J. Biochem. 126, 387-394 (1999)

# Determination of the Solution Structure of the N-Domain Plus Linker of Antarctic Eel Pout Antifreeze Protein RD3

Kazunori Miura,* Satoru Ohgiya,† Tamotsu Hoshino,† Nobuaki Nemoto,‡ Masato Odaira,† Katsutoshi Nitta,* and Sakae Tsuda†,[1]

*Division of Biological Sciences, Graduate School of Science, Hokkaido University, Kita 8 Nishi 5, Kita-ku, Sapporo 060-0808; †Bioscience and Chemistry Division, Hokkaido National Industrial Research Institute (HNIRI), 2-17-2-1 Tsukisamu-Higashi, Toyohira-Ku, Sapporo 062-8517; and ‡Varian Japan, Varian Japan Sumitomo Shibaura Bldg., 4-16-36 Shibaura, Minato-ku, Tokyo 108-0023

Received May 10, 1999; accepted June 4, 1999

**RD3, a new antifreeze protein (AFP) extracted from antarctic eel pout is a single polypeptide divided into homologous N-terminal (residues Asn[1]-Glu[64]) and C-terminal (residues Ser[74]-Glu[134]) domains, each of which has a high sequence identity with Type III AFP. A 9-residue linker (-D[65]GTTSPGLK[73]-) connects these two domains in tandem and is thought to play a significant role in defining the nature of the intact molecule. The present paper shows for the first time the solution structure and preliminary $^{15}$N-NMR backbone dynamics data of the N-domain plus the linker of recombinant RD3 protein (RD3-N1: residues 1-73) by employing homo- and heteronuclear multidimensional NMR spectroscopy. Forty converged structures of RD3-N1 were successfully calculated by using a total of 958 NMR-derived structural restraints. It was found that the N-domain of RD3-N1 has a globular form comprising six $\beta$-strands, three type III turns, and several loops, which stabilize a flat, ice-binding site formed on one side of this domain. Further, the linker portion appears to have a definitive structure, which is independent of the globular N-domain. This definitive linker is roughly divided into two short strands, -D[65]GTTSP[70]- and -G[71]LK[73]-, which are bent around -T[67]TSPG[71]- at an angle of approximately 60°. This bending motif of the linker may function to orient the two ice-binding sites of the N- and C-domains of RD3 in the same direction, leading to their simultaneous interactions with the ice crystal surface.**

**Key words: antifreeze protein RD3, linker, NMR structure.**

Antifreeze protein (AFP) exists in blood plasma and body fluids of some marine fishes living in polar and near polar waters where the temperature is subzero, frequently as low as −1.9°C. AFP has a unique ability to adsorb to a hexagonal ice crystal to inhibit its growth, which results in depression of the freezing point of water, leading to protection of tissues from freezing injury (*1*). Such depression of the freezing point is 500 times higher than that of the colligative salts on a molar basis and occurs non-colligatively because of AFP-induced thermal hysteresis: a disparity between the melting and freezing points of the solution. AFP also affects the morphology of the ice crystal growth: it creates bipyramidal ice crystals in the solution (*2*). The thermal hysteresis and bipyramidal ice crystals are commonly observed for various types of AFP.

Among the variety of AFPs, the most advanced knowledge has been obtained for fish AFPs, which are grouped into four types (I, II, III, and IV). Type I AFP, isolated from some species of flounder and sculpin, is composed of alanine-rich repetitions of an eleven amino acid unit, Thr-$(Xxx)_2$-Asx-$(Xxx)_7$, where Xxx is mostly alanine (*3*), which forms an amphipathic $\alpha$-helical structure (*4, 5*). Type I AFP binds selectively to the ({2 0 (−2) 1}) bipyramidal planes in the direction of <0 1 (−1) 2>, implying that the 16.5 Å spacing of the $i$, $i+11$ threonine residues is essential for the specific interaction with ice, because it matches 16.7 Å distance of the water molecules on the ice surface (*6*).

Type II AFP, isolated from sea raven, smelt, and herring (*7, 8*), is a cysteine-rich (five disulfide bonds) protein having a high sequence identity with the carbohydrate-recognition domain of $Ca^{2+}$-dependent (C-type) lectin. Type II AFPs of smelt and herring require $Ca^{2+}$ ions for their activity, while that of sea raven does not (*8, 9*). The NMR structure of sea raven AFP is composed of two $\alpha$-helices and nine $\beta$-strands in two $\beta$-sheets (*10*). The ice-binding site of Type II AFP is considered to interact with the ({1 1 (−2) 1}) plane of the ice crystal (*11*).

Type III AFP, isolated from ocean pout and arctic and antarctic eel pout (*12, 13*) has no bias in specific amino acids. Sönnichsen *et al.* have determined the NMR solution structure of one isoform of Type III AFP (*i.e.*, QAE isoform) (*14*). This structure comprises several short and

[1] To whom correspondence should be addressed. E-mail: tsuda@hniri.go.jp, Phone: +81-857-8912, Fax: +81-857-8983

Abbreviations: AFP, antifreeze protein; HSQC, hetero nuclear single quantum coherence spectroscopy; DQF-COSY, double-quantum filtered correlation spectroscopy; NOESY, nuclear Overhauser enhancement spectroscopy; TOCSY, total correlation spectroscopy; SCUBA, simulated cross-peaks under bleached alphas; DSS, 2,2-dimethyl-2-silapentane-5-sulfonate sodium salt; SA, simulated annealing.

irregular $\beta$-sheets and one helical turn, which stabilize a flat plane constructed on one side of the molecule. The polar side-chain atoms of $Glu^{9}$, $Asn^{14}$, $Thr^{15}$, $Thr^{18}$, and $Gln^{44}$ are located on this plane so as to bind complementarily with the water molecules spaced regularly on the prism ({1 0 (−1) 0}) plane of the hexagonal ice crystal. Jia *et al.* and Yang *et al.* revealed X-ray crystal structures of the QAE isoforms (QAE and HPLC-3), whose overall folding topologies and ice-binding residues are almost identical to those of the NMR structure (*14-16*). Sönnichsen *et al.* and Jia *et al.* pointed out that $Asn^{14}$, a non-planar residue located on the ice-binding site of Type III AFP, is involved in an initial protein-ice interaction at the intersection of the prism and basal planes. DeLuca *et al.* determined the X-ray structures of six Type III AFP mutants whose $Ala^{16}$ located at the center of ice-binding surface was replaced with several different amino acids (*17*). Sönnichsen *et al.* and DeLuca *et al.* suggested that the high-affinity AFP-ice interaction is caused by (i) hydrogen bonding, (ii) the van der Waals interaction of the non-polar side chains of AFP, and (iii) the entropic effect resulting from release of bound water from both ice and AFP (*14, 17*).

Type IV AFP, isolated from longhorn sculpin (*18*), is a glutamic acid-rich protein and has high homology with several apolipoproteins whose structural motif is the $\alpha$-helix bundle. The formation of the helical structure of Type IV AFP has recently been evidenced by CD experiments (*19*).

For the AFPs whose tertiary structures have been determined (Type I, II, and III), two different ice-binding manners have been proposed. For Type I and II AFPs, the active surface is postulated to interact with the bipyramidal plane ({2 0 (−2) 1}) (Type I) and ({1 1 (−2) 1}) (Type II) (*5, 10*). For Type III AFP, the active surface is thought to interact with the prism plane ({1 0 (−1) 0}), which involves binding of $Asn^{14}$ to the intersection of the prism and basal planes ({0 0 0 1}) (*13, 14*). These differences imply that the antifreeze mechanism on the ice crystal growth is not simple but has variations.

A new species of antifreeze protein, named RD3, isolated from blood plasma of antarctic eel pout (*Lycodichthys dearborni*) (*20*), was found to be a single polypeptide comprising homologous N-domain ($Asn^{1}$-$Glu^{64}$) and C-domain ($Ser^{74}$-$Glu^{134}$) domains connected through a 9-residue linker ($D^{65}GTTSPGLK^{73}$). Each domain has over 80% sequence identity with the ordinary Type III AFP monomer, and 80% identity with the other domain. All of the key residues for the antifreeze activity identified in the Type III AFP monomer ($Gln^{9}$, $Asn^{14}$, $Thr^{15}$, $Ala^{16}$, $Thr^{18}$, and $Gln^{44}$) are preserved in the two domains. RD3 is the only known species of AFP which forms such an intramolecular dimer, and it possesses 1.9 times higher activity than the other Type III AFPs on a molar basis (*20*). These properties of RD3 raise the question of its overall structure-function relationship; and the answer will provide a clue to understanding the detailed ice-binding mechanism of AFP. However, the structure of such an intramolecular dimer connecting through a (flexible) linker is very difficult to analyze, because such proteins are generally difficult to crystallize and give severe overlapping of the NMR signals. Even if the structure of such protein can be determined, the linker portion is sometimes obscured or impossible to observe [*e.g.*, nitrate-nitrate response regulator protein, NarI (PDB # 1RNL, Brookhaven National Laboratories, Long Island, NY, USA)]. One approach to this problem has been to examine only the unit domain plus linker portion, and this in combination with a study of the intact molecule provides accurate descriptions of the linker structure and the resultant overall structure-function relationship. Here we present the solution structure and preliminary $^{15}N$-NMR backbone dynamics data of the recombinant N-terminal half domain plus the 9-residue short linker portion of RD3 (denoted RD3-NI) by employing the techniques of heteronuclear multidimensional NMR spectroscopy. The data show that the N-domain forms a globular structure which has high similarlity with the Type III AFP monomer. In addition, it appears that the linker forms a definitive structure independently of the N-domain, which comprises a bending motif having a definite orientation. This paper discusses the role of the bending linker with regard to its contribution to the overall structure of the intact RD3 molecule and its antifreeze activity.

## EXPERIMENTAL PROCEDURES

*Purification of Recombinant RD3-NI*—The optimal codons were chosen to construct DNA fragment encoding RD3-NI (*i.e.*, $Asn^{1}$-$Lys^{73}$ of RD3) by reference to the codon usage of the *tuf* AB gene (*21*). The codons of the translation enhancer (*22*), Shine-Dalgalno sequence (*23*), and methionine were attached upstream of the initiation codon. The *Pst*I and *Eco*RI recognition sites at the 5′ and 3′ termini of the DNA fragment were used for the codon insertion into the expression vector pKK223-3UC constructed with pKK223-3 (Pharmacia Biotech, Uppsala, Sweden) and pUC119 (Takara Shuzo, Kyoto). The designed DNA sequence is shown in Fig. 1.

*Escherichia coli* JM105 transformed with the constructed RD3-NI expression plasmid was grown in YT medium (*24*) containing 0.1 M of ampicillin at 30°C to reach $A_{600}$ value of 2.5. This transformant was incubated for 12 h after addition of 0.02% (w/v) IPTG to induce the expression, and harvested by centrifugation. It was then lysed by sonication for 2 min on ice with TES buffer containing 25 mM Tris-HCl (pH 7.5), 50 mM sucrose, 10 mM EDTA, and 0.1 mM PMSF. The lysate was centrifuged ($21,000 \times g$) for 30 min at 4°C. The supernatant was subsequently centrifuged ($130,000 \times g$) for 60 min at 4°C. The obtained material was applied to DEAE Sephacel column (35 × 160 mm) equilibrated with TES buffer. The void fraction was concentrated by addition of ammonium sulfate. The precipitate obtained was dissolved in TES buffer, then dialyzed against TES buffer containing 0.2 M NaCl. The resulting solution was concentrated by ultrafiltration, then applied to Superdex 30 pg column (16 × 700 mm) equilibrated with TES buffer containing 0.2 M NaCl. The final fraction containing RD3-NI was purified by reversed-phase liquid chromatography using TSKgel ODS-80Ts attached to PD-8020 HPLC equipment (Tosoh, Tokyo). The $^{15}N$-labeled RD3-NI was obtained from the transformant cultured in M10 minimal medium (*24*) containing $^{15}N$-labeled $NH_4Cl$ and by using the expression vector pKK223-3UC. Approximately 100 mg of non-labeled RD3-NI and 15 mg of $^{15}N$-labeled RD3-NI were prepared from 2.4 and 6.0 liters of the transformant cultures, respectively. The purity of the sample was checked by tricine-SDS-PAGE (*25*) with Coomassie

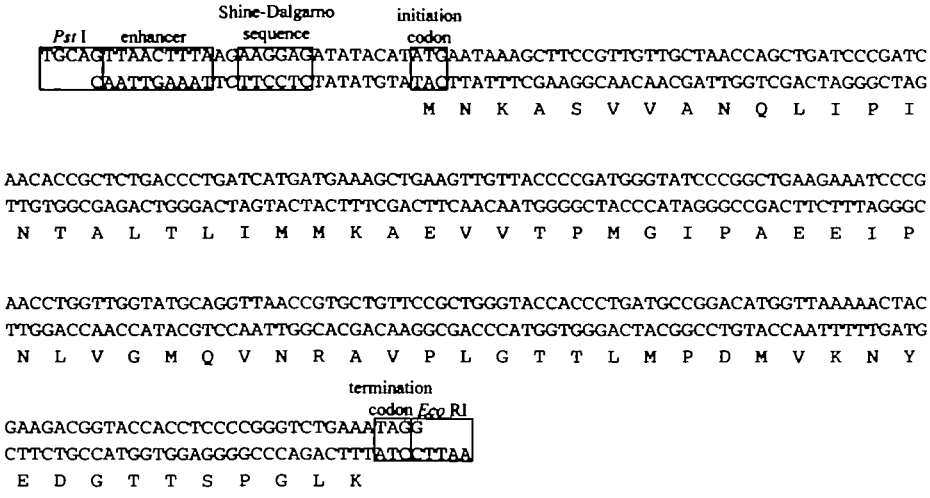


Fig. 1. **DNA and amino acid sequences of recombinant RD3-Nl.** Boxed regions indicate *Pst*I and *Eco*RI recognition sites, an enhancer sequence, a Shine-Dalgarno sequence, initiation and termination codons.

Brilliant Blue staining. The amino acid sequence of the purified RD3-Nl was verified by using an ABI 491 Protein Sequencer (Applied Biosystems, Foster City, CA, USA). The antifreeze activity of the recombinant RD3-Nl was examined by observation of bipyramidal ice crystal formation and thermal hysteresis.

Thermal hysteresis was measured by observation of ice crystal morphology through a Leica DMLB100 photomicroscope equipped with a Linkam LK600 temperature controller. The AFP solution was frozen momentarily with liquid nitrogen, which creates several ice crystal seeds in the solution, then warmed to 0-4°C so as not to break the seeds. This solution was cooled again on the sample stage of the photomicroscope, and the growth of the ice crystal seeds was monitored. We defined thermal hysteresis as the difference between the initial growth temperature and melting temperature of the seed ice crystal.

*NMR Experiments*—For 2D- and 3D-NMR experiments performed in $H_2O$ solution, 4 mg of lyophilized unlabeled or $^{15}N$-labeled RD3-Nl was dissolved in 0.5 ml of $H_2O$ ($H_2O$:$D_2O$=9:1) containing 25 mM of KCl and 1 mM of $NaN_3$ to give a final concentration of 1 mM (pH 6.7). For the experiments in $D_2O$ solution, 4 mg of the non-labeled sample was dissolved in 100% $D_2O$ containing 25 mM of KCl and 1 mM of $NaN_3$ to give a final concentration of 1 mM (pH 6.7).

The NMR experiments were performed on JNM-Alpha and Unity-500 spectrometers at two different temperatures (4 and 30°C). The following sets of 2D- and 3D-NMR data were acquired for the spectral assignments of the $^1H$- and $^{15}N$-resonances: (i) DQF-COSY (*26*); (ii) TOCSY (50-100 ms) (*27*); (iii) NOESY (60-150 ms) (*28*); (iv) {$^{15}N$-$^1H$} HSQC (*29*); (v) $^{15}N$-edited NOESY (60-120 ms); (vi) $^{15}N$-edited TOCSY (60-100 ms) (*30*). The 3D-HNHA (*31*) experiments were performed to obtain the $^3J$-coupling constants between NH and $C_\alpha H$ protons. The {$^{15}N$-$^1H$} $T_1$, {$^{15}N$-$^1H$} $T_2$, and {$^{15}N$-$^1H$} NOE experiments were performed for preliminary estimation of the $^{15}N$-NMR relaxation parameters ($\tau_m$, $\tau_\epsilon$, $S^2$, and $R_{ex}$) (*32*). For the spectra recorded in $H_2O$, water suppression was achieved by presaturation (1.0-2.0 s) using SCUBA (*33*) pulse sequence.

In most of the 2D and 3D experiments, the $^1H$-acquisition dimension was centered at the water frequency [5.01 ppm (4°C) and 4.67 ppm (30°C)] with a spectral width of 6,000 Hz. For the 2D-NMR DQF-COSY, TOCSY, and NOESY, 512 complex $t_2$ data points were collected with 256 $t_1$ increments. The data were zero-filled to a final matrix size of 2,048×1,024 and were Fourier-transformed with a 90°-shifted sine-bell window function. For the {$^{15}N$-$^1H$} HSQC experiments, a total of 512 complex $t_2$ data points ($^1H$) were collected with 128-256 $t_1$ ($^{15}N$) increments. The $^{15}N$ time domain data were extended by zero-filling or linear prediction to give 512 complex points prior to Fourier transformation. The $^{15}N$-edited NOESY and $^{15}N$-edited TOCSY were recorded with 512 ($^1H$)×128 ($^1H$)×32 ($^{15}N$) complex data points and with a spectral width of 1,600 Hz for the $^{15}N$ dimension. The 3D-HNHA spectra were recorded with spectral width of 8,000 Hz for the $^1H$ dimensions and 1,500 Hz for the $^{15}N$ dimension. A total of 512 complex points were collected with 160 $t_1$ ($^1H$) and 96 $t_2$ ($^{15}N$) increments. The {$^{15}N$-$^1H$} $T_1$, {$^{15}N$-$^1H$} $T_2$ and {$^{15}N$-$^1H$} NOE experiments were performed with a spectral width of 8,000 Hz for the $^1H$ dimension and 1,650 Hz for the $^{15}N$ dimension. A total of 1,024 complex $t_2$ data points ($^1H$) were collected with 192 or 256 $t_1$ ($^{15}N$) increments. The $^1H$ chemical shifts were referenced from 2,2-dimethyl-2-silapentane-5-sulfonate sodium salt (DSS). The $^{15}N$ chemical shifts were referenced from an external $^{15}NH_3$ at 0.00 ppm, which was indirectly determined from resonance frequency of DSS using the frequency ratio $\Xi(^{15}N/^1H)=0.101329118$ (*34*). All of the NMR data were processed and analyzed on a SGI Power Indigo2 workstation (Silicon Graphics, Mountain View, CA, USA) using NMRPipe(*35*) and PIPP (*36*) software.

*Structural Calculations*—For the use of interproton distance restraints, NOEs were divided into three groups having the ranges of 1.6-2.7, 1.6-3.5, and 1.6-5.0 Å (lower to upper limits). For the pseudo-atom corrections, the upper limit boundary was increased by 1.0 Å. The $^3J_{NH\text{-}H\alpha}$ coupling constants were estimated from the 3D-HNHA experiment and used for definition of the *phi* angle restraints based on the Karplus equation: $^3J_{NH\text{-}H\alpha}=6.51\cos^2(phi-60^\circ)-1.76\cos(phi-60^\circ)+1.60$ (*31*). The hydrogen-bonding restraints obtained using the hbplus program (*37*) were set at 1.8±0.5 Å for the HN-O distance and 2.8±0.5 Å for N-O distances. In the initial calculation stage, 100 starting structures from an extended conformation of RD3-

Nl were calculated by using the simulated annealing (SA) protocol in X-PLOR 3.851 (*38*) with heating for 60 ps and cooling for 30 ps. Approximately 70% of the initial structures were converged. The structure refinement was carried out by starting with the 50 lowest-energy converged structures using the X-PLOR SA protocol by applying the NOE restraints plus the hydrogen-bonding and the *phi*-angle restraints with heating for 30 ps and cooling for 20 ps. The energy-minimized average structure of RD3-Nl was calculated using the SA protocol with the following steps: (i) 2,000 cycles of restrained energy minimization and (ii) 800 cycles of unrestrained energy minimization. All structural calculations were performed on an SGI Power Indigo2 workstation. The Molscript (*39*), Raster 3D (*40*), and MOLMOL (*41*) programs were used for the structural descriptions.

## RESULTS AND DISCUSSIONS

The recombinant RD3-Nl creates bipyramidal ice crystals like those observed for ordinary Type III AFPs (*12, 13*). The thermal hysteresis measurement using the photomicroscope shows that the level of antifreeze activity of RD3-Nl is identical to that obtained for the QAE isoform (*14*). The antifreeze activity of RD3-Nl is not significantly changed in the pH range between 4 and 10.

The chemical shift assignment of RD3-Nl was carried out using the standard strategy (*42*) by 2D- and 3D-NMR experiments performed at 4 and 30°C. The $^{1}H$- and $^{15}N$-resonances of the residues of RD3-Nl except for seven prolines ($Pro^{12}$, $Pro^{29}$, $Pro^{33}$, $Pro^{38}$, $Pro^{50}$, $Pro^{57}$, and $Pro^{70}$) and aromatic resonances of $Tyr^{63}$ were unambiguously assigned using the sets of 2D-spectra recorded in $H_2O$. Figure 2 shows well-separated {$^{15}N$-$^{1}H$} HSQC spectrum of the recombinant RD3-Nl at 4°C, in which the assignments are indicated for 65 out of 66 expected cross peaks originating from the backbone NH groups and all of 16 cross peaks from the side-chain $NH_2$ groups of asparagines and glutamines. The failure to observe the cross peak of $Asn^{1}$ can presumably be ascribed to rapid exchange with solvent. The assignments of the seven prolines and the aromatic group of $Tyr^{63}$ were performed using the spectra in $D_2O$. Strong $d_{\alpha\alpha}$ $(i-1, i)$ and weak $d_{\alpha\beta}$ $(i-1, i)$ NOEs are typically observed for *cis*-bonded residues between $Xxx^{i-1}$ and $Pro^{i}$, where $i$ is the residue number of proline and Xxx is any residue. Because these characteristic NOEs were observed between $Thr^{28}$ and $Pro^{29}$, we determined the $Thr^{28}$-$Pro^{29}$ peptide bond to be *cis*. The other prolines ($Pro^{12}$, $Pro^{33}$, $Pro^{38}$, $Pro^{50}$, $Pro^{57}$, and $Pro^{70}$) gave neither strong $d_{\alpha\alpha}$ $(i-1, i)$ nor weak $d_{\alpha\beta}$ $(i-1, i)$ NOEs, indicating that the peptide bonds between them and their preceding residues were each trans. The final assignments of RD3-Nl at 4°C were deposited in the BioMagResBank (accession number 4199). The assignments of the globular N-domain portion of RD3-Nl are in good agreement with the corresponding residues of ordinary Type III AFP (*43*).

On the basis of the assignments, the solution structure of RD3-Nl at 4°C was calculated using NMR-derived 958 experimental restraints. We obtained a total of 870 NOEs, of which 460 were intraresidual ($|i-j|=0$), 223 sequential ($|i-j|=1$), 49 medium-range ($1<|i-j|<=4$), and 138 long-range ($|i-j|>4$, where $i$ and $j$ are residue number of two atoms). The 48 dihedral *phi* angle restraints were obtained by measuring $^{3}J_{NH\text{-}H\alpha}$ coupling constants in the 3D-HNHA experiment. In addition, we used 20 hydrogen bond restraints using the hbplus program for the structural calculation (*38*). Figure 3 shows 40 converged structures of RD3-Nl, which have no distance violations greater than 0.2 Å and no dihedral restraint violations greater than 2°. The RMSD of the backbone atoms evaluated for the globular domain (colored in blue: $Ser^{4}$-$Glu^{64}$) of the 40 structures against the averaged structure is 0.83±0.14 Å for the backbone atoms ($C_\alpha$, C′, and N) and 1.53±0.17 Å for all non-hydrogen atoms. The relatively poor convergency of the calculated structures is presumably due to the errors of the intensity calibration of $^{15}N$-edited NOESY. As shown in Fig. 3, the 9-residue linker portion (colored red: $D^{65}$-$K^{73}$) forms a definitive structure having a specific orientation,

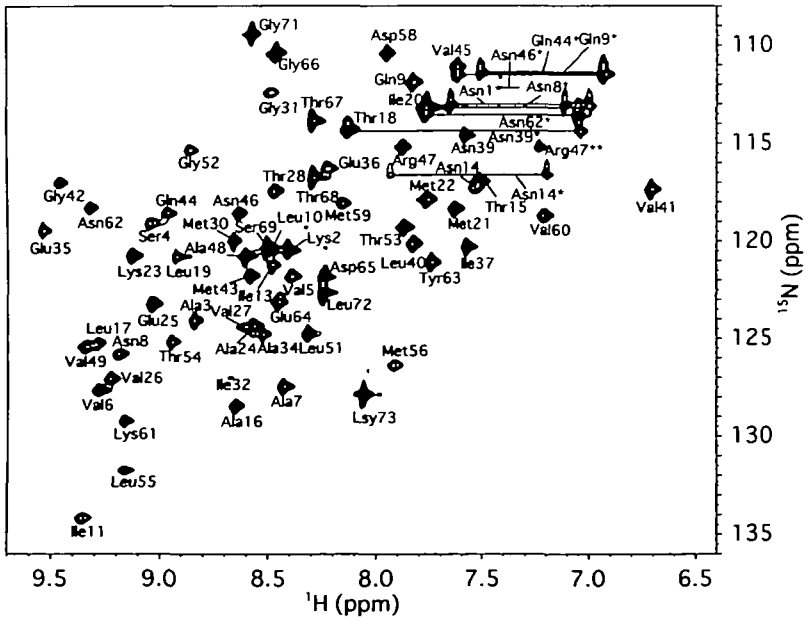


**Fig. 2. {$^{15}N$-$^{1}H$} HSQC spectrum of RD3-Nl obtained at 4°C.** Side-chain resonances of glutamines and asparagines are indicated by single asterisks. Folded-resonance of $Arg^{47}$ ($\varepsilon$-NH) is indicated by a double asterisk.

although it is less well-converged (RMSD = 2.24 ± 0.73 Å for the backbone atoms) for its location at the flexible C-terminal end. The Ramachandran plot obtained with the PROCHECK program (*44*) for residues other than glycine and proline shows that 81 and 19% of the *phi* and *psi* angles fall into the most favored regions and additional allowed regions, respectively. There is no residue distributed in generously allowed regions or disallowed regions. The detailed statistics of the structural calculation are listed in Table I. The atomic coordinates for the final structures of RD3-Nl and the sets of restraints have been deposited with the Brookhaven Protein Data Bank (PDB # 3NLA and 3RDN).

For identification of the secondary structure of RD3-Nl, we employed the program VADAR (*45*), the criteria of which are based on *phi*/*psi* angles, positions of $C_\alpha$ atoms, and the hydrogen-bonding pattern. It appears that the globular domain of RD3-Nl comprises six $\beta$-strands (residues 3-11, 15-18, 22-26, 31-33, 43-49, and 53-55) and three type III turns (18-21, 36-39, and 56-59). The content of the $\beta$-sheet of RD3-Nl is about 47%. It is found that the locations of these secondary structural elements of the globular domain of RD3-Nl are closely similar to those of the QAE isoform NMR structure (PDB # 1KDF) and the X-ray structure (PDB # 1MSI). The RMSD value of the backbone structures between 1KDF, 1MSI, and RD3-Nl was estimated to be 1.17 ± 0.23 Å, indicating that the calculated structure of the globular domain of RD3-Nl has almost equal quality and identical topology with the previously determined structures of Type III AFPs.

Figure 4 shows the ribbon representation of RD3-Nl, in

TABLE I. **Structural statistics of the 40 SA structures of RD3-Nl.**

| | |
|---|---|
| X-PLOR potential energies (kcal·mol$^{-1}$) | |
| $E_{total}$ | 105.76 ± 5.29 |
| $E_{bond}$ | 2.70 ± 0.37 |
| $E_{angle}$ | 85.54 ± 3.17 |
| $E_{improper}$ | 11.47 ± 0.28 |
| $E_{VDW}$ | 4.70 ± 1.45 |
| $E_{NOE}$ | 1.28 ± 1.00 |
| $E_{CDIH}$ | 0.06 ± 0.06 |
| RMSD from idealized geometry | |
| Bonds (Å) | 0.002 ± 0.000 |
| Angles (°) | 0.523 ± 0.010 |
| Impropers (°) | 0.380 ± 0.005 |
| RMSD from distance restraints (Å) | |
| All [910] | 0.005 ± 0.002 |
| Intraresidue [460] | 0.000 ± 0.000 |
| Sequential [223] | 0.004 ± 0.003 |
| Medium-range [49] | 0.001 ± 0.002 |
| Long-range [138] | 0.004 ± 0.002 |
| Hydrogen bond [40] | 0.004 ± 0.003 |
| RMSD from dihedral restraints (°) | |
| 48 *phi* | 0.134 ± 0.057 |
| Total number of restraint violations | |
| NOE > 0.2 Å | 0 |
| Dihedral > 2° | 0 |

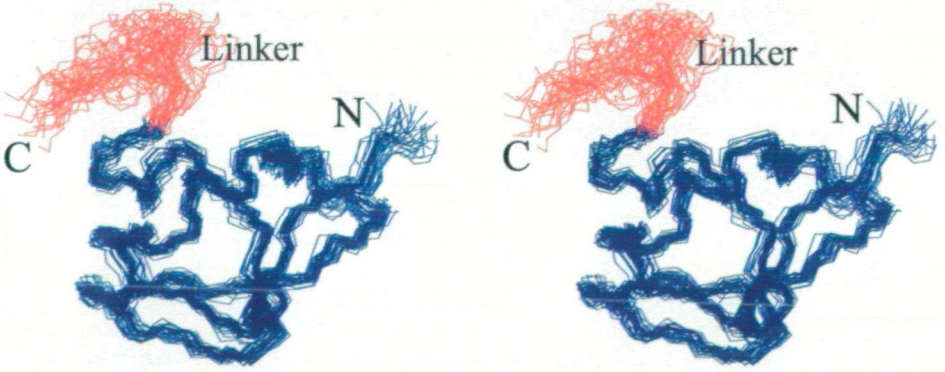


Fig. 3. **Stereoview of the solution structure of RD3-Nl at 4°C.** The backbone atoms (N, $C_\alpha$, and C′) of the family of 40 structures are superimposed and displayed in the "rods" representation. A well-defined globular domain ($Ser^4$-$Glu^{64}$) is colored blue, and the linker portion ($D^{65}$-$K^{73}$) colored red. The linker is less well-defined than the globular domain. The figure was produced with the programs Molscript (*39*) and raster3D (*40*).

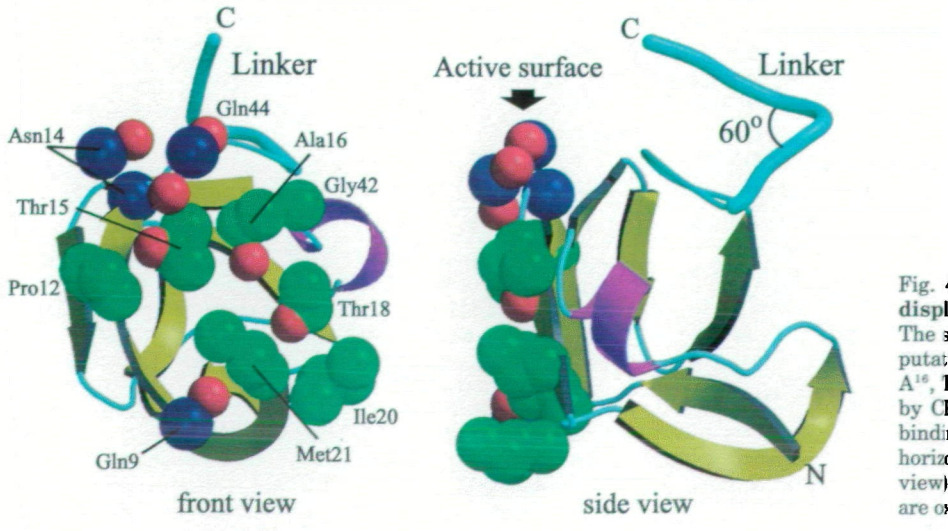


Fig. 4. **NMR solution structure of RD3-Nl displayed in the "ribbon" representation.** The side chains of the residues located on the putative ice-binding surface ($Q^9$, $P^{12}$, $D^{14}$, $T^{15}$, $A^{16}$, $T^{18}$, $I^{20}$, $M^{21}$, $G^{42}$, and $Q^{44}$) are represented by CPK. **Left**: view perpendicular to the ice-binding surface (front view). **Right**: view horizontal to the ice-binding surface (side view). The atoms colored red, blue, and green are oxygen, nitrogen, and carbon, respectively.

which the side chain atoms involved in the ice-binding site (see below) are represented by CPK. It is found that the positions of the polar groups (side chains of $Gln^{9}$, $Asn^{14}$, $Thr^{15}$, $Thr^{18}$, and $Gln^{44}$, and main chain carbonyl group of $Ala^{16}$) in RD3-NI are located so as to fit to the positions of the oxygen atoms of the prism plane of ice crystal. The non-polar side chain groups of $Pro^{12}$, $Thr^{15}$, $Ala^{16}$, $Thr^{18}$, $Ile^{20}$, $Met^{21}$, and $Gly^{42}$ fill the gaps of the above polar residues, which will lead to stabilization of the polar groups with appropriate spacing for the ice-binding. As a consequence, the above side chain atoms form a remarkably flat surface on one side of the globular domain, as seen in the right-hand side of Fig. 4. This flat surface is thought to be a putative ice-binding site, which is similarly assumed for the ordinary Type III AFPs (*14-16*). It is thought that the polar group spacings make enthalpic contributions to the ice-binding through hydrogen bonding, and the planarity makes an entropic contribution through the release of bound water from both ice and AFP. The planarity also leads to a van der Waals interaction between the ice crystal and AFP (*15*). The $Ile^{20}$ and $Gly^{42}$ on the ice-binding surface of RD3-NI are replaced with $Val^{20}$ and $Ser^{42}$ in QAE isoform, respectively. It is found that the $Ile^{20}$ and $Gly^{42}$ do not alter the planarity of the active surface of RD3-NI; such alteration may lead to a significant change in its antifreeze activity. It is noted that a mutant of QAE isoform whose $Ser^{42}$ is replaced with glycine (S42G) has full antifreeze activity (*46*). The side chain hydroxyl group of $Asn^{14}$ of RD3-NI (Fig. 4) is located slightly outside of the active surface, similarly to the case of QAE isoform. This residue has been considered to be necessary for the initial interaction between one edge of the AFP active surface and an intersection of the prism and basal planes of the ice crystal (*15*).

It appears that the 9-residue linker portion ($D^{65}$GTTSPGLK$^{73}$) of RD3-NI forms a definitive structure independently of the globular N-domain (Fig. 3). It assumes a bent form having an angle of approximately 60° (Fig. 4). The independence of the linker from the globular portion is suggested by observation of only one interresidue NOE contact ($Tyr^{63}$-$Pro^{70}$) between the linker and the N-domain. The linker is composed of two short strands, -$D^{65}$GTTSP$^{70}$- and -$G^{71}$LK$^{73}$-, which are bent around -$T^{67}$TSPG$^{71}$-. Our preliminary $^{15}$N-NMR relaxation data[2] shows that the order parameter ($S^2$) (*47*) of the backbone motion becomes smaller along the sequence from $Asp^{65}$ to $Thr^{68}$, while it becomes larger along the -$Gly^{71}$-$Leu^{72}$- sequence. Such a U-shapes $S^2$ profile of the linker supports the existence of the above two short strands that are bent around $Ser^{69}$. This bending property of the linker is considered result from the inherent nature of the amino acid sequence around $Ser^{69}$, which favors a turn conformation, as evidenced by Chou and Fasman's secondary structure prediction (*48*). The independence of the linker in the RD3-NI molecule raises the possibility that the linker keeps its independent nature in the intact RD3 molecule and restricts the orientation of the two domains to some extent. It should be noted that RD3 possesses 1.9 times higher antifreeze activity than the ordinary Type III AFP monomer (*20*), implying that any direct or cooperative interaction exists between the two domains.

[2] Miura, K., Ohgiya, S., Hoshino, T., Nemoto, N., Nitta, K., and Tsuda, S. unpublished data.

One well-known intramolecular dimer similar to RD3 is a dumbbell-shaped calcium regulatory protein, troponin C, which comprises two homologous globular N- and C-terminal half domains (*49*). The two domains are connected by a linker portion which is composed of two $\alpha$-helical parts, D- and E-helices. It has been well-demonstrated that each helix retains exactly the same structural motif and orientation (*e.g.*, interhelical angle) when troponin C is proteolytically divided into the N- and C-terminal half fragments (*50*). Hence, in the present case, a possible 3D-structure of intact RD3 will be modeled when the linker portion of RD3 is assumed to possess virtually the same bending motif as formed in RD3-NI. In our most updated model structure of RD3, the homologous N- and C-terminal globular domains of RD3 come close in tandem through the 60 degree bending linker so as not to associate directly. The property of the globular domain not to favor association is revealed by the present observation of only monomer peaks of RD3-NI (Fig. 2) even under high concentration (8 mg/ml). Further, our preliminary $^{15}$N-NMR relaxation data shows that overall tumbling motional speed of RD3-NI is typical value of a monomer of this size ($\tau_m = 8.4$ ms). In the model structure, the bending linker is able to bring the two flat ice-binding surfaces of the N- and C-domains into the same orientation. With such a shape, the two ice-biding surfaces of the RD3 molecule can interact simultaneously with the same prism plane of the hexagonal ice crystal. The model can account for the 1.9 times higher antifreeze activity of RD3 compared with Type III AFP monomer (*20*). However, it cannot explain the initial interaction of $Asn^{14}$ with the intersection of the prism and basal planes; $Asn^{14}$ is sandwiched between the two globular domains so that the interaction is structurally inhibited. Hence, it is not clear whether the initial ice-binding of RD3 occurs in the same way as with the Type III monomer; the linker property might be significant for this interaction. Of course, a final conclusion should be drawn after clarifying structure of the intact RD3 molecule, which will support, refute, or require modification of the present model of RD3 structure.

To summarize, this paper determines the 3D structure of the N-terminal globular domain plus linker portion of a new species of antifreeze protein, RD3, the globular domain of which appears to form a similar tertiary structure to the ordinary Type III AFP monomer. In addition, the linker appears to assume a definitive bending form having a specific orientation that is independent of the globular N-domain.

The authors are grateful to Ai Miura for upkeeping the NMR spectrometer.

## REFERENCES

1. 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
2. Yeh, Y. and Feeney, R.E. (1996) Antifreeze proteins: Structures and mechanisms of function. *Chem. Rev.* **96**, 601-617
3. DeVries, A.L. and Lin, Y. (1977) Structure of a peptide antifreeze and mechanism of adsorption to ice. *Biochim. Biophys. Acta* **495**, 388-392
4. Yang, D.S.C., Sax, M., Chakrabartty, A., and Hew, C.L. (1988) Crystal structure of an antifreeze polypeptide and its mechanistic implications. *Nature* **333**, 232-237
5. Scieri, F. and Yang, D.S.C. (1995) Ice-binding structure and

mechanism of an antifreeze protein from winter flounder. *Nature* **375**, 426-431
6. Knight, C.A., Cheng, C.C., and DeVries, A.L. (1991) Adsorption of alpha-helical antifreeze peptides on specific ice crystal surface planes. *Biophys. J.* **59**, 409-418
7. Ng, N.F.L. and Hew, C.L. (1992) Structure of an antifreeze polypeptide from the sea raven. Disulfide bonds and similarity to lectin-binding proteins. *J. Biol. Chem.* **267**, 16069-16075
8. Ewart, K.V., Rubinsky, B., and Fletcher, G.L. (1992) Structural and functional similarity between fish antifreeze proteins and calcium-dependent lectins. *Biochem. Biophys. Res. Commun.* **185**, 335-340
9. Ewart, K.V., Yang, D.S.C., Ananthanarayanan, V.S., Fletcher, G.L., and Hew, C.L. (1996) $Ca^{2+}$-dependent antifreeze proteins. Modulation of conformation and activity by divalent metal ions. *J. Biol. Chem.* **271**, 16627-16632
10. Gronwald, W., Loewen, M.C., Lix, B., Daugulis, A.J., Sönnichsen, F.D., Davies, P.L., and Sykes, B.D. (1998) The solution structure of Type II antifreeze protein reveals a new member of the lectin family. *Biochemistry* **37**, 4712-4721
11. Wierzbicki, A., Madura, J.D., Salmon, C., and Sönnichsen, F.D. (1997) Modeling studies of binding of sea raven type II antifreeze protein to ice. *J. Chem. Inf. Comput. Sci.* **37**, 1006-1010
12. Hew, C.L., Wang, N.C., Joshi, S., Fletcher, G.L., Scott, G.K., Hayes, P.H., Buettner, B., and Davies, P.L. (1988) Multiple genes provide the basis for antifreeze protein diversity and dosage in the ocean pout, *Macrozoarces americanus*. *J. Biol. Chem.* **263**, 12049-12055
13. Schrag, J.D., Cheng, C.C., Panico, M., Morris, H.R., and DeVries, A.L. (1987) Primary and secondary structure of antifreeze peptides from arctic and antarctic zoarcid fishes. *Biochim. Biophys. Acta* **915**, 357-370
14. Sönnichsen, F.D., DeLuca, C.I., Davies, P.L., and Sykes, B.D. (1996) Refined solution structure of type III antifreeze protein: hydrophobic groups may be involved in the energetics of the protein-ice interaction. *Structure* **4**, 1325-1337
15. Jia, Z., DeLuca, C.I., Chao, H., and Davies, P.L. (1996) Structural basis for the binding of a globular antifreeze protein to ice. *Nature* **384**, 285-288
16. Yang, D.S., Hon, W.C., Bubanko, S., Xue, Y., Seetharaman, J., Hew, C.L., and Sicheri, F. (1998) Identification of the ice-binding surface on a type III antifreeze protein with a "flatness function" algorithm. *Biophys. J.* **74**, 2142-2151
17. DeLuca, C.I., Davies, P.L., Ye, Q., and Jia, Z. (1998) The effects of steric mutations on the structure of type III antifreeze protein and its interaction with ice. *J. Mol. Biol.* **275**, 515-525
18. Deng, G., Andrews, D.W., and Laursen, R.A. (1997) Amino acid sequence of a new type of antifreeze protein, from the longhorn sculpin *Myoxocephalus octodecimspinosis*. *FEBS Lett.* **402**, 17-20
19. Deng, G. and Laursen, R.A. (1998) Isolation and characterization of an antifreeze protein from the longhorn sculpin, *Myoxocephalus octodecimspinosis*. *Biochim. Biophys. Acta* **1388**, 305-314
20. Wang, X., DeVries, A.L., and Cheng, C.C. (1995) Antifreeze peptide heterogeneity in an antarctic eel pout includes an unusually large major variant comprised of two 7 kDa type III AFPs linked in tandem. *Biochim. Biophys. Acta* **1247**, 163-172
21. Ikemura, T. (1982) Correlation between the abundance of yeast transfer RNAs and the occurrence of the respective codons in protein genes. Differences in synonymous codon choice patterns of yeast and *Escherichia coli* with reference to the abundance of isoaccepting transfer RNAs. *J. Mol. Biol.* **158**, 573-597
22. Olins, P.O. and Rangwala, S.H. (1989) A novel sequence element derived from bacteriophage T7 mRNA acts as an enhancer of translation of the *lacZ* gene in *Escherichia coli*. *J. Biol. Chem.* **264**, 16973-16976
23. Kozak, M. (1983) Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. *Microbiol. Rev.* **47**, 1-45
24. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York
25. Schägger, H. and von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368-379
26. Rance, M., Sørensen, O.W., Bodenhausen, G., Wagner, G., Ernst, R.R., and Wüthrich, K. (1983) Improved spectral resolution in cosy $^1$H NMR spectra of proteins via double quantum filtering. *Biochem. Biophys. Res. Commun.* **117**, 479-485
27. Braunschweiler, L. and Ernst, R.R. (1983) Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy. *J. Magn. Reson.* **53**, 521-528
28. Jeener, J., Meier, B.H., Bachmann, P., and Ernst, R.R. (1979) Investigation of exchange processes by two dimensional NMR specrtoscopy. *J. Chem. Phys.* **71**, 4546-4553
29. Palmer III, A.G., Cavanagh, J., Wright, P.E., and Rance, M. (1991) Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy. *J. Magn. Reson.* **93**, 151-170
30. Kay, L.E., Marion, D., and Bax, A. (1989) Practical aspects of 3D heteronuclear NMR of proteins. *J. Magn. Reson.* **84**, 72-84
31. Vuister, G.W. and Bax, A. (1993) Quantitative J correlation: A new approach for measuring homonuclear three-bond $J(H^NH^\alpha)$ coupling constants in $^{15}$N-enriched proteins. *J. Am. Chem. Soc.* **115**, 7772-7777
32. Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Forman-Kay, J.D., and Kay, L.E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by $^{15}$N NMR relaxation. *Biochemistry* **33**, 5984-6003
33. Brown, S.C., Weber, P.L., and Mueller, L. (1988) Toward complete $^1$H NMR spectra in proteins. *J. Magn. Reson.* **77**, 166-169
34. Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L., and Sykes, B.D. (1995) $^1$H, $^{13}$C and $^{15}$N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* **6**, 135-140
35. Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277-293
36. Garrett, D.S., Powers, R., Gronenborn, A.M., and Clore, G.M. (1991) A common sense approach to peak-picking in two-, three-, and four-dimensional spectra using automatic computer analysis of contour diagrams. *J. Magn. Reson.* **95**, 214-220
37. McDonald, I.K. and Thornton, J.M. (1994) Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777-793
38. Brünger, A.T. (1992) *X-PLOR Manual*, version 3.1, Yale Univ. Press, Cambridge, MA.
39. Kraulis, P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946-950
40. Merrit, E.A. and Murphy, M.E.P. (1994) Raster3D Version 2.0, a program for photorealistic molecular graphics. *Acta Crystallogr.* **D50**, 869-873
41. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics* **14**, 51-55
42. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York
43. Chao, H., Davies, P.L., Sykes, B.D., and Sönnichsen, F.D. (1993) Use of proline mutants to help solve the NMR solution structure of type III antifreeze protein. *Protein Sci.* **2**, 1411-1428
44. Morris, A.L., MacArthur, M.W., Hutchinson, E.G., and Thornton, J.M. (1992) Stereochemical quality of protein structure coordinates. *Proteins* **12**, 345-364
45. Wishart, D.S., Willard, L., Richards, F.M., and Sykes, B.D. (1994) *VADAR: A Comprehensive Program for Protein Structural Evaluation*, Version 1.2. University of Alberta, Edmonton, Alberta, Canada
46. Chao, H.L., Sönnichsen, 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 Sci.* **3**, 1760-1769

47. Lipari, G. and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.* **104**, 4546-4559
48. Chou, P.Y. and Fasman, G.D. (1978) Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol.* **47**, 45-148
49. Herzberg, O. and James, M.N.G. (1985) Structure of the calcium regulatory muscle protein troponin-C at 2.8 Å resolution. *Nature* **313**, 653-659
50. Zot, A.S. and Potter, J.D. (1987) Structural aspects of troponin-tropomyosin regulation of skeletal muscle contraction. *Annu. Rev. Biophys. Chem.* **16**, 535-559