backbone constraints. This includes the



side chains of residues Lys 18, Glu 22, Arg 37, Asp 1 and Ser 4. This may prove more effective in maintaining the AFPs  $\alpha$ -helical conformation over longer simulation periods. Problems with fraying helix termini and backbone distortions were observed in a number of simulation studies and these may adversely affect the study of the AFP's interaction with ice.

## **CHAPTER 6 APPLICATION OF THE MOLECULAR REPLACEMENT STRATEGY TO OTHER $\alpha$ -HELICAL STRUCTURES**

The simplicity of predominantly  $\alpha$ -helical structures such as the winter flounder AFP, the peptitergent PD<sub>1</sub> (Schafmeister, Miercke and Stroud, 1993) and the model four helix bundle protein  $\alpha$ 1 (Hill, Anderson, Wesson, DeGrado & Eisenberg, 1990), would seem to make them well suited for structure determination using molecular replacement techniques with idealized  $\alpha$ -helices as search models. However, a survey of the literature, shows that such an approach is not common. To date, only the peptitergent PD<sub>1</sub> structure has been correctly determined using this approach and a successful outcome can be attributed in great part to the crystal system's convenient packing arrangement with one molecule in the asymmetric unit. Our results clearly demonstrate why  $\alpha$ -helical structures are likely to pose serious problems for solution by the aforementioned technique. One problem arises from the symmetry of the  $\alpha$ -helical structure itself. This makes distinguishing correct helix directions and rotations about the helix axes impossible since a sufficient number of the side chain conformations which can break rotational ambiguity are not usually known beforehand. Furthermore, in the winter flounder AFP example reported here, the knowledge of the strong

conformational preference of the Thr side chains was only helpful after significant refinement of the backbone structure had been performed. Determining correct translation solutions can also present a problem depending on the tilt of the  $\alpha$ -helical molecules with respect to the target symmetry axis. The packing arrangement of the AFP molecules, with relatively perpendicular helix tilts with respect to the unique crystallographic axis (off by  $\sim 15^\circ$ ) represents a more difficult extreme while the near parallel alignment (off by  $\sim 13^\circ$ ) of the peptitertgen PD1 relative to the unique axis of its crystal system represents a simpler extreme for translation determination. Difficulties encountered in rotation and translation searches are obviously compounded by having a greater number of molecules in the asymmetric unit and the deviation of the search model from the actual structure presents an even greater problem. Both of these problems were encountered in the AFP structure determination and were surmounted by a multi-parameter search and refinement scheme. While an  $R_{\text{factor}}$  comparison after the SA refinement stage was sufficient to distinguish the correct solution from a large number of potential MR solutions, the use of a heavy atom derivative to filter and correlate the origin of MR solutions provided a helpful and complementary means to this end.

By combining a careful visual interpretation of Patterson map features (Yang, Sax, Chakrabartty & Hew, 1988) with a rational search of unknown parameters followed by SA refinement of the better solutions, it should be possible to solve the

structure of other simpler lone  $\alpha$ -helical proteins and possibly more complex systems without the aid of isomorphous replacement techniques.

## **CHAPTER 7 TESTING THE PROPOSED ICE BINDING MODEL**

The proposed ice binding model is based on the hypothesis that the crystal structure of the AFP reflects its ice binding conformation. Three approaches to test the binding model have been designed and are currently under way. The first approach is to determine if the conformation of the ice binding side chains observed in the crystal structure are also observed in solution. The second and third approaches involve disrupting AFP structural features deemed important by the proposed model to observe the effect on AFP function. In the two latter approaches, respectively, site directed mutagenesis and the natural structural diversity within the  $\alpha$ -helical class of AFPs are to be employed.

### **7.1 NMR studies on mini AFP constructs**

In collaboration with Dr. Pele Chong of Connaught Laboratories, Dr. Guang Zhu from Hong Kong University of Science and Technology, and Dr. Ananthanarayanan of McMaster University, studies into the solution conformation of model AFP peptides have been initiated. The peptide sequences to be studied are shown in figure 7.1. These are versions of natural AFPs which have been truncated in order to simplify NMR analyses and to reduce the cost of peptide synthesis. Model peptides 1 to 4 contain the winter flounder N- and C-terminal capping sequences



(needed to maintain  $\alpha$ -helical stability) flanking a single ice binding motif. Peptides 1 and 2 possess a fully intact Thr/Leu/Asn IBM (from winter flounder) and a Thr/Lys/Asp IBM (from most other  $\alpha$ -helical AFPs) respectively. Model peptides 3 and 4 are Leu to Ala and Lys to Ala IBM mutants of peptides 1 and 2 respectively. Using NMR methods, we hope to show that Asn/Leu and Asp/Lys IBM side chain pairs associate in solution to restrain side chain rotations. In the Leu to Ala and Lys to Ala mutants, the proposed intra-molecular side chain stabilizing interactions are not possible and hence the conformation of the Asn and Asp IBM side chains are predicted to be relatively more mobile.

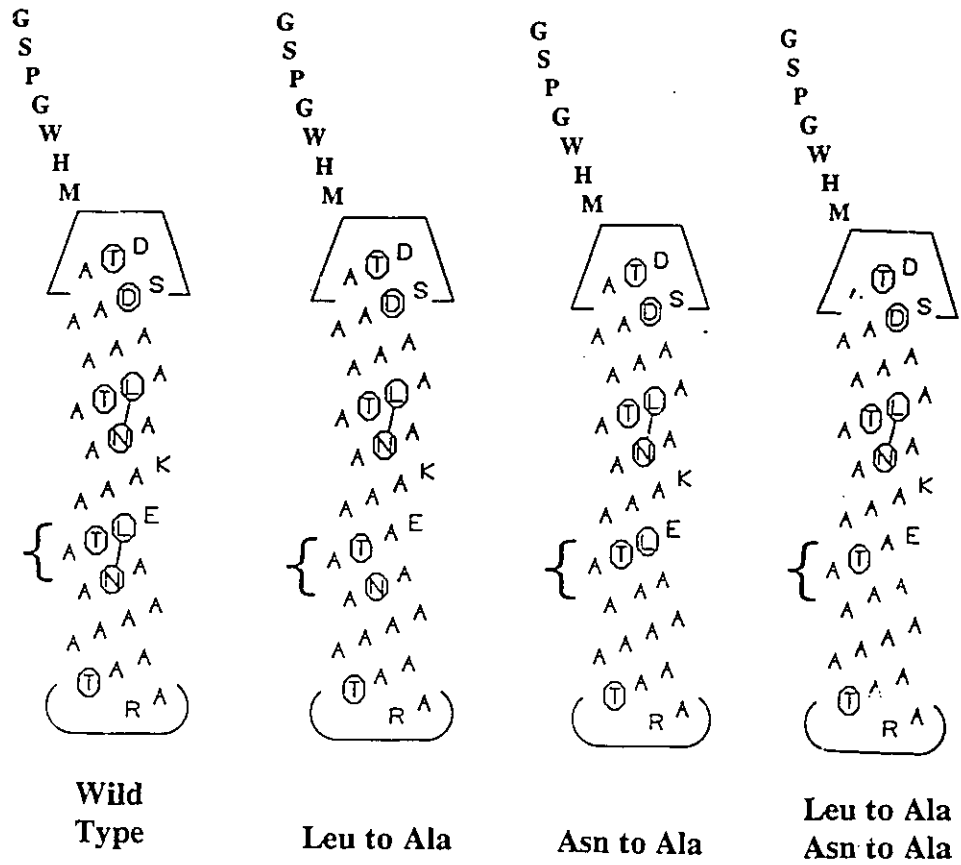
In addition to the aforementioned solution studies into the side chain conformation of the winter flounder and homologous IBM sequences, NMR and crystallization studies have been initiated on the two potential N-cap sequences found conserved amongst the sculpin  $\alpha$ -helical AFPs (chapter 3.1). These N-cap sequences have been incorporated into model peptides 5 and 6 respectively (figure 7.1) along with the C-cap sequence of the winter flounder AFP. In addition to uncovering a potential N-cap structure, it is hoped that the studies on model peptide 6 will also reveal the conformation of the yet undefined ice binding side chains located within the N-terminal sequence (chapter 4.2). This will help to test whether rigid and flat ice binding structures are a general feature of all the  $\alpha$ -helical AFPs and also to uncover the origin of the Arctic sculpin-3/shorthorn sculpin-8 AFP's unique ice binding specificity

(Knight Driggers & DeVries, 1991). All six peptides have been synthesized and purified. NMR and crystallization experiments are currently under way.

## 7.2 Site directed mutagenesis on recombinant AFP

Previous mutation studies have shown that the substitution with alanine of both leucine residues within the central IBMs of the winter flounder AFP had the effect reducing antifreeze activity by 33%. This observation in combination with modeling studies led to the proposal that the leucine residues function independently and directly to interact with ice (Wen & Laursen 1992a+b). In contrast, we believe that the resultant loss of AFP function from leucine substitution is due to the loss of a stabilizing effect on the conformation of the spatially adjacent (i+4) asparagine residues. As a test, we predict that the effect of leucine residues on function is strictly dependent on the presence of the paired Asn residues. Furthermore we predict that Leu and Asn side chains alone may be more detrimental to antifreeze function than if both of the paired residues were absent. This would account for the observation that the residues are present or absent as pairs from all recognizable IBMs (fig 1.3). Figure 7.2 shows a number of winter flounder AFP mutants that have been designed to test the aforementioned hypothesis. The predicted order of AFP antifreeze effectiveness is WT >> Ala Ala >> Ala Leu  $\cong$  Ala Asn. Such a result is not predicted by earlier ice binding models.



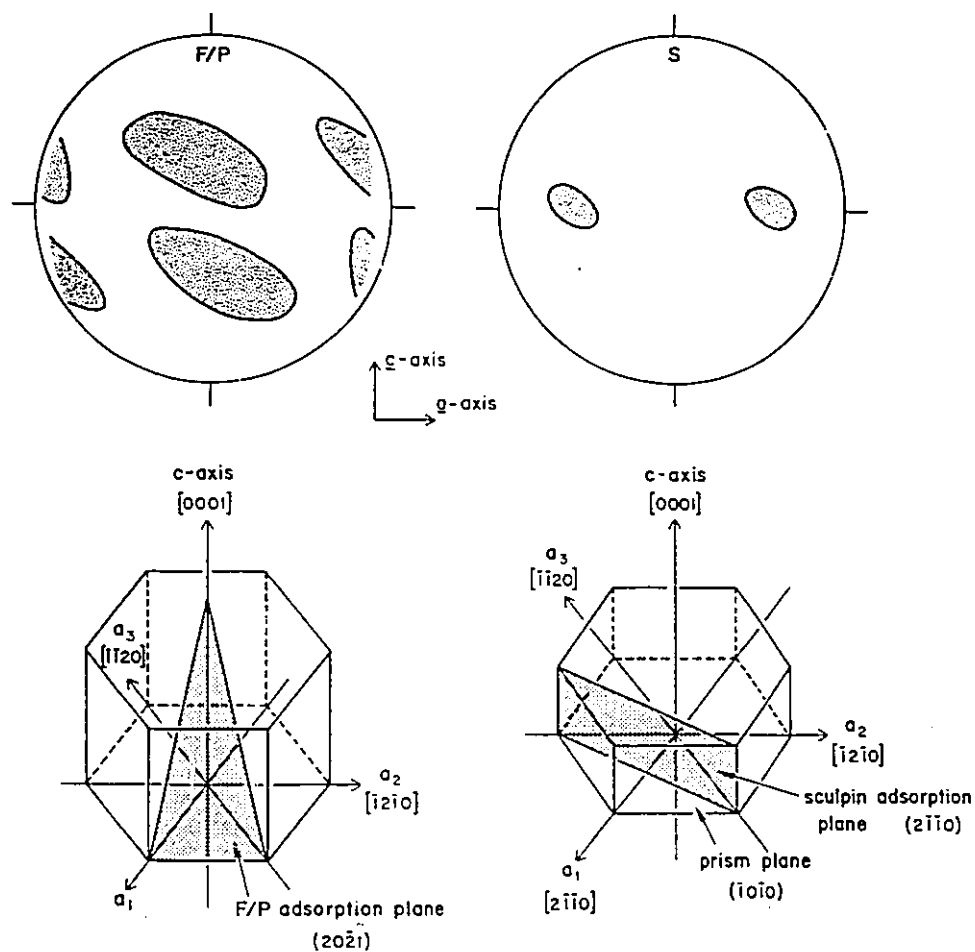


**Figure 7.2** Wild type and site directed mutants of a recombinant winter flounder AFP. After liberating the AFPs from their GST fusion partners, seven non-native amino acids remain attached to the amino termini and the carboxy termini are not amidated.

An expression system for the winter flounder AFP in *E. coli* has been developed to facilitate the aforementioned mutagenesis studies (Xue, Lepinsky, Cacky, Chen, Sicheri, Sung & Yang, work in progress). The winter flounder AFP is expressed as a C-terminal fusion to Glutathione-S-Transferase (GST). Low levels of antifreeze activity are detectable for the fusion protein and a large increase in activity is observed when the AFP is cleaved from the GST carrier protein with thrombin. Seven non-native amino acids remain attached to the amino terminus of the AFP after thrombin cleavage of which one amino acid is the aromatic residue tryptophan (fig 7.2). The presence of the aromatic residue facilitates accurate quantitation of peptides by absorbance measurements @ 280nm. An additional difference between the recombinant AFP and the native serum derived winter flounder AFP is the absence of a carboxy terminus amide modification. The four AFP variants in figure 7.2 have been designed and successfully overexpressed. Purification of the AFPs and the detailed measurement of antifreeze activities are in progress.

### **7.3 Prediction of ice binding specificities.**

AFP binding specificity was first demonstrated by Knight and coworkers (1991) using elegant experiments whereby a single oriented crystal of ice was placed in dilute solutions of AFP and grown to form a large hemisphere (figure 7.3 top). These single hemispherical crystals of ice were then placed into a cold room and the top layers of ice



**Figure 7.3** Ice binding specificity assay. The top figures, left to right respectively, display the adsorption patterns of winter flounder HPLC8 and sculpin AFPs on a single hemispherical crystal of ice. Bottom figures represent the binding planes and alignments of the AFPs as extrapolated from their respective adsorption patterns. Reproduced from Knight, Cheng, and Devries, 1991.

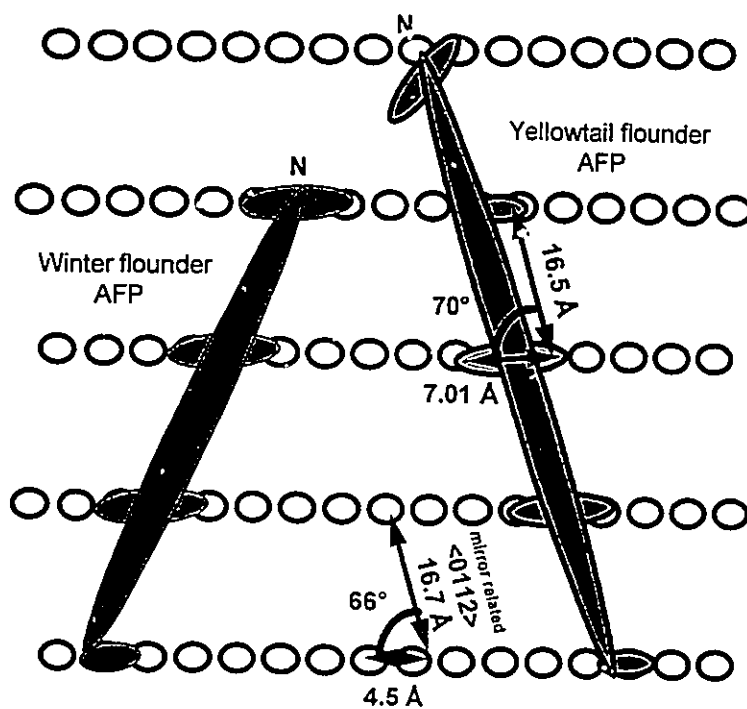
were sublimed away with time to reveal white patches of proteins that correspond to the absorption sites of the AFP. Because the original ice crystal was oriented with respect to its crystal axes, the protein patches could be used to extrapolate the binding planes of the AFP on ice (figure 7.3 bottom). In studies performed on the structurally elongated AFPs (i.e. the carbohydrate containing AFPs of cod and members of the  $\alpha$ -helical class of AFPs) an asymmetrically shaped binding patch also revealed the binding orientation of the AFP on its binding plane. Such studies showed that the winter flounder and Alaskan plaice AFPs binds along  $\langle 01\bar{1}2 \rangle$  axes of the  $\{20\bar{2}1\}$  ice planes, that a shorthorn sculpin AFP (very similar to Arctic sculpin-3 and shorthorn sculpin-8) binds to the  $\{2\bar{1}\bar{1}0\}$  ice planes and that the smaller carbohydrate containing AFPs (isoforms 7 & 8) bind parallel to the  $a$ -axes on the  $\{\bar{1}0\bar{1}0\}$  ice planes (Knight, Cheng & DeVries, 1991; Knight, Driggers & DeVries, 1993). Subsequent studies demonstrated that the synthetic D isomeric form of the winter flounder AFP binds along the *mirror related*  $\langle 01\bar{1}2 \rangle$  axes of the  $\{20\bar{2}1\}$  ice plane (Laursen and Knight, 1994). Such a result was expected since the D and L isoforms of the winter flounder AFP are perfect mirror images of themselves. When D and L isoforms of the winter flounder AFP were mixed, a criss-crossing binding pattern on the  $\{20\bar{2}1\}$  ice planes was found.

Two  $\alpha$ -helical AFPs, one from yellowtail flounder and one from winter flounder; have yet to be assayed for their ice binding specificity and these provide a test of our understanding of the ice recognition process. As noted in chapter 3.2, the

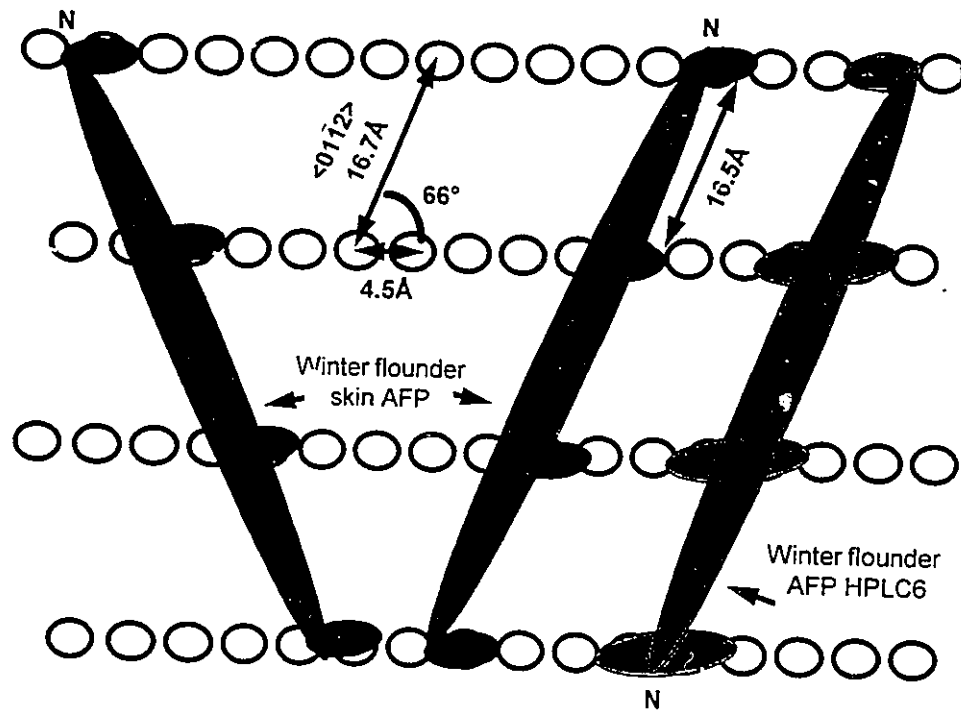
yellowtail flounder AFP is different from all other  $\alpha$ -helical AFPs in possessing a different radial separation between Thr and Asp/Lys IBM residues. The difference in the radial separation should result in a drastic change in the relative positioning of the ice binding group with respect to the helix axis as shown in figure 7.4. Based on the inferred spatial arrangement of ice binding groups, we predict that yellowtail flounder will bind with the same orientational preference as the D isomeric form of the winter flounder AFP.

The second AFP which has yet to be characterized for its ice binding specificity was recently discovered in the laboratory of Dr. Choy Hew at the University of Toronto. This AFP is unique in having all of its IBMs incomplete with only Thr IBM residues (personal communication). Although all of its IBMs are incomplete, this AFP is still functional. Because of the incomplete state of the AFP's IBMs, the ice binding structure is linear which leads us to predict that the AFP will bind along the  $\langle 01\bar{1}2 \rangle$  axes and also the *mirror* related axes of the  $\{20\bar{2}1\}$  ice plane (Fig. 7.5).

Collaborative studies have been initiated to test the described binding specificity predictions. The newly discovered winter flounder AFP is being purified in Dr. Choy Hew's laboratory while the yellowtail AFP (serum kindly provided by Dr. Garth Fletcher) is being purified in Dr. Daniel Yang's laboratory. Using the purified AFP samples, binding assays are to be performed by Dr. Charles Knight from the University of Colorado. Purification experiments are now in progress.



**Figure 7.4** Predicted binding specificity of the yellowtail flounder AFP. Cartoon representations show the complementarity between AFP structures and the topology of the  $\{20\bar{2}1\}$  ice plane. AFP molecules, shown in dark shading with amino termini labeled, display the positioning and geometry of ice binding groups relative to the helix axis. Water molecules that form the prominent ridges on the ice plane are shown as circles. The yellowtail AFP molecule on the right is aligned along the 'mirror related'  $\langle 01\bar{1}2 \rangle$  binding axis while the winter flounder AFP molecule on the left is aligned along the  $\langle 01\bar{1}2 \rangle$  axis.



**Figure 7.5** Predicted binding specificity of a newly discovered skin AFP from winter flounder. Cartoon representations shows the complementarity between AFP structures and the topology of the {20 $\bar{2}$ 1} ice plane. AFP molecules, shown in dark shading with amino termini labeled, display the positioning and geometry of ice binding groups relative to the helix axis. Water molecules that form the prominent ridges on the ice plane are shown as circles. The two skin winter flounder AFP molecules on the left are aligned along both the 'mirror related'  $\langle 01\bar{1}2 \rangle$  and the  $\langle 01\bar{1}2 \rangle$  axes while the winter flounder AFP molecule (HPLC6) on the right is aligned along the  $\langle 01\bar{1}2 \rangle$  axis.

## CHAPTER 8. CONCLUSIONS

The 1.5Å x-ray crystal structure of the winter flounder AFP, isoform HPLC6, has been determined and a new ice binding model is proposed. Mutagenesis, NMR, and ice binding specificity experiments to test the ice binding model have been proposed and are currently underway. The molecular replacement method used to solve the inherently difficult AFP structure may prove useful for the solution of other  $\alpha$ -helical structures.



## GLOSSARY

The following are descriptions of crystallographic terms, functions and methods highlighted in the body of this work with bold italics that pertain to the x-ray crystal structure determination of the winter flounder AFP. Much of the information contained in this section was obtained from readings of Protein Crystallography (Blundell & Johnson, 1976), X-Ray Structure Determination (Stout & Jensen, 1989) and the *PHASES* manual (Furey, 1992).

**ASIR (Automated Solvent Inclusion and Refinement):** A suit of programs for the automatic iterative selection and refinement of solvent molecules in macromolecular x-ray crystal structures. The core program of this suit of programs determines by distance, hydrogen bonding and B-factor cutoff criteria whether potential difference Fourier peaks are to be selected for inclusion as water molecules in crystallographic refinement. Originally designed for use with the *PROLSQ* refinement program, shell programs have been written to allow the core program to work with *XPLOR*.

**Electron density ( $\rho$ ):** Electron distribution within a 3D crystal structure.

$$\rho(x,y,z) = 1/V \sum_h \sum_k \sum_l F_{hkl} e^{-2\pi i (hx+ky+lz)}$$

**Friedel Mates:** The paired structure factors  $F_{hkl}$  and  $F_{\bar{h}\bar{k}\bar{l}}$ , originating from the same crystallographic planes, are referred to as Friedel mates. The difference between the magnitudes of the paired structure factors arises from the anomalous scattering properties of heavy atoms.

**Figure of Merit (FOM):** A measure of the phasing quality of a heavy atom derivative. FOM is the cosine of the variance of the phase probability distributions. An ideal distribution, FOM=1, has a very sharp phase probability distribution while the other extreme, FOM=0.0, has a flat phase probability distribution. For protein structure determination, the phasing information of a derivative is excellent if FOM > 0.85, good if FOM is > 0.75, poor < 0.6. This parameter has direct bearing on the interpretability of an electron density map.

**Harker Plane:** In the  $P2_1$  space group, the Patterson UW cross section through  $V=0.5$  represents the Harker Plane. Major peaks on a Harker section represent cross vectors between symmetry related heavy atoms sites. From the major peak positions, the placement of the heavy atoms in the unit cell can be determined. For a Harker peak at the fractional Patterson positions U,W, the real space position is  $X=1/2U$ ,  $Z=1/2V$ .

**Isomorphous Replacement:** A method for solving the phase problem in x-ray crystal structure determination. The method requires searching for heavy atom derivatives in which the heavy atoms bind at relatively few sites in the unit cell with minimal changes to the structure of the crystal contents (i.e. it must bind isomorphously). If heavy atoms bind isomorphously, the difference between a native and heavy atom derivative structure factor amplitudes (after relative Wilson scaling) is due largely to the presence of the heavy atoms. This difference represents the isomorphous phasing signal. In addition to the isomorphous signal, each heavy atom produces an anomalous phasing signal which arises from the anomalous scattering characteristics of heavy atoms. This is measured within the derivative diffraction data itself as the difference in the intensities of Friedel mates. With some heavy atoms, the anomalous signal is too weak to provide useful phasing information. The isomorphous and anomalous signals each provide phasing information that reduces the phase ambiguity for each structure factor to two possible solutions. One of the phase solutions is correct while the other which varies with each distinct phasing source is incorrect. Because the anomalous and isomorphous phasing signals are distinct, phasing signals from a single heavy atom derivative can be used to resolve the phase ambiguity. This is termed the Single Isomorphous Replacement Anomalous Scattering (SIRAS) method. More often, multiple derivatives are employed to solve phase ambiguity using isomorphous signals alone. This is termed the

Multiple Isomorphous Replacement (MIR) method. The use of anomalous and isomorphous phasing information from a number of derivatives is termed the Multiple Isomorphous Replacement Anomalous Scattering (MIRAS) method. In the winter flounder AFP structure solution, the anomalous and isomorphous signal of two derivatives with the same binding sites were employed for phasing purposes. Because the derivatives shared the same binding sites the phasing information from the two derivatives were very similar. This combined with a weak anomalous signal resulted in relatively poor electron density maps.

**Luzzati Plot:** A statistical plot for the estimate of overall coordinate error. Luzzati statistics are based on the assumption that the disagreement between the observed and calculated structure factor amplitudes, as measured by the  $R_{\text{factor}}$ , is due only to errors in the atomic coordinates. If errors other than these are present, the Luzzati plot gives an over estimation of coordinate error. An example of an error that has a detrimental effect on  $R_{\text{factor}}$  and results in an over estimation of coordinate error is the incompleteness of the structural model.

**Mathews Coefficient:** With the solvent content of protein crystals averaging close to 43%, the number of protein molecules in the asymmetric unit can be estimated from the

molecular weight of the protein (from which the molecules displaced volume is estimated) and the volume of the unit cell.

**Molecular Replacement:** The conventional molecular replacement approach to determining a protein crystal structure with *XPLOR* involves solving successively the rotation and then the translation components which define the orientation of a search model in the unit cell. If greater than one molecule is present in the asymmetric unit, multiple rotation and translation solutions are to be expected. For solving the rotation component, two steps are employed.

Step I. Two Patterson maps are calculated using the search model and the observed structure factors ( $F_{\text{obs}}$ ). The top peaks of the model Patterson map are then selected using criteria including intensities, vector lengths and the total number of peaks. The selected peaks are then systematically rotated with respect to the stationary  $F_{\text{obs}}$  Patterson map over a specified search range. At each rotation sampling (eularian angles  $\theta_1\theta_2\theta_3$ ), the value of the  $F_{\text{obs}}$  Patterson map at the positions of the rotated peaks is checked. Agreement between the two maps for a particular rotation is assessed using the sum of the superimposed peak products. When the rotation is such that maximum agreement between the maps is observed the summation of the peak products is usually at its maximum. In most cases the correct solution is not obvious due to a high signal to noise ratio.

Step II. The top rotation solutions as identified in step I are clustered (i.e. grouped if sufficiently similar) and then refined versus the Patterson correlation coefficient (PC). This stage helps to filter the correct solution. After PC refinement, strong solutions corresponding to the correct rotation orientations for each of the molecules in the asymmetric unit should be apparent.

For solving the translation component, the search molecule is rotated to the top PC refined solution and then translated along the specified unit cell axes for the specified range and increments. At each sampled translation position, the PC value is calculated directly. When the search is completed, the correct solution(s) should be apparent. If more than two molecules in the asymmetric unit are present, a translation search must then be performed to correlate the origins of the molecules.

In the solution of the winter flounder structure, a conventional approach proved unsuccessful and a novel MR strategy was employed. This involved using two search molecules simultaneously in a three parameter translational search strategy. Both PC and  $R_{\text{factor}}$  functions were employed to assess the quality of the translation placements.

**Origin correlation:** In the  $P2_1$  space group, the origin along the Y axis is not defined by crystal symmetry but is arbitrarily chosen. In addition four origins can be chosen

with respect to the X and Z axes (0.0, 0.0, 0.0; 0.0, 0.0, 0.5; 0.5, 0.0, 0.0; 0.5, 0.0, 0.5). In order to compare two structures with different relative origins, a common origin must be selected. A heavy atom derivative provides an opportunity to correlate the origins of structures by providing a common point of reference. To compare the different molecular replacement solutions of the winter flounder AFP crystal structure, origins were correlated by first determining the position of heavy atom sites in a TbNO<sub>3</sub> derivative with respect to each model structure. This was done through the calculation of a difference Fourier map. With the heavy atom positions determined, each model was then shifted by translation to make the heavy atoms sites coincident.

**Patterson Correlation Coefficient (PC):** A crystallographic function measuring the agreement between calculated and observed structure factor amplitudes. The PC function is usually more sensitive to errors than the  $R_{\text{factor}}$  function when a structure is at an early stage of refinement or if the structural model is not complete.

$$PC = \frac{\sum ab - (\sum a \sum b) / N}{\{[\sum a^2 - (\sum a)^2 / N]^{1/2}\} \{[\sum b^2 - (\sum b)^2 / N]^{1/2}\}}$$

$$\text{where } a = |F_o|^2 \text{ and } b = |F_c|^2$$

**Patterson function:** Features of a Patterson density (P) map correspond to the interatomic vectors of the crystal contents.

$$P(xyz) = 1/V \sum_h \sum_k \sum_l |F_{hkl}|^2 \cos 2\pi (hx + ky + lz)$$

**Phasing Power:** A signal to noise measure of the phasing quality of a heavy atom derivative. Good if >2.0, marginal =1.0, below 1.0 useless.

$$\text{Phasing Power} = \text{r.m.s. } |F_H| / E \quad (E = \text{lack of closure})$$

**Phase extension:** In performing solvent flattening, options exist to predict the intensity and phase of reflections absent from the phased set of reflections. When performed in the winter flounder AFP structure determination, the continuity of electron density maps features was greatly improved.

**R<sub>iso</sub>:** Fractional isomorphous difference. A measure of the agreement between native and heavy atom derivative data sets. This parameter gives a good indication of heavy atom substitution if binding is isomorphous.



$R_{\text{iso}} = \sum | |F_{PH}| - |F_P| | / \sum |F_P|$ , where  $F_{PH}$  and  $F_P$  are the structure factors of the derivative and native data sets respectively.

**$R_{\text{cullis}}$**  : A measure of how good the refinement of heavy atom positions has progressed. Based on centric reflections only.

$R_{\text{cullis}} = \sum | |F_H| - (|F_{PH}| - |F_P|) | / \sum |F_H|$ , where  $F_H$  is the calculated heavy-atom structure factor;  $F_{PH}$  and  $F_P$  are the structure factors of the derivative and native data sets respectively. Typical values for derivatives range from 0 to 60%.

**$R_{\text{factor}}$** : A measure of the agreement between observed and calculated (model) structure factor amplitudes.

$R_{\text{factor}} = \sum |F_O| - |F_C| / \sum |F_C|$  where  $F_O$  and  $F_C$  are the observed and calculated structure factor amplitudes respectively.

**$R_{\text{free}}$** : A statistical measure of the accuracy and progress of crystallographic refinement. Same as  $R_{\text{factor}}$  but only structure factor amplitudes that are excluded from refinement are included in the calculation. Before commencement of refinement 10-15% of reflections are randomly set aside for monitoring of  $R_{\text{free}}$ . As the structure is

refined versus the other 90% of the reflections, it is expected that  $R_{\text{free}}$  should also improve. If at any stage of refinement  $R_{\text{free}}$  does not improve while the  $R_{\text{factor}}$  does, the accuracy of the structure is not improving and refinement should be stopped.

**$R_{\text{sym}}$ :** A measure of the agreement between equivalent reflections within a single data set.

$$R_{\text{sym}} = \frac{\sum |I_i - \langle I \rangle|}{\sum I_i}$$
 where  $I_i$  is the intensity of an individual reflection and  $\langle I \rangle$  is the mean intensity of that reflection.

**Refinement:** Optimization of positional (and also B-factors) parameters against a target function using a powell minimization. Commonly the target function is  $R_{\text{factor}}$  or the PC. When the ratio of parameters to be refined to the number of observable to be refined against is too great, as is the case for most protein structure refinements, observables are supplemented with geometrical observations from other protein structures. These are included in the form of stereo-chemical constraints. A weighting scheme is employed to place stereo-chemical and x-ray terms on similar scales. Various types of refinement can be performed. These include:

i. **Rigid body:** Refinement of three rotation and three translation parameters for each identified object. The geometry and distances between the atoms within an object are held fixed.

ii. **Positional:** Refinement of three translation parameters (x,y,z) for each specified atom.

iii. **Simulated Annealing:** Kinetic energy is given to the macromolecule under refinement in a manner analogous to a dynamics simulation. This serves to increase the radius of convergence of the refinement procedure. A high temperature is specified from which level the structure is slowly cooled. Following this cooling, 30-40 cycles of positional refinement are performed.

**Restrained isotropic B-factor refinement:** To prevent B-factor parameters from becoming unrealistically large or negative due to a limited number of observables, the B-factors of atoms under refinement can be restrained. In *XPLOR*, B-factors of non backbone atoms are restrained to a default level of  $2.5\sigma$  of the total non-backbone atom B-factor variation. The restraint level for backbone atoms is  $1.5\sigma$ . The more relaxed restraint for non-backbone atoms reflects the higher mobility of side chains atoms.

**Self Rotation search:** Cross rotation search between an  $F_o$  Patterson map and itself. Used to identify the non-crystallographic symmetry axes relating the different molecules in the asymmetric unit. Performed in the same manner as a cross rotation search between  $F_c$  and  $F_o$  Patterson maps.

**Simulated annealing 'omit' maps:** A technique used to remove model bias from an electron density map. Model bias refers to errors incorporated in a model structure during refinement that are difficult to detect and remove. This problem is more severe after the refinement of lower resolution structures where the ratio of refined parameters to observables is low. The technique involves omitting a small portion of a structure and performing one cycle of SA refinement. The kinetic energy imparted to the remainder of the structure is expected to randomize atomic coordinates to remove imbedded bias. If the omitted section was in error, the correct position of atoms should be revealed.

**Solvent Flattening:** A means of improving poor electron density maps arising from errors in phase estimation. The rationale for the method is based on correcting electron density map features that are unrealistic. For example, solvent regions should be relatively featureless due to solvent disorder, and no regions of the map should have electron densities below zero. The method of solvent flattening involves distinguishing

automatically or manually the solvent boundaries in the unit cell. For automatic determination an estimate of the solvent content of the crystal is required. Electron density regions corresponding to solvent regions are then flattened to remove unrealistic features through a modification of phases. Phases are also modified to remove negative electron density from all regions of the map. The result of solvent flattening is that the overall phasing is improved such that electron density corresponding to regions of the protein structure are more interpretable. Such a result was observed in the winter flounder AFP structure determination.

## APPENDIX A XPLOR SHELL SCRIPT FOR A MULTI-PARAMETER SEARCH

```

remarks using two search models oriented with best rotations from the single
remarks molecule molecular replacement search studies, a multiparamter
remarks translation search is performed.

parameter @/usr/local/xplor/toppar/param19x.pro end { read parameters }

structure @/users/mmcu/yang/flid/ala370.psf end { read structure file }
coor @/users/mmcu/yang/flid/ala370min3.pdb

coor rotate euler ( 245.74 64.54 255.63 ) end {this is the top rotation solution }
                                         {for the helix up direction }

{ in order to generate the 2nd molecule we first renumber the 1st }
vector do ( resid=encode(decode(resid)+1000)) ( all )

structure @/users/mmcu/yang/flid/ala370.psf end {read structure file for second }
                                         { molecule }
coor @/users/mmcu/yang/flid/ala370min3.pdb
coor rotate euler (107.14 63.14 283.90 ) selection= ( resid 1:37 ) end
{above is the top rotation solution solution for the helix down direction}

vector do (b=15.0) (all)
evaluate ($wa=10000.) { this is the weight for the XREF energy term }
                    { in this case it is arbitrary since we're not }
                    { combining it with another energy term }

xrefin
{===>} { unit cell for crystal }
a=37.9749 b=36.7800 c=21.7373 alpha=90. beta=101.0812 gamma=90.

{===>}
symmetry=(x,y,z) { operators for crystal symmetry P2(1) }
symmetry=(-x,y+1/2,-z)

SCATter ( chemical C* )
2.31000 20.8439 1.02000 10.2075 1.58860 .568700 .865000 51.6512 .215600
SCATter ( chemical N* )
12.2126 .005700 3.13220 9.89330 2.01250 28.9975 1.16630 .582600 -11.529
SCATter ( chemical O* )
3.04850 13.2771 2.28680 5.70110 1.54630 .323900 .867000 32.9089 .250800
SCATter ( chemical S* )
6.90530 1.46790 5.20340 22.2151 1.43790 .253600 1.58630 56.1720 .866900
SCATter ( chemical P* )
6.43450 1.90670 4.17910 27.1570 1.78000 0.52600 1.49080 68.1645 1.11490
SCATter ( chemical FE* )
11.1764 4.61470 7.38630 0.30050 3.39480 11.6729 0.07240 38.5566 0.97070

{===>}
nreflections=5000
reflection @/users/mmcu/yang/flid/afphp6cx.hkl { read reflections }
{===>}
resolution 10.0 3.0 { resolution range }

```

```

reduce
do amplitude ( fobs = fobs * heavy(fobs - 2.0*sigma) ) { sigma cutoff }
fwind=0.1=100000

{===>}
method=fft
fft
memory=2000000 { fft method with memory statement }
end

wa=$wa
target=E2E2 { specify target used for both PC-refinement }
{ and translation search }
mbins=20 { number of bins used for E calculation }

tolerance=0. lookup=false { this makes the minimizer happy }

end

flags exclude * include xref end { only use XREF energy term }

{===>} set display=/usr/tmp/yang/sicheri/sldsrchPC3.list1 end

coor fractionalize end

evaluate ($1=0)
evaluate ($2=0)
evaluate ($3=0)

set echo=off end
set message=off end

coor translate vector=( 0.0 0.0 0.250 ) selection=(all) end
{position AFP molecules on the diagonal ridges}

coor copy end

{do loop over s1 24 steps}
while ($1 < 24) loop s1
evaluate ($a=$1/48)
evaluate ($ma=-$a)

evaluate ($2=0)
{do loop over s2 48 steps}
while ($2 < 48) loop s2
evaluate ($b=$2/48)
evaluate ($mb=-$b)

evaluate ($3=0)
{do loop over y2 37 steps}
while ($3 < 37) loop y2
evaluate ($c=$3/37)

coor translate vector=( $a 0.0 $ma ) selection=( resid 1:37 ) end
coor translate vector=( $b 0.0 $mb ) selection=( resid 1001:1037 ) end

```

```
coor translate vector=( 0.0 $c 0.0 ) selection=( resid 1001:1037 ) end

coor orth end
  xrefin
  update
  print target
end
energy end
evaluate ($arg= 1-$xref/$wa)
display $1 $2 $3 $arg
coor frac end

      coor swap end
      coor copy end

evaluate ($3=$3+1)
end loop y2

evaluate ($2=$2+1)
end loop sl2

evaluate ($1=$1+1)
end loop sl1
stop
```



## APPENDIX B XPLOR SHELL SCRIPT FOR THE REFINEMENT OF CLUSTERED SOLUTIONS

```
remarks      read in the top clustered solutions of sldsrchPC3.list
remarks
remarks      this contains clustered solutions obtained using a
remarks      a modification of the peak search algorithm from the PHASES
remarks      package
remarks
remarks      then rigidbody minimize
remarks      then positional refine
remarks      then SA refine
remarks
remarks      Dihedral constraints to maintain the helical nature of the AFP
remarks      are also included but have little influence on the outcome of results
remarks

set display=/usr/tmp/yang/sicheri/ref.list8 end

parameter @/usr/local/xplor/toppar/param19x.pro end

structure @/users/mmcu/yang/flid/ala370.psf end
coord @/users/mmcu/yang/flid/ala370min3.pdb
coord rotate euler (107.14 63.14 283.90 ) end

{ in order to generate the 2nd molecule we first renumber the 1st }
  vector do ( resid=encode(decode(resid)+1000)) ( all )

structure @/users/mmcu/yang/flid/ala370.psf end
coord @/users/mmcu/yang/flid/ala370min3.pdb
coord rotate euler (107.14 63.14 283.90 ) selection= ( resid 1:37 ) end

vector do (b=15.0) (all)
vector do (q=1.0) (all)

xrefin

{ unit cell for crystal }
  a=37.9749 b=36.7800 c=21.7373 alpha=90. beta=101.0812 gamma=90.

{ operators for crystal symmetry P2(1) }
  symmetry=(x,y,z)
  symmetry=(-x,y+1/2,-z)

SCATter ( chemical C* )
  2.31000 20.8439 1.02000 10.2075 1.58860 .568700 .865000 51.6512 .215600
SCATter ( chemical N* )
  12.2126 .005700 3.13220 9.89330 2.01250 28.9975 1.16630 .582600 -11.529
SCATter ( chemical O* )
  3.04850 13.2771 2.28680 5.70110 1.54630 .323900 .867000 32.9089 .250800
SCATter ( chemical S* )
  6.90530 1.46790 5.20340 22.2151 1.43790 .253600 1.58630 56.1720 .866900
SCATter ( chemical P* )
  6.43450 1.90670 4.17910 27.1570 1.78000 0.52600 1.49080 68.1645 1.11490
SCATter ( chemical FE* )
```

```

11.1764 4.61470 7.38630 0.30050 3.39480 11.6729 0.07240 38.5566 0.97070

nreflections=5000
reflection @/users/mmccu/yang/fld/afphp6cx.hkl { read reflections }

resolution 10.0 3.0 { resolution range }
reduce
do amplitude ( fobs = fobs * heavy(fobs - 2.0*sigma) ) {sig cut}
fwind=0.1=10000

method=fft
fft
memory=2000000 { fft method with memory statement }
end

evaluate ($wa=20118)
wa=$wa

end

coor fractionalize end

coor translate vector=( 0.0 0.0 0.250 ) selection=( resid 1:37) end
coor translate vector=( 0.0 0.0 0.250 ) selection=( resid 1001:1037) end

coor orthogonalize end

coor copy end { save these coordinates for }
{ the start of following cycles }

!!!!!!!!!!!! build the phi and psi constraints!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!

evaluate($77=1.) {impose dihedral constraints}
While ($77 < 37.) loop dihedrall { for whole molecule in }
evaluate($78=$77+1.) { alpha helical form }
constraints dihedral
assign (resid $77 and name n)
(resid $77 and name ca)
(resid $77 and name c)
(resid $78 and name n) 20.0 -47.5 2.5 2
assign (resid $77 and name c)
(resid $78 and name n)
(resid $78 and name ca)
(resid $78 and name c) 20.0 -57.4 2.5 2
scale=1.0
end
end
evaluate($77=$77+1.)
end loop dihedrall

```

```

evaluate($87=1001.)                                {impose dihedral constraints}
While ($87 < 1037.) loop dihedral2                { for whole molecule in }
  evaluate($88=$87+1.)                            { alpha helical form }
  constraints dihedral
  assign (resid $87 and name n)
    (resid $87 and name ca)
    (resid $87 and name c)
    (resid $88 and name n) 20.0 -47.5 2.5 2
  assign (resid $87 and name c)
    (resid $88 and name n)
    (resid $88 and name ca)
    (resid $88 and name c) 20.0 -57.4 2.5 2
  scale=1.0
  end
  end
  evaluate($87=$87+1.)
end loop dihedral2

!***** loop PEAK *****

!***** part A position the molecules *****

evaluate ($counter=0) { this counts the elements in a line }
evaluate ($number=0) { this counts the number of lines }

for $1 in ( @/users/mmcu/yang/1d/sldsrchPC3.peak8 ) loop peak

  evaluate ($counter=$counter+1) { this counts the elements in a line }

  if ($counter=1) then evaluate($sld1=$1)
  elseif ($counter=2) then evaluate($sld2=$1)
  elseif ($counter=3) then evaluate($t2= $1)

  elseif ($counter=4) then
    evaluate ($number=$number+1) { this counts the number of lines }
    evaluate ($pkght=$1)
    evaluate ($counter=0)

    evaluate ($a=$sld1/48.) { fractionalize sld1 sld2 and t2}
    evaluate ($ma=-$a)
    evaluate ($b=$sld2/48.)
    evaluate ($mb=-$b)
    evaluate ($c=$t2/37.)
    evaluate ($mc=-$c)

  coor fractionalize end

  coor translate vector=( $a 0.0 $ma ) selection=( resid 1:37 ) end
  coor translate vector=( $b 0.0 $mb ) selection=( resid 1001:1037 ) end
  coor translate vector=( 0.0 $c 0.0 ) selection=( resid 1001:1037 ) end

  coordinates orthogonalize end

  xref
  update
  print rfactor
  evaluate ($r0=$result)
  end

```

```
!write coordinates          output= /usr/tmp/yang/sicheri/refine.pdb8a end
```

```
!***** PART A PRIME WRITE THIS REFERENCE POINT
```

```
coor swap end             { fit coordinates to starting structure in }
  vector do (vx=x) ( all ) { order to measure the orientation of the }
  vector do (vy=y) ( all ) { PC-refined structure }
  vector do (vz=z) ( all ) { the arrays vx, vy, vz are used as temporary}
                          { stores in order to keep the starting }
                          { coordinates }

```

```
coor fit
selection= (resi 1:37)
end
```

```
evalutate ($xmc=$x)      { the COOR FIT statement stores the angles }
evalutate ($ymc=$y)      { in the symbol $theta1, $theta2, $theta3 }
evalutate ($zmc=$z)      { $x,$y,$z }
evalutate ($th1mc=$theta1)
evalutate ($th2mc=$theta2)
evalutate ($th3mc=$theta3)
```

```
coor fit
selection= (resi 1001:1037)
end
```

```
evalutate ($xmd=$x)
evalutate ($ymd=$y)
evalutate ($zmd=$z)
evalutate ($th1md=$theta1)
evalutate ($th2md=$theta2)
evalutate ($th3md=$theta3)
```

```
vector do (x=vx) ( all )
vector do (y=vy) ( all )
vector do (z=vz) ( all )
```

```
{ print information: orientation of rotation }
{ function peak ($t1, $t2, $t3), orientation }
{ after PC-refinement ($theta1, $theta2, }
{ $theta3), index of the rotation function, }
{ rotation function value, PCs for initial, }
{ and rigid body refined structures. }
```

```
coor swap end
```

```
!write coordinates          output= /usr/tmp/yang/sicheri/refine.pdb8b end
```

```
!***** PART B rigid body minimize vs RESIDUAL *****
```

```
xref
  mbins=20             { number of bins used for E calculation }
  tolerance=0.0
  lookup=false        { this makes the minimizer happy }
end
```

```

flags exclude * include xref end          { only use XREF energy term }

      minimize rigid
      group=(resid 1:37)
      group=(resid 1001:1037)
      nstep=150
      drop=40
      end

xref
update
print rfactor
evaluate ($r1=$result)
end

!***** PART C minimize with MOLECULAR CONSTRAINTS *****

flags exclude *
      include xref bond angle dihe impr vdw elec pvdw pele cdih end

parameter nbonds repel=0.89 rcon=4. end end
      minimize rigid
      group=(resid 1:37)
      group=(resid 1001:1037)
      nstep=40
      drop=40
      end

      minimize powell
      nstep=40
      drop=40
      end

parameter nbonds repel=0.0 end end
      minimize powel
      nstep=40
      drop=100.
      end

xref
update
print rfactor
evaluate ($r2=$result)
end

!write coordinates                          output= /usr/tmp/yang/sicheri/refine.pdb8c end

!***** PART D simulated annealing routine *****

xref
      tolerance=0.2
      lookup=true
      end

set seed=4323324368 end

```

```

vector do (vx=maxwell(2000.0)) (all)
vector do (vy=maxwell(2000.0)) (all)
vector do (vz=maxwell(2000.0)) (all)
vector do (fbeta=100.0) (all)
evaluate ($666=2000)
while ($666 > 300.0 ) loop heat
dynamics verlet
timestep = .0005
nstep=50
iasvel=current
nprint=5 iprfrq=0
tcoupling=true tbath=$666
end
evaluate ($666=$666-25)
end loop heat

xrefine
tolerance=0.0
lookup=false
end

minimize powell
nstep=120
drop=10
end

write coordinates                               output= /usr/tmp/yang/sicheri/refine.pdb8 end

xref
update
print rfactor
evaluate ($r3=$result)
end

!*****then compare each of the helices to the original structure*****

coor swap end      { fit coordinates to starting structure in }
vector do (vx=x) ( all ) { order to measure the orientation of the }
vector do (vy=y) ( all ) { PC-refined structure }
vector do (vz=z) ( all ) { the arrays vx, vy, vz are used as temporary}
                    { stores in order to keep the starting }
                    { coordinates }

coor fit
selection= (resi 1:37)
end

evalutate ($xma=$x) { the COOR FIT statement stores the angles }
evalutate ($yma=$y) { in the symbol $theta1, $theta2, $theta3 }
evalutate ($zma=$z) { $x,$y,$z }
evalutate ($th1ma=$theta1)
evalutate ($th2ma=$theta2)

```

```

evalutate ($th3ma=$theta3)

  coor fit
  selection= (resi 1001:1037)
  end

evalutate ($xmb=$x)
evalutate ($ymb=$y)
evalutate ($zmb=$z)
evalutate ($th1mb=$theta1)
evalutate ($th2mb=$theta2)
evalutate ($th3mb=$theta3)

vector do (x=vx) ( all )
vector do (y=vy) ( all )
vector do (z=vz) ( all )

      { print information: orientation of rotation }
      { function peak ($t1, $t2, $t3), orientation }
      { after PC-refinement ($theta1, $theta2, }
      { $theta3), index of the rotation function. }
      { rotation function value, PCs for initial, }
      { and rigid body refined structures. }

coor copy end

display $sld1 $sld2 $t2 $r0 $r1 $r2 $r3

display $xmc $ymc $zmc $th1mc $th2mc $th3mc
display $xmd $ymd $zmd $th1md $th2md $th3md

display $xma $yma $zma $th1ma $th2ma $th3ma
display $xmb $ymb $zmb $th1mb $th2mb $th3mb

set echo=off end
set message=off end

  end if
end loop peak

stop

```

## REFERENCES

- Ahmed, A.I., Osuga, D.T., & Feeney, R.E. (1973). Antifreeze glycoprotein from an Antarctic fish. *J. Biol. Chem.* **248**, 8524-8527.
- Ananthanarayanan, V. S. (1989). Antifreeze proteins: structural diversity and mechanism of action. *Life Chemistry Rep.* **7**, 1-32 (1989).
- Ananthanarayanan, V.A. and Hew, C.L. (1977). Structural studies on the freezing point depressing protein of the winter flounder *Pseudopleuronectes americanus*. *Biochem. Biophys. Res. Comm.* **74**, 685-689 (1977).
- Baker, E.N. and Hubbard, R.E. (1984). Hydrogen bonding in globular proteins. *Prog. Biophys. molec. Biol.* **44**, 97-179.
- Brown, R.A., Yin, Y., Burcham, T.S., and Feeney, R.E. (1985). Direct evidence for antifreeze glycoprotein adsorption onto an ice surface. *Biopolymers* **24**, 1265-1270 (1985).



Brunger, A.T. (1992). *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR*. New Haven, Yale Univ. Press, CT, USA.

Brunger, A.T. (1992). Free R Value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-475.

Brunger, A.T. (1990). Extension of Molecular Replacement: a new search strategy based on patterson correlation refinement. *Acta Cryst.* **A46**, 46-57.

Chakrabarty, A., Ananthanarayanan, V.A., and Hew. C.L. (1989). Structure-function relationships in a winter flounder antifreeze polypeptide. *J. Biol. Chem.* **264**, 11307-11312.

Chakrabarty, A., Yang, D.S.C., and Hew. C.L. (1989). Structure-function relationships in a winter flounder antifreeze polypeptide. *J. Biol. Chem.* **264**, 11313-11316.

Chakrabarty, A. and Baldwin, R.L. (1995) Stability of  $\alpha$ -helices. *Adv. Prot. Chem.*

Chakrabarty, A. and Hew, C.L. (1991) The effect of enhanced  $\alpha$ -helicity on the activity of a winter flounder antifreeze polypeptide. *Eur. J. Biochem.* **202**, 1057-1063.

Chao, H., Davies, P.L., Sykes, B.D. and Sonnichsen, F.S (1993). Use of proline mutants to help solve the NMR solution structure of type III antifreeze protein. *Protein Science* **259**, 1154-1157.

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

Chou, K.C. (1992). Energy-optimized structure of antifreeze protien and its binding mechanism. *J Mol. Bio.* **223**, 509-517.

Chou, P.Y. and Fasman, G. D. (1974). Prediction of protein conformation. *Biochemistry* **13**, 222-245.

Davies, P.L., Roach, A.H. and Hew, C.L. (1982). DNA sequence coding for an antifreeze protein precursor from winter flounder. *Proc. Natl Acad. Sci. USA* **79**, 335-339.

DeVries, A.L. (1984). Role of glycopeptides and peptides in inhibition of crystallization of water in polar fishes. *Phil. Trans. R. Soc. Lond. B* **304**, 575-588.

DeVries, A.L. (1983). Antifreeze peptides and glycopeptides in cold-water fishes. *Ann. Rev. Physiol.* **45**, 245-260.

Duman, J.G. and Olser, T.M. (1993). Thermal hysteresis protein activity in bacteria, fungi, and phylogenetically diverse plants. *Cryobiology* **30**, 322-328.

Duman, J.G., Patterson, J.L., Kozak., J.J., and DeVries, A.L.(1980). Isopiestic determinations of water binding by fish antifreeze glycoproteins. *Biochim. Biophys. Acta.* **626**, 332-336.

Ewart, V.K., Rubinsky, B. & Fletcher, G.L. (1992). Structural and functional similarity between fish antifreeze proteins and calcium-dependant lectins. *Biophys. Res. Commun.* **185**, 335-340.

Ewart, V.K., Fletcher, G.L. (1993). Herring antifreeze protein: primary structure and evidence for a C-type lectin evolutionary origin. *Mol. Mar. Biol. Biotechnol.* 2(1): 20-27.

Feeney, R.E. and Yeh, Y (1989). Antifreeze Proteins from fish blood. *Adv. Protein chem.* 32, 191-282

Fourney, R.M., Joshi, S.B., Kao, M.H., & Hew, C.L. (1984). Heterogeneity of antifreeze polypeptides from the Newfoundland winter flounder, *Pseudopleuronectes americanus*. *Can.J. Zool.* 62, 28-33.

Furey, W. & Swaminathan, S. (1990). *American Crystallographic Association Meeting Abstracts*, New Orleans LA, PA33.

Haschemeyer, A.E.V., Guschlbauer, W., and DeVries, A.L. (1977). Water binding by antifreeze glycoproteins from antarctic fish. *Nature* 269, 87-88.

Hew, C.L., Chakrabartty, A., and Yang, D.S.C. *Integration and Control of Metabolic Processes: Pure and Applied Aspects*, O.L. Kon *et al.* Eds. (ICSU Press, 1987) pp. 299-309.

Hew, C.L., Wang, N.C., Yan, S., Cai, H., Sclater, A. and Fletcher, G. L. (1986). Biosynthesis of antifreeze polypeptides in the winter flounder. *Eur. J. Biochem.* **160**, 267-272.

Hew, C. L. and Yang, D. S. C. (1992). Protein interaction with ice. *Eur. J. Biochem.* **203**, 33-42.

Hill, C.P., Anderson, D.H., Wesson, L., Degrado W.F., and Eisenberg, D. (1990). Crystal structure of  $\alpha 1$ : implications for protein design. *Science* **249**, 543-546.

Jones, T.A., Zou, J.Y., Cowan, S. W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta. crystallogr.* **A47** 110-119.

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

Knight, C.A., Cheng, C.C. and DeVries, A.L. (1991). Adsorption of  $\alpha$ -helical antifreeze proteins on specific ice crystal surface planes. *Biophys. J.* **59**, 409-418.

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

Knight, C.A. and DeVries, A.L. (1989). Melting inhibition and superheating of ice by an antifreeze glycopeptide. *Science* **245**, 505-507.

Knight, C.A., and Duman J.G. (1986). Inhibition of recrystallization of ice by insect thermal hysteresis proteins: a possible cryo protective role. *Cryobiology* **23**, 256-262.

Lal, M., Clark, A.H., Lips, A., Ruddock, J.N., and White, D.N.J. (1993). Inhibition of ice crystal growth by preferential peptide adsorption: a molecular modelling study. *Faraday Discuss.* **95**, 299-306.

Laursen, R.A., Wen, D. & Knight, C.A. (1994). Enantioselective adsorption of the D- and L forms of an  $\alpha$ -helical antifreeze polypeptide to the {2021} planes of ice. *J. Am. Chem. Soc.* **116**, 12057-12058.

Luzzati, V. (1952). Statistical treatment of errors in the determination of crystal structures. *Acta Cryst.* **5**, 802-810. Translated March 1990 by Eric Fauman, UC San Francisco, Dept. Biochemistry and Biophysics.

Madura, J.D., Wierzbicki, A., Harrington, J.P., Maughon, R.H., Raymond, J.A., and Sikes, C.S. (1994). Interactions of the D- and L-Forms of winter flounder antifreeze peptide with the {201} planes of ice. *J. Am. Chem. Soc.* **116**, 417-418.

Marqusee, S. and Baldwin, R.L. (1987). Helix stabilization by Glu- Lys<sup>+</sup> salt bridges in short peptides of *de novo* design. *Proc. Natl. Acad. Sci. USA.* **84**, 8898-8902 (1987)

McDonald, S. M. Brady, J.W. and Clancy, P. (1993). Molecular dynamics simulations of a winter flounder antifreeze polypeptide in aqueous solution. *Biopolymers* **33**, 1481-1503.

McGregor, M. J., Islam, S. A., and Sternberg, M. J. E. (1987). Analysis of the relationship between side-chain conformation and secondary structure in globular proteins. *J. Mol. Biol.* **198**, 295-310.

- Megaw, H.D. (1934). Cell dimensions of ordinary and 'heavy' ice. *Nature, Lond.* **134**, 900-901.
- Merutka, G. and Stellwagen, E. (1991). Effect of amino acid ion pairs on peptide helicity. *Biochemistry* **30**, 1591-1594.
- Ng, N.F. & Hew, C.L. (1992). Structure of an antifreeze polypeptide from sea raven. *J. Biol. Chem.* **267**, 16069-16075.
- Pickett, M, Scott, G., Davies, P., Wang, N., Joshi, S. and Hew, C.(1984). Sequence of an antifreeze protein precursor. *Eur. J. Biochem.* **143**, 35-38.
- Piela, L., Nemethy, G., Scheraga, H.A. (1987). Conformational constraints of amino acid side chains in  $\alpha$ -helices. *Biopolymers* **26**, 1273-1286.
- Presta, L.D. and Rose, G.D. (1988). Helix signals in proteins. *Science* **240**, 1632-1641.
- Raymond, J.A., and DeVries, A.L. (1977). Adsorption inhibition as a mechanism of freezing resistance in polar fishes. *Proc. Natl. Acad. Sci. USA* **74**, 2589-2593.



Rao, B.N.N. and Bush, C.A. (1987). Comparison by  $^1\text{H-NMR}$  spectroscopy of the conformation of the 2600 Dalton antifreeze glycopeptide of polar cod with that of high molecular weight antifreeze glycoprotein. *Biopolymers* **26**,1227-1244.

Richardson, J.S. and Richardson, D.C. (1988). Amino acid preferences for specific locations at the ends of  $\alpha$  helices. *Science* **240**, 1648-1652.

Schafmeister, C.E., Miercke, L.J.W., and Stroud, R.M. (1993). Structure at 2.5Å of a designed peptide that maintains solubility of membrane proteins. *Science* **262**, 734-738.

Sicheri, F.S. and Yang, D.S.C. (1995). Ice binding structure and mechanism of an antifreeze protein from winter flounder. *Nature (in press)*.

Sonnichsen, F.D, Sykes, B.D. and Davies, P.L. (1995). Comparative modeling of the three-dimensional structure of type II antifreeze protein. *Protein Science* **4**, 460-471.

Sonnichsen, F.D, Sykes, B.D., Chao, H., and Davies, P.L. (1993). The non-helical structure of antifreeze protein type III. *Science* **259**, 1154-1157.

Stout, H.G. & Jensen, L.H. (1989). X-ray structure determination. A practical guide. (2nd ed.), John Wiley and Sons, New York, pp. 178-222.

Terwilliger, T.C, Kim, S.H. and Eisenberg, D. (1987). Generalized method of determining heavy-atom positions using the difference Patterson function. *Acta Crystallogr. Sect. A: Found. Crystallogr.* **43**, 1-5.

Thanki N., Thornton, J.M., and Goodfellow J.M. (1988). Distributions of water around amino acid residues in proteins. *J Mol. Biol.* **202**, 637-657.

Tong, H, Berghuis, A.M., Chen, J. and Luo, Y. (1994). ASIR: an automatic procedure for determining solvent structure in protein crystallography. *J. Appl. Cryst.* **27**, 421-426.

Urrutia, M.E., Duman, J.G. and Knight C.A. (1992). Plant thermal hysteresis proteins. *B.B.A.* **1121**, 199-206.

Wang, B.-C. (1985). Resolution of phase ambiguity in macromolecular crystallography. *Methods Enzymol.* **115**, 90-112.

Wen, D. and Laursen, R.A. (1992a). Structure-function relationships in an antifreeze polypeptide. *J. Biol. Chem.* **267**, 14102-14108.

Wen, D., and Laursen, R.A. (1992b). A model for binding of an antifreeze polypeptide to ice. *Biophys. J.* **63**, 1659-1662.

Wilson, P.W., Beaglehole, D. and DeVries, A.L. (1993). Antifreeze glycopeptide adsorption on single crystal ice surfaces using ellipsometry. *Biophys. J.* **64**, 1878-1884.

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

Yang, D.S.C., Chung, Y.J., Chen, P., Rose, J.P., and Hew, C.L. (1986). Single crystals of an antifreeze polypeptide from winter flounder. *J. Mol. Biol.* **189**, 725.

Yang, D.S.C., Sax, M., Chakrabarty, A., and Hew, C.L. (1988). Crystal structure of an antifreeze polypeptide and its mechanistic implications. *Nature* **333**, 232-237.