*Protein Science* (1998), 7:1555–1563. Cambridge University Press. Printed in the USA.


# Conformational and dynamic properties of a 14 residue antifreeze glycopeptide from Antarctic cod

ANDREW N. LANE,[1] LISA M. HAYS,[2] ROBERT E. FEENEY,[3] LOIS M. CROWE,[2] AND JOHN H. CROWE[2]

[1]Division of Molecular Structure, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, United Kingdom
[2]Section of Molecular and Cellular Biology, University of California, One Shields Avenue, Davis, California 95616
[3]Department of Food Science and Technology, University of California, One Shields Avenue, Davis, California 95616

(RECEIVED November 19, 1997; ACCEPTED April 7, 1998)

**Abstract**

The $^{1}H$ and $^{13}C$ NMR spectra of a 14-residue antifreeze glycopeptide from Antarctic cod (*Tetramatomus borchgrevinki*) containing two proline residues have been assigned. $^{13}C$ NMR relaxation experiments indicate motional anisotropy of the peptide, with a tumbling time in water at 5 °C of 3–4 ns. The relaxation data and lack of long-range NOEs are consistent with a linear peptide undergoing significant segmental motion. However, extreme values of some coupling constants and strong sequential NOEs indicate regions of local order, which are most evident at the two ATPA subsequences. Similar spectroscopic properties were observed in the 16-residue analogue containing an Arg–Ala dipeptide added to the C-terminus. Molecular modeling also showed no evidence of long-range order, but the two ATPA subsequences were relatively well determined by the experimental data. These motifs were quite distinct from helical structures or $\beta$ turns commonly found in proteins, but rather resemble sections of an extended polyproline helix. Thus, the NMR data provide a description of the local order, which is of relevance to the mechanism of action of the antifreeze activity of the antifreeze glycopeptides as well as their ability to protect cells during hypothermic storage.

**Keywords:** antifreeze glycopeptides; conformation; local order; NMR relaxation

The blood of fish that live in cold seas, such as the Antarctic cod, contains high concentrations (≈35 g/L) of glycoproteins that prevent ice crystal growth by a noncolligative mechanism (for reviews see Feeney, 1988; Yeh & Feeney, 1996). There are eight known fractions of antifreeze glycoproteins (AFGP) that range in molecular mass from 33.7–2.6 kDa (DeVries et al., 1970; Ahlgren & DeVries, 1984), and each consists of a number of quasi repeating structures. The smallest peptide, which is the most abundant in terms of molarity and mass, is AFGP fraction 8 (AFGP-8). The major component of this fraction is a 14 residue glycopeptide that contains 4 identical disaccharides:

$$\text{AAT*AAT*PAT*AAT*PA} \quad (1)$$

where the asterisk denotes the glycosylated residue. The four threonines are glycosylated at the C$\beta$ position with Gal$\beta$1-3GalNAc$\alpha$1-Thr. The minor components differ in that the proline at position 7 is substituted by alanine (30%) and the alanine at position 11 is replaced by proline (20%). The longer sequence AAT*AAT*PAT*AAT*PARA from *Eleginus gracilis* is homogeneous in sequence and slightly more active in its antifreeze activity than AFGP-8 from *Tetramatomus borchgrevinki* (Burcham et al., 1986b).

The mechanism by which these peptides prevent growth of ice crystals is important for several reasons. It is of intrinsic interest to understand how these proteins exert their antifreeze properties. Furthermore, AFGPs are beginning to have practical importance for cryopreservation (Rubinsky et al., 1991; Arav et al., 1993) and hypothermic storage (Hays et al., 1996) of biomaterials that are otherwise difficult to stabilize. For instance, human blood platelets are stored in blood banks at room temperature for a maximum of five days, after which they are discarded. Recent work has shown that the platelets can be stored at 4 °C in the presence of AFGP for at least three weeks with excellent recovery when they are rewarmed (Tablin et al., 1996). Thus, an understanding of properties of these proteins that lead to stabilization of biomaterials is of immediate interest. The best available evidence suggests that the proteins interact at or in the ice-water interface during freezing (Hew & Yang, 1992; Yeh & Feeney, 1996) or with the phospholipid bilayers that comprise biological membranes during cooling (Hays et al., 1996; Tablin et al., 1996). These interactions must

Reprint requests to: Andrew N. Lane, Division of Molecular Structure, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, United Kingdom; e-mail: alane@nimr.mrc.ac.uk.

*Abbreviations:* AFGP, antifreeze glycopeptide; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; DQF-COSY, double quantum filtered correlation spectroscopy; HMQC, heteronuclear multiple quantum coherence; RMSD, root-mean-square deviation; CD, circular dichroism.

depend on the availability of appropriate peptide functional groups for binding to the ice-water interface or membranes, and therefore, the mechanism must take into account the conformational properties of the peptide.

Several studies have addressed the conformational properties of these peptides in solution. Circular dichroism has conclusively shown that there is no detectable $\alpha$-helix present (Bush et al., 1984), unlike in other antifreeze proteins (Davies & Sykes, 1997). However, there are several alternative models for the preferred conformations of these peptides (Bush et al., 1984; Bush & Feeney, 1986; Rao & Bush, 1987; Dill et al., 1992; Drewes & Rowen, 1993; Yeh & Feeney, 1996), so that models of the ice-peptide interactions are highly ambiguous. In this article, we report an NMR examination of the conformational and dynamical properties of AFGP-8.

## Results

### *NMR assignments*

The spin systems of the amino acids and sugars were assigned using a combination of NOESY, COSY, and TOCSY in both $H_2O$ and $D_2O$, and HMQC in $D_2O$. Figure 1 shows a NOESY spectrum in $H_2O$ of the exchangeable protons and a DQF–COSY spectrum of the aliphatic region. The assignment of spin systems is straightforward for this small molecule (Fig. 1A), despite the strong overlap in parts of the spectrum. The NOESY spectrum (Fig. 1B) shows substantial NOEs between the Ala NH and the Thr NH. These are sequential interactions, and are consistent with a small value of $\phi$ on average, rather than an extended conformation. All spin systems were assigned, and nearly complete site-specific assignments were obtained from sequential NOEs (exceptions were A1 and A10) (Table 1). The Ala residues adjacent to the two Pro residues were readily assignable owing to resolved NH shifts and sequential NOEs. The overlap of the other NH resonances made complete assignment of the Ala residues more difficult. However, sequential NOEs between protons of Ala($i$) and Thr($i$ + 1) made it possible to assign four of the remaining Ala residues unambiguously. Connectivities in NOESY and ROESY spectra confirmed that the GalNAc sugars are connected to the threonine residues.

The minor peaks visible in the spectra arise from the small degree of sequence heterogeneity. The sequential NH–NH and intraresidue NHAc–ThrNH NOEs were also observed in the 16-mer Arg–Ala analogue (not shown), which shows that the conformational properties of the major species are the same for both molecules, and that the presence of minor species in the AFGP-8 preparation does not interfere with the conformational analysis.

The $\alpha$-GalNAc resonances are nondegenerate; there are four sets of resonances for most of the protons. Most of the shifts are different from those found in the disaccharide Gal$\beta$1-3GalNac-ol (van Halbeek et al., 1982). The H1 chemical shifts are different from those in either GalNAc or in the disaccharide, which is presumably due to the glycosidic linkage to Thr. However, the variation in shift of the anomeric proton from residue to residue suggests that there are differences in local conformation as the chemical environment is otherwise the same for each residue. In contrast, there is only one set of resonances for the $\beta$-Gal residues, and the shifts are very similar to those of $\beta$Gal in Gal$\beta$1-3GalNAc-ol (van Halbeek et al., 1982). This suggests that these latter residues extend into solution and interact only weakly with the peptide.

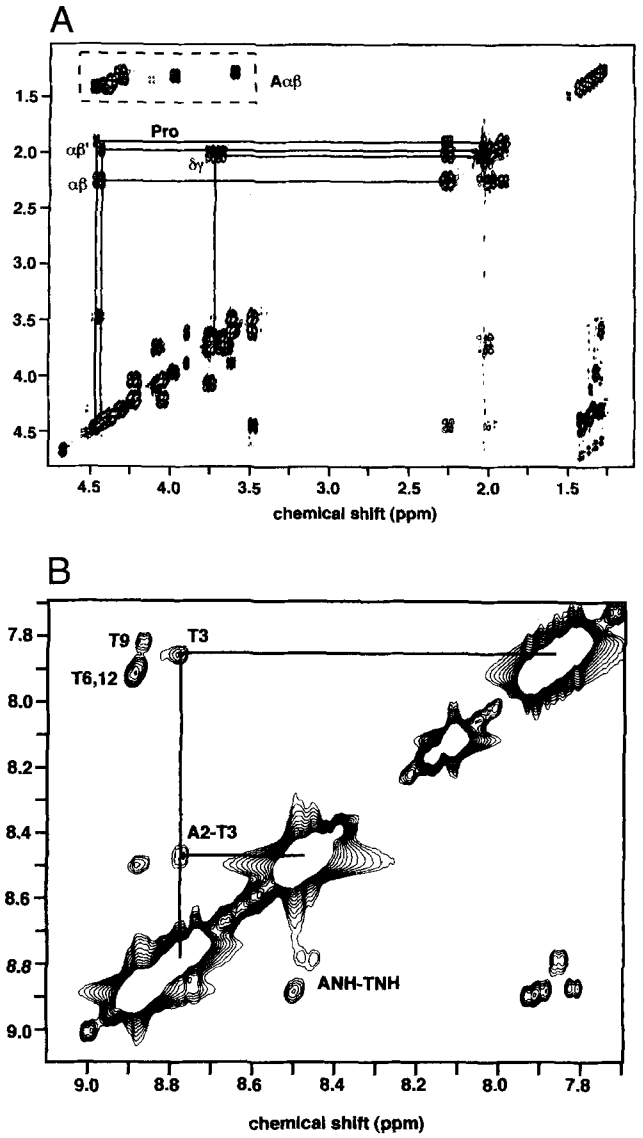


**Fig. 1.** Two-dimensional NMR spectra of AFGP-8 at 5 °C. The spectra were recorded as described in Materials and methods. **A:** DQF-COSY spectrum recorded at 11.75 T in $D_2O$ in showing cross peaks of the sugar and amino acid residues. The acquisition times in $t_2$ and $t_1$ were 0.68 and 0.1 s, respectively. The free induction decays were zero-filled once in $t_2$ and twice in $t_1$ and apodized with a Gaussian function in both dimensions. **B:** NOESY spectrum in $H_2O$, showing NOEs between exchangeable protons. The spectrum was recorded at 14.1 T with acquisition times in $t_2$ and $t_1$ of 0.4 and 0.048 s, and a mixing time of 100 ms.

Figure 2 shows a $^{13}C$-$^{1}H$ HMQC spectrum of the glycopeptide recorded at natural abundance. Given the assigned proton resonances, it is straightforward to assign the carbon resonances (Table 1). Note that, as for the proton resonances, the carbon shifts of the $\beta$-Gal residues are degenerate. The $^{1}H$ and $^{13}C$ shifts of the Ala residues are similar to those of random coil peptides (Wüthrich, 1986; Wishart et al., 1995), whereas the CH$\alpha$ shifts and especially the NH of the four Thr residues show significant deviations from random coil shifts (>0.1 and >0.5 ppm, respectively). The NOESY spectra show that the C$\delta$H of the two Pro residues are close to the $\alpha$CH of the preceding Thr residue, and they also have shifts different from the random coil values (>0.1 ppm).

**Table 1.** *NMR assignments of AFGP-8 at 5 °C*[a]

| Residue | NH | α | β | γ | δ |
|---|---|---|---|---|---|
| Ala 1 | — | 3.69(50.5) | 1.32(19.1[b]) | — | — |
| Ala 2 | 8.44 | 4.35(51.5) | 1.45(19.1[b]) | — | — |
| Thr 3 | 8.78 | 4.45(59.8) | 4.33(79.0) | 1.32(21.2) | — |
| Ala 4 | 8.49 | 4.40(51.5) | 1.43(19.1[b]) | | |
| Ala 5 | 8.50 | 4.45(53.0) | 1.40(19.1[b]) | — | — |
| Thr 6 | 8.90 | 4.65(58.1) | 4.31(76.8) | 1.36(21.2) | — |
| Pro 7 | — | 4.48(62.6) | 2.25(31.8) 1.90 | 2.01(27.1) | 3.75(50.5) |
| Ala 8 | 8.48 | 4.39(51.7) | 1.37(19.1[b]) | — | — |
| Thr 9 | 8.87 | 4.50(59.7) | 4.30(79.0) | 1.29(21.3) | — |
| Ala 11 | 8.50 | 4.45 (52.0) | 1.43(19.1[b]) | — | — |
| Thr 12 | 8.89 | 4.66(58.2) | 4.31(76.5) | 1.36(21.3) | — |
| Pro 13 | — | 4.45(62.6) | 2.21(31.8) 1.98 | 2.01(27.1) | 3.70(50.5) |
| Ala 14 | 8.11 | 3.96(53.5) | 1.32(19.1[b]) | — | — |

| Sugars | 1 | 2 | 3 | 4 | 5 | 6 | NH | Ac |
|---|---|---|---|---|---|---|---|---|
| α-GalNAc3 | 4.90(101.3) | 4.23(71.5) | 4.07(79.9) | 4.1(73.7) | | 3.77(63.8) | 7.80 | 2.00(24.8) |
| α-GalNAc6 | 4.98(101.2) | 4.25(71.5) | 4.10(77.9) | 4.1(73.7) | | 3.77(63.8) | 7.91 | 2.00(24.8) |
| α-GalNAc9 | 4.93(101.7) | 4.24(71.5) | 4.08(77.9) | 4.1(73.7) | | 3.77(63.8) | 7.82 | 2.00(24.8) |
| α-GalNAc12 | 5.02(101.7) | 4.24(71.5) | 4.08(77.9) | 4.1(73.7) | | 3.77(63.8) | 7.88 | 1.98(24.8) |
| β-Gal | 4.47(107.5) | 3.49(73.0) | 3.63(74.9) | 3.90(71.1) | 3.65(77.6) | 3.77(63.8) | — | — |

[a]Assignments were made as described in the text. Values in parentheses are $^{13}C$ chemical shifts.
[b]Alanine methyl groups are essentially degenerate.

The Thr residues are quite distinct, have slowly exchanging NH (both backbone and NAc), and large $^3J_{NH\alpha CH}$. In contrast, many of the Ala NH are nearly degenerate, and the protons exchange relatively rapidly.

The NOESY spectrum shows that the NH of the acetamide side chain of the GalNac residues has a large coupling constant ($^3J >$ 10 Hz) to the H2 of the sugar ring, indicating restricted rotation, and a trans conformation with respect to H2. Similarly, the threonine residues also have a large (>8 Hz) value of $^3J_{NH\text{-}\alpha CH}$, consistent with a value of $\phi$ characteristic of an extended conformation (i.e., ca. −120°). In contrast the Ala residues, where resolution permits, showed smaller values of $^3J_{NH\text{-}\alpha CH}$ (ca. 5–6 Hz), which is consistent either with averaging about $\phi$ or some intermediate unique value (e.g., −90°) (Wüthrich, 1986; Schwalbe et al., 1997). However, the rate of exchange of the Ala NH with water, as observed in the ROESY spectrum, is for most residues considerably greater than for the threonine residues, or the NHAc in the GalNAc residues. This indicates that the Ala residues do not participate in hydrogen bonds, which would be consistent with averaging (and see below).

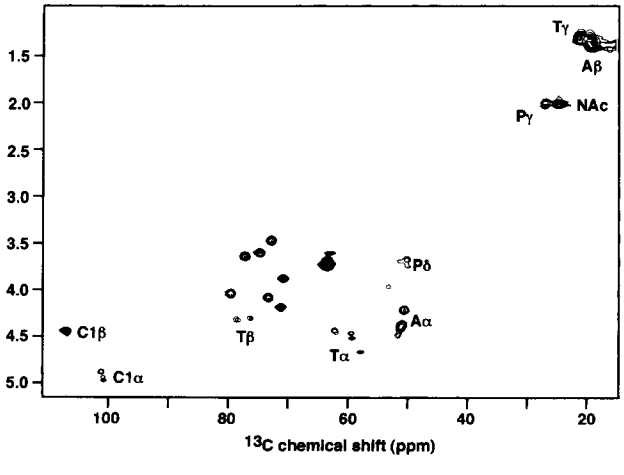


**Fig. 2.** HMQC spectrum of AFGP-8. The spectrum was recorded at 9.4 T in $D_2O$ at natural abundance and 5 °C as described in Materials and methods. The acquisition times were 0.512 s in $t_2$ and 0.098 s in $t_1$. The free induction decays were zero-filled once in $t_2$ and twice in $t_1$ and apodized with a Gaussian function in both dimensions.

*Dynamics*

At 10 °C and lower, the $^1H$ NOEs were all negative, and quite large, implying effective correlation times that are large compared with $1/\omega$, i.e., $\tau \gg 0.3$ ns. This was confirmed by analysis of the $^{13}C$ relaxation data (Table 2). For example, the relaxation rate constants for the carbon sites in the pyranose ring of β-Gal were very similar. Application of the model of Lipari and Szabo (1982) indicates an overall correlation time of around 3.2 ns, and a moderate order parameter of 0.6–0.7. This magnitude of the order parameter is required to account for the significant NOE, which otherwise would be near to the minimum for an overall correlation time of 3 ns. The significantly lower order parameter of the C6 position is consistent with an additional rotation about C5–C6. The smaller order parameters observed for the methyl groups is as

**Table 2.** *$^{13}C$ relaxation data for AFGP-8 at 5 °C*[a]

| Carbon | | $R_1$ ($s^{-1}$) | $R_2$ ($s^{-1}$) | NOE | $\tau_0$ (ns) | $\tau_i$ (ns) | $S^2$ | $E$ (%) |
|---|---|---|---|---|---|---|---|---|
| βGal | 1 | 4.0 | 14 | 1.4 | 3.4 | 0.14 | 0.8 | 0.36 |
| | 2 | 3.4 | 10.3 | 1.5 | 3.2 | 0.09 | 0.60 | 1.1 |
| | 3 | 3.4 | 11.4 | 1.5 | 3.4 | 0.09 | 0.65 | 0.93 |
| | 4 | 3.4 | 10.8 | 1.5 | 3.2 | 0.07 | 0.65 | 1.8 |
| | 5 | 3.9 | 11.2 | 1.5 | 3.1 | 0.11 | 0.70 | 1.4 |
| | 6 | 3.1 | 8.9 | 2.0 | 3.1 | 0.11 | 0.35 | 0.7 |
| αGalNAc | 1 | 3.9 | 17.5 | 1.3 | 3.9 | 0.26 | 0.9 | 0.8 |
| | 4 | 4.0 | 17.7 | 1.5 | 4.5 | 0.28 | 0.8 | 4.5 |
| Ala | α | 3.2 | 8 | 2.0 | 3.6 | 0.14 | 0.35 | 1.6 |
| Ala | β | 1.3 | 2.5 | 2.0 | 2.5 | 0.03 | 0.15 | 2.3 |
| Thr | α | 4.2 | 9 | 1.5 | 2.3 | 0.1 | 0.65 | 0.29 |
| Thr | γ | 1.3 | 2.6 | 2.0 | 2.6 | 0.03 | 0.15 | 1.2 |
| Pro | β | 2.8 | nd | 1.55 | 4.8 | nd | 0.65 | 1.4 |
| Pro | γ | 1.9 | 8 | 1.68 | 4.5 | 0.05 | 0.35 | 2.4 |
| Pro | δ | 1.8 | nd | 1.3 | 4.8 | nd | nd | — |
| NAc | Me | 0.67 | 1.1 | 1.5 | 3.1 | 0.02 | 0.05 | 1.9 |

[a]The relaxation data were determined at 9.4 T as described in the text. The estimated error on $R_1$ is ±5% for sugar carbons and methyl groups and ±10% otherwise; the precision of the NOE measurements was ±5% and for $R_2$ ± 10%.

expected, owing to rapid rotation about the C-methyl bond. Most of the order parameters for the backbone atoms are smaller than is observed in globular proteins (Nicholson et al., 1996), which is consistent with significant segmental motion on the nanosecond time scale. Further, the variation of $\tau_0$ with position indicates that the assumption of an isotropic rotor does not fully account for the relaxation data.

The effective tumbling time is between 2.5 and 4.5 ns, with a median value of about 3 ns. The rotational correlation time expected for a glycopeptide of relative molecular mass 2,600, based on the Stokes–Einstein relation for an unhydrated sphere, would be only about 1.3–1.5 ns in $D_2O$ at 5 °C. Globular proteins and peptides generally have somewhat larger rotational correlation times than expected for an unhydrated sphere (Daragan & Mayo, 1997). The presence of different correlation times is consistent with a nonsymmetric, elongated molecule (see below). In the presence of segmental flexibility, the effective correlation times should be complicated time averages over all of the conformers present in the solution. The minimum rotational correlation time was 2.5 ns, which would be consistent with substantial rotational anisotropy or extensive hydration, such as via the carbohydrate moiety (see below).

### *NH exchange and hydration*

Presaturation of the solvent signal caused some loss of intensity of the exchangeable protons, particularly those of the Ala NH that resonate near 8.5 ppm. The ROESY spectrum (Fig. 3) confirms that these protons exchange the most readily, though it is still possible to observe cross peaks to αCH in the DQF-COSY spectrum recorded with presaturation. The Thr NH also exchange to some extent, whereas the NH of the N-acetyl groups and the NH of Ala14 showed essentially no exchange with solvent on the second timescale. Clearly there is a substantial range of exchange rates in the molecule, with some residues showing relatively ready exchange (most of the Ala residues), which presumably therefore are not strongly hydrogen bonded. In contrast, the NAc groups, and the NH of Ala14, are less accessible to the solvent.

There is also a broad peak at ≈5.7 ppm (Fig. 3) that becomes increasingly exchange broadened as either the temperature or pH is raised, and is not present in $D_2O$. As all the amide protons have been accounted for, the only remaining exchangeable protons are hydroxyls. Hydroxyl protons typically resonate in this region of the spectrum (ca. 6.5–7 ppm in RNA, and 5–6 ppm in carbohydrates at low temperatures) (Leroy et al., 1985; Poppe & van Halbeek, 1994; Conte et al., 1996). In the arginine-containing peptide, the side-chain guanidino NH and $NH_2$ resonances were also visible, and showed negligible exchange with the solvent (not shown). However, the two aminoproton signals showed mutual exchange on the 100 ms timescale, indicating rotation about the C–N bond causing averaging of the four signals into two (Nieto et al., 1997). The relatively slow solvent exchange of these protons and slow rotations in the side chain are consistent with structural interactions of the residue, as has been observed in proteins (Nieto et al., 1997).

Interactions between water and nonexchangeable protons were also observed, including positive NOEs to all of the methyl groups and the βH of the Pro residues, and small negative NOEs to the αCH (Fig. 3). This is consistent with the amino acid residues being exposed to solvent, and argues against a compact folded structure.

### *Sequential NOEs and coupling constants*

The NH of the Thr residues showed $^3J_{NH\alpha CH} > 8$ Hz and $^3J_{\alpha\beta} < 3$ Hz. In contrast, in the cases where they could be measured, $^3J_{NH\alpha CH} \approx 6$ Hz for Ala, which is consistent with averaging about

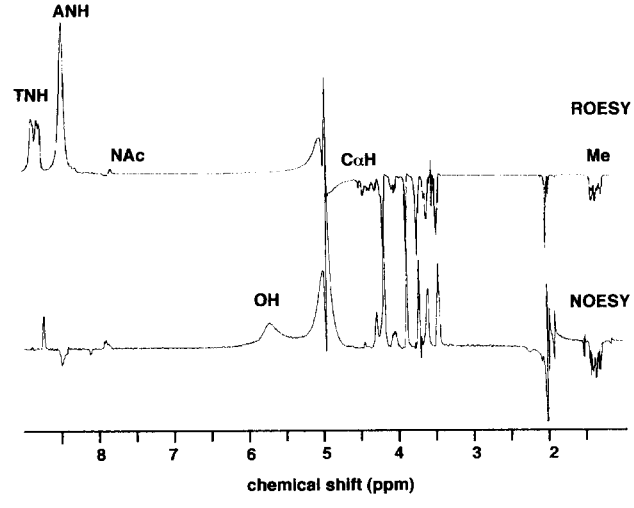


**Fig. 3.** ROESY and NOESY cross sections of AFGP-8. The spectra were recorded at 14.1 T, 5 °C as described in Materials and methods. The mixing times were 100 ms (NOESY) and 50 ms (ROESY). The spin-lock field strength in the ROESY experiment was 3.73 kHz. The acquisition times in $t_2$ and $t_1$ were 0.37 and 0.05 s, respectively, in both experiments. The free induction decays were zero-filled once in $t_2$ and twice in $t_1$ and apodized with a Gaussian function in both dimensions. The cross sections through the water resonance show the relative amount of exchange of different protons (positive peaks) and ROEs to nonexchangeable protons (negative peaks).

$\phi$. In addition, all four NAc-Gal residues showed $^3J_{NH} > 9$ Hz, implying a trans orientation of the NH with respect to the sugar C2H. The strong TNH-NH(Ac), relatively large value for $^3J_{NH\alpha CH}$ and the small $^3J_{\alpha\beta}$, indicates a reasonably well-ordered conformation of the Thr residues. Furthermore, the orientation of the NAc-Gal residue with respect to the Thr is probably relatively well ordered by virtue of the numerous NOEs connecting these fragments (Table 3).

Two of the Ala residues also showed NOEs to the adjacent Pro residues of the kind $d_{\alpha N}$ and $d_{\beta N}$, which suggest an extended conformation at the two PA steps (Wüthrich, 1986). Furthermore, NOEs between the T$\alpha$CH and the neighboring P$\delta$H are consistent with an extended conformation for the two TP steps, if the Pro is in the trans conformation. The $^{13}$C NMR shifts for the Pro residues indeed indicate a trans peptide bond, indicating a structured component near the two TPA subsequences. This is compatible with the slow exchange of the peptide NH. However, the relative sequential NOE intensities for the two TP steps at a mixing time of 100 ms were not equal, indicating that despite the identity of the local sequence, the conformation and/or dynamics may differ somewhat. Finally, $d_{\alpha N}$ connectivities were observed for AT steps. Hence, the limited data available are consistent with local regions of ordered structure within an overall extended, flexible structure.

**Table 3.** *Experimental constraints*[a]

| Proton a | Proton b | Strength |
|---|---|---|
| T3NH | A2$\alpha$CH | m |
| | N-Ac-Gal NH | s |
| | N-Ac-Gal H2 | w |
| | T3$\beta$CH | w |
| T3$\alpha$CH | T3$\beta$CH | s |
| A5NH | T6NH | m |
| A5$\alpha$H | T6NH | m |
| T6NH | N-Ac-Gal NH | s |
| | N-Ac-Gal H2 | w |
| | T6$\beta$CH | w |
| T6$\alpha$H | P7$\delta$H | m |
| P7$\alpha$H | A8NH | m |
| P7$\beta$H | A8NH | w |
| A8$\alpha$H | T9NH | m |
| T9NH | N-Ac-Gal NH | s |
| | N-Ac-Gal H2 | w |
| | T9$\beta$CH | w |
| T9$\alpha$CH | T9$\beta$CH | s |
| A11NH | T12NH | m |
| A11$\alpha$H | T12NH | m |
| T12NH | N-Ac-Gal NH | s |
| | N-Ac-Gal H2 | w |
| | T12$\beta$CH | w |
| T12$\alpha$CH | T12$\beta$CH | s |
| T12$\alpha$CH | P13$\delta$CH | m |
| P13$\alpha$H | A14NH | s |
| P13$\beta$H | A14NH | m |
| $\beta$-GalH1 | NAc-GalH3 | s |
| $\phi$ | Thr | $-70$ to $-160°$ |
| $\chi_1$ | Thr | 1 to 100° |

[a]Strength refers to the relative cross-peak intensity in NOESY spectra as described in Materials and methods; s = strong, m = medium, and w = weak.

*Model structures*

The relatively large negative $^1$H-$^1$H NOEs indicate that the molecular tumbling rate is well within the slow motion limit. The carbon relaxation data are consistent with a molecule rotating with a correlation time of ca. 3–4 ns at 5 °C, possibly anisotropically, with substantial internal motions; the molecule is relatively flexible. The presence of the bulky disaccharide every third residue makes it impossible for the peptide to fold randomly. Essentially, all that is possible are extended structures, helices, or a compact structure with the disaccharides on the outside. The presence of the two Pro residues rules out an extensive $\alpha$-helical arrangement, which is supported by the reported absence of a helical CD signal (Bush et al., 1984). A compact folded structure is also unlikely given the absence of any long-range NOEs or ROEs under any experimental conditions tested, despite relatively narrow resonances (i.e., such that chemical exchange and line broadening cannot account for small NOEs).

The relative intraresidue and inter-residue NOE intensities for the Thr residues, the magnitude of the coupling constant, and NOEs to the Gal NAc residues make it extremely probable that these residues are relatively well ordered. Given the observed sequential NOEs, ordering is also probable about the two TPA subsequences (see above). For the 14 residue peptide, and given that the peptide bonds of both Pro residues are trans, then there are only 28 degrees of freedom in the peptide (as $\chi_1$ for Ala is not defined), plus 2 degrees for freedom for each disaccharide. However, although the total number of restraints (54) exceed the number of degrees of freedom in the system, because of incomplete sequential assignments of the molecule, the uneven distribution of constraints along the molecule, the lack of any long-range constraints, and the presence of conformational dynamics make it impossible to produce a unique family of structures. Nevertheless, the available information makes it reasonable to find classes of models that are compatible with the data, and indicate the preferred conformational preferences in solution.

An initial model structure of the glycosylated peptide as an extended chain was built using InsightII as described in Materials and methods. The peptide was then folded into different initial secondary structures and then refined by restrained energy minimization and restrained molecular dynamics calculations using the constraints given in Table 3, and the information available from three-bond coupling constants, as described in Materials and methods.

Models starting from the left-handed $\alpha$ helix or the $3_{10}$ helix had a very high restraint energy that was not greatly relieved by energy minimization or high-temperature MD sampling. These starting models always became trapped in a local minimum that still had very large restraint violations. We conclude that such structures are incompatible with the experimental data, and are not likely to be present in significant amounts in solution. Both extended chain and the right-handed $\alpha$ helix, while starting with a high potential energy and large restraint violations, refined to structures having low potential energies and insignificant restraint violations. As there are no long- or medium-range NOE constraints, the conformational preferences in the models rely entirely on sequential and intraresidue information. It is not surprising therefore that the regions that are conformationally restricted are the two ATPA steps, where sequential NOEs were observed (Table 3), whereas in general the AA steps were not determined by the data, and are the source of long-range disorder in the structures. Thus, the pairwise

RMSD values for the backbone atoms of the whole structures are 4.9 ± 1.5 Å, and the RMSD between the refined structures and one in which no restraints were applied was 5.17 ± 1.1 Å. However, the RMSD values for the ATPA and ATAA subsequences were much smaller. We will denote these subsequences as S1 (A2T3A4A5), S2(A5T6P7A8), S3(A8T9A10A11), and S4(A11T12P13A14). Owing to a lack of data, the first segment, S1, is not defined. Figure 4 shows superpositions of the three substructures S2, S3, and S4 from different runs of the restrained dynamics, and the statistics of the converged structures are given in Table 4. The RMSD value for the backbone atoms was 0.83 ± 0.28 Å for S4 and 1.38 ± 0.43 Å for the heavy atoms. These values are substantially smaller than the RMSD to the structure minimized without constraints (1.56 ± 0.11 and 2.43 ± 0.21, respectively). This shows that the restraints rather than the force field are responsible for the clustering in conformational space.

As shown in Figure 4 and Table 4, the other two subsequences are less well defined, though they both have backbone features in common. In terms of the backbone RMSD, the structures are determined in the order S4 > S2 > S3, for equal number of local constraints. It may be that S3 is the least well determined because there is an additional degree of freedom in the backbone compared with the two Pro-containing subsequences, where $\phi$ is essentially fixed. There may be increasing order toward the C-terminus, which is supported by the observation that in the 16-residue analogue, which is the same sequence but with a C-terminal Arg–Ala extension, shows evidence of structure in the Arg side chain (see above). The higher RMSD value for all heavy atoms is in part due to the poorer determination of the orientation of the disaccharides with respect to the peptide backbone, as seen in Figure 4.

Although the superposition of the overall structures is poor owing to the lack of long- and medium-range interactions, individual model structures have a global feature in common, namely that they appear as flattened sausages. This is in keeping with the observed anisotropy of the molecule, and accounts in part for the internal flexibility. Thus, the molecules are extended rather than globular.

**Table 4.** *Statistics of local structures of AFGP-8*[a]

| $U_{pot}$ (kJ mol$^{-1}$) | $U_f$ (kJ mol$^{-1}$) | $U_b$ (kJ mol$^{-1}$) |
|---|---|---|
| −21.9 ±7.1 | 0.19 ±0.07 | 3.37 ±0.19 |
| | RMSD/Å | |
| Segment | Backbone | All heavy |
| AFGP-8 | 4.95 ± 1.5 | 5.7 ± 1.1 |
| S2 | 1.04 ± 0.41 | 2.30 ± 0.63 |
| S3 | 1.47 ± 0.39 | 1.86 ± 0.44 |
| S4 | 0.83 ± 0.28 | 1.38 ± 0.43 |

[a] $U_{pot}$ is the total potential energy, $U_f$ is the residual forcing energy from the experimental restraints, $U_b$ is the residual bond violation energy, and the RMSD was calculated pairwise.

## Discussion

Numerous local structures have been proposed for this class of antifreeze glycopeptides, including the $\gamma$-turn (Drewes & Rowen, 1993) and the polyproline helix (Bush et al., 1984; Bush & Feeney, 1986). We have already shown that the $3_{10}$ and $\alpha$ helices are not compatible with the experimental data (see above). The $\phi,\psi$ angles of the S3 and S4 segments, which are the most well-determined parts of the molecule, are incompatible with either $\beta$- or $\gamma$-turns. The closest secondary structure element that has been described is the polyproline helix.

The NMR data show that there are parts of the molecule that adopt preferred conformations in solution, and that these are in a more extended conformation than in a compact secondary structure. The data also show that the molecule as a whole is on average extended, as there was no evidence of any long- or medium-range interactions in the NOESY spectra. Furthermore, the molecule exhibits some segmental motion. It is notable that the amide exchange rates, particularly those of the Ala residues, indicate that they are not protected by any structural features such as H-bonding, and are likely to be rather flexible on average. It is not possible therefore to provide a picture of a unique conformation for this peptide in solution, nor is it meaningful to do so. A complete description of the glycopeptide would have to take into account in detail the conformational averaging, and would involve an ensemble approach such as recently described for denatured lysozyme (Schwalbe et al., 1997). However, because the ATPA subsequences do have a preferred conformation, it is useful to focus on them, as it is possible that the disaccharide moiety interacts with the ice crystal or membrane surface. It may be reasonable that this pre-

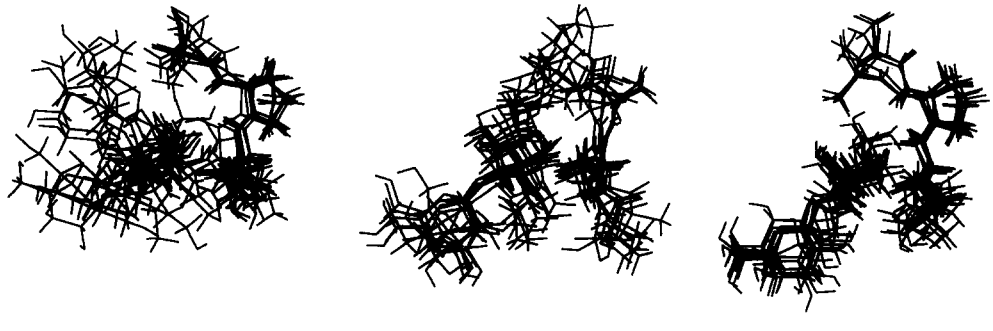

**Fig. 4.** Preferred conformations of the AFGP-8 peptide. Segments 2 (A5–A8), 3 (A8–A11), and 4 (A11–A14) are shown left to right as superpositions of eight structures.

ferred conformation also exists for the AATA subsequences, though the population on average may be less than for ATPA by virtue of the extra degree of freedom in the peptide backbone. This seems to be the case for segment S3, which clearly shows features in common with the best-determined segment S4. Thus, although a Pro residue may reduce the local conformational flexibility, it appears not to be essential for the formation of this kind of local secondary structure. However the presence of the Pro seems also no guarantee of greater rigidity, as the NOE intensities involving the two Pro residues are not equal. Whether this reflects differences in mean conformation or different dynamics is not clear at present. According to their NMR spectra, the longer antifreeze glycopeptides, which do not contain Pro, must also be flexible (Bush & Feeney, 1986). It is possible that the local segmental flexibility observed in these small glycopeptides is responsible for the flexibility seen in the longer molecules.

It has been suggested (Mimura et al., 1992) that the NH of the NAc group may make a stabilizing hydrogen bond to the Thr carbonyl oxygen, which would account for the relatively slow exchange of these amide protons. However, such an orientation would place the Thr NH far from the NAc NH, which is not in agreement with the observed strong NOE between these protons. Indeed, this NOE is of similar strength in all four Thr residues, indicating that the presence of the Pro does not directly influence the preferred conformation of the glycosylated Thr residue.

It is likely that the Ala residues are the sites of segmental flexibility in the molecule, so that the conformations of ATA sequences are probably less well defined in solution. It would also be easy to rotate the ATXA subsequences to place them all on the same face of the molecule. It is certainly possible that there is a significant population in solution of such a structure over the entire length of the molecule, which would place all the sugars on one face of the molecule, as described by Bush et al. (1984), Bush and Feeney (1986), and Rao and Bush (1987). In one model of the ice-binding mechanism, hydroxyl groups from the carbohydrates interact with the ice surface (Knight et al., 1993). As far as binding to an ice facet, however, it would not be necessary for such a structure to be present. If only two disaccharides were on the same side of the molecule on forming an initial complex, it would be possible for the remainder of the molecule to rotate so as to bring the other disaccharides into play. The addition of the Arg–Ala dipeptide to the end of the 14-mer increases the antifreeze activity (Burcham et al., 1986b). The arginine residue appears to be at least partly structured, and the extra H-bonding capacity of the guanidino side chain could serve to provide an additional anchor for the C-terminal part of the peptide chain to interact with ice. The detailed mechanism of action of the AFGPs may be different from other antifreeze proteins that do not contain sugars. These proteins are folded, relatively rigid, fall into four classes having different secondary and tertiary structures, and bind to different ice facets (Davies & Sykes, 1997). The primary structural feature they have in common is a flat ice-binding surface. The mechanism of action of the antifreeze proteins remains unsolved. It is not clear what role flexibility has in the action of the AFGPs, though it is possible that this circumvents the need to bind to a flat surface.

## Materials and methods

AFGP fraction 8 was isolated from the Antarctic cod, *T. borchgrevinki*, as previously described (DeVries et al., 1970), then lyophilized. AFGPs were further purified by acetone precipitation according to a recent protocol (Hays et al., 1996). The major sequence of AFGP-8 is

$$\begin{array}{l} 1 \quad 5 \quad 10 \\ \text{AATAATPATAATPA} \end{array} \qquad (2)$$

AFGP-8 is in fact a mixture of peptides; approximately 30% proline at position 7 is substituted by Ala, and 20% of the Alas at position 11 are replaced by proline. This heterogeneity gives rise to additional minor peaks in the NMR spectra. We have concentrated only on the major, proline-containing species. Additional experiments were carried out on the 16-mer isolated from *E. gracilis* (Burcham et al., 1986a):

$$\begin{array}{l} 1 \quad 5 \quad 10 \quad 15 \\ \text{AATAATPATAATPARA} \end{array} \qquad (3)$$

This peptide is homogeneous in sequence.

### *NMR spectroscopy*

AFGP-8 (20 mg) was dissolved in $H_2O$ or $D_2O$ (0.62 mL) to give a concentration of 12 mM. NMR spectra were recorded at different temperatures and magnetic field strengths as follows.

NMR spectra were recorded at 14.1 T on a Varian Unity spectrometer and at 11.75 T on a Varian Unityplus spectrometer. For spectra in $H_2O$, the solvent signal was suppressed using the pulsed gradient echo Watergate method (Piotto et al., 1992). Two-dimensional spectra were recorded using the hypercomplex method for obtaining phase-sensitive data (States et al., 1982). NOESY spectra were recorded at 0, 5, and 10 °C with mixing times of 50, 100, 200, and 500 ms. TOCSY spectra were recorded using MLEV-17 (Bax & Davis, 1985), with a typical spin-lock strength of 8 kHz for a duration of 50 ms. A ROESY spectrum was recorded using an unmodulated mixing pulse of 50 ms duration and 4 kHz $B_1$ field strength. DQF-COSY spectra were recorded using standard methods.

Heteronuclear NMR spectra were recorded at 9.4 T on a Bruker AM400 spectrometer. A natural abundance HMQC (Bax & Hawkins, 1983) spectrum was recorded with proton detection to assign the carbon signals. Chemical shifts were referenced to internal 2,2 dimethylsilapentane-5-sulfonate at 0 ppm.

### *Relaxation times*

$^{13}C$ relaxation times were recorded using inversion recovery, Hahn spin echo, and gated decoupler experiments for $T_1$, $T_2$, and the NOE, respectively. The relaxation data for $T_1$ and $T_2$ were analyzed by nonlinear regression to the appropriate exponential equations, and the NOE was calculated from the ratio of the intensity of the difference spectrum to that in the control (i.e., no NOE) spectrum. Dynamic parameters were analyzed using the formalism of Lipari and Szabo (1982) as follows. The target function was defined as

$$T = (R_1c - R_1o)^2/R_1o^2 + (R_2c - R_2o)^2/R_2o^2(\text{NOE}c - \text{NOE}o)^2/\text{NOE}o^2 \qquad (4)$$

where $R_1$ and $R_2$ are the spin-lattice and spin-spin relaxation rate constants, respectively, and the scripts $o$ and $c$ refer to observed and calculated, respectively. The calculated relaxation parameters

were determined according to the standard formulas for dipolar relaxation:

$$R_1 = a/r^6 CH[J(\Delta\omega) + 3J(\omega_c) + 3J(\omega_H) + 6J(\Sigma\omega)] \quad (5)$$

$$R_2 = a/2r_{CH}^6[4J(0) + J(\Delta\omega) + 3J(\omega_c) + 3J(\omega_H) + 6J(\Sigma\omega)] \quad (6)$$

$$NOE = 4[6J(\Sigma\omega) - J(\Delta\omega)]/R_1 \quad (7)$$

where $r_{CH}$ is the carbon-proton bond length (1.095 Å) and $a = 3.56$ Å$^6$ ns$^{-1}$. The spectral density functions are defined as

$$J(\omega, \tau) = \tau/(1 + \omega^2\tau^2) \quad (8)$$

In the Lipari–Szabo model, the relaxation rates are given as the sum as two terms.

$$R_i = S^2R_i(\omega, \tau_0) + (1 - S^2)R_i(\omega, \tau_e) \quad (9)$$

where $S^2$ is an order parameter, $\omega$ is the Larmor frequency, $\tau_0$ is the rotational correlation time for overall tumbling. $\tau_e$ is a composite correlation time, $\tau_0\tau_i/(\tau_0 + \tau_i)$, where $\tau_i$ is the correlation time for internal motion that must be fast on the Larmor timescale. For each resolved carbon atom, the data were searched exhaustively over all three parameters in intervals of 0–1 for $S^2$, 0–0.4 ns for $\tau_i$, and 1–5 ns for $\tau_0$, and the calculated values of the relaxation parameters were compared with the observed values as described by Equation 4. This procedure identifies all sets of parameters that are consistent with the data, and an estimate of the range for each parameter within some criterion can be estimated. We have used the criterion that all parameters for which the calculated values lie within one standard deviation of the experimental values are accepted.

### *Hydration*

Exchange of NH and $NH_2$ protons with water was monitored by NOESY and ROESY spectroscopy in $H_2O$. The relative magnitude and sign of the cross peaks with water defines the timescale of the exchange. Similarly, the sign and magnitude of cross peaks between nonexchangeable protons and water were determined to assess aspects of hydration of the glycopeptide as described in detail for DNA and RNA (Conte et al., 1996; Lane et al., 1997).

### *Conformation*

Conformational features were assessed by measuring three-bond coupling constants in 1D, NOESY, and DQF-COSY spectra, and from NOE intensities, all within the framework of information about the dynamics of the system. Molecular models were built using InsightII (MSI Inc., San Diego, California). An extended polypeptide having the appropriate sequence was built. Gal$\beta$ 1–3 GalNAc$\alpha$1-disaccharides were constructed and appended to the O$\gamma$ atom of each Thr residue. Conformations were searched using different restrained molecular dynamics protocols. Different starting structures were constructed, including extended chain, left- and right-handed $\alpha$ helices, and the $3_{10}$ helix. In all of these starting structures, the initial potential energy was very high, and the restraint energy was also large, indicating that these models poorly represent the kinds of conformations present in solution. The initial models were first energy minimized for 500 cycles (using conjugate gradients) without the experimental constraints, using the AMBER force field with the extensions for carbohydrates as developed by Homans (1990) and implemented in the program DISCOVER (MSI Inc., San Diego, California). No electrostatic terms were used, and all other contributions to the energy were scaled to unity (except the 1,4 interactions, which were scaled by 0.5). All calculations were carried out in vacuuo, which in the absence of electrostatic interactions permits a wide range of conformations to be sampled efficiently. The system was then heated to 600 K for 1 ps, which effectively randomizes the coordinates. The constraints were then applied, followed by cooling to 300 K and a further 25 ps restrained dynamics. The system was finally energy minimized with 5,000 cycles of conjugate gradient refinement. NOESY cross-peak intensities at different mixing times at which they were observed were classified as strong (50 ms), medium (100, 200 ms), and weak (500 ms), corresponding to upper bounds of 2.6, 3.5, and 4.2 Å, respectively. In addition the weak NOEs were given a lower bound of 3 Å. The force constant for both upper and lower bound was set at 50 kcal mol$^{-1}$ Å$^{-2}$. The $\phi$ angles (Thr residues only) were restrained in the range $-150°$ to $-80°$ with a force constant of 50 kcal mol$^{-1}$ rad$^{-2}$. $\chi_1$ (Thr residues) were restrained between 1 and 100°, also with a force constant of 50 kcal mol$^{-1}$ rad$^{-2}$.

Structures were compared within Insight II, and the torsion angles were listed for determining those conformational features that were consistent from structure to structure.

## Acknowledgments

This work was supported by the MRC, NIH Grant IRO1HL57810-01, and NSF grant No. IBN 93-08581. NMR experiments were carried out at the MRC Biomedical NMR Centre, Mill Hill.

## References

Ahlgren JA, DeVries AL. 1984. Comparison of antifreeze glycopeptides from several Antarctic fishes. *Pol Biol 3*:93–97.

Arav A, Rubinsky B, Fletcher G, Seren E. 1993. Cryogenic protection of oocytes with antifreeze proteins. *Mol Repr Devel 36*:488–493.

Bax A, Davis DG. 1985. MLEV-17 based two-dimensional homonuclear magnetisation transfer spectroscopy. *J Magn Reson 65*:355–360.

Bax A, Hawkins BL. 1983. Correlation between proton and nitrogen-15 chemical shifts by multiple quantum NMR. *J Magn Reson 55*:301–315.

Burcham TS, Osuga DT, Narasinga Rao BN, Bush CA, Feeney RE. 1986a. Purification and primary sequences of the major arginine-containing antifreeze glycopeptides from the fish of *Eleginus gracilis*. *J Biol Chem 261*:6384–6389.

Burcham TS, Osuga, DT, Yeh Y, Feeney RE. 1986b. A kinetic description of antifreeze glycoprotein activity. *J Biol Chem 261*:6390–6397.

Bush CA, Feeney RE. 1986. Conformation of the glycotripeptide repeating unit of antifreeze glycoprotein of polar fish as determined from the fully assigned n.m.r. spectrum. *Int J Pept Prot Res 28*:386–397.

Bush CA, Ralapati S, Matson SM, Osuga DT, Yeh Y, Feeney RE. 1984. Conformation of the antifreeze glycoprotein of polar fish. *Arch Biochem Biophys 232*:624–631.

Conte MR, Conn GL, Brown T, Lane AN. 1996. Hydration of the RNA duplex r(CGCAAAUUUGCG)$_2$ determined by NMR. *Nucl Acids Res 24*:3693–3699.

Daragan VA, Mayo KH. 1997. Motional analysis of protein and peptide dynamics using $^{13}$C and $^{15}$N NMR relaxation. *Prog NMR Spectr 31*:63–105.

Davies PL, Sykes BD. 1997. Antifreeze proteins. *Curr Op Str Biol 7*:828–834.

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

Dill K, Huan L, Bearden DW, Feeney RE. 1992. Structural studies of Antarctic fish antifreeze glycoproteins by one- and two-dimensional NMR spectroscopy. *J Carb Res 11*:499–517.

Drewes JA, Rowen KL. 1993. Evidence for a γ-turn in antifreeze glycopeptides. *Biophys J 65*:985–991.

Feeney RE. 1988. Inhibition and promotion of freezing: Fish antifreeze proteins and ice-nucleating proteins. *Comm Agric Food Chem 1*:147–181.

Hays LM, Feeney RE, Crowe LM, Crowe JH, Oliver AE. 1996. Antifreeze glycoproteins inhibit leakage from liposomes during thermotropic phase transitions. *Proc Natl Acad Sci USA 93*:6835–6840.

Hew CL, Yang DC. 1992. Protein interaction with ice. *Eur J Biochem 203*: 33–42.

Homans SW. 1990. A molecular mechanics force field for the conformational analysis of oligosaccharides: Comparison of theoretical and crystal structures of Man α1-3 Man β1-4 GlcNac. *Biochemistry 29*:9110–9118.

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

Lane AN, Jenkins TC, Frenkiel TA. 1997. Hydration and solution structure of $d(CGCAAATTTGCG)_2$ and its complex with propamidine from NMR and molecular modelling. *Biochim Biophys Acta 1350*:205–220.

Leroy J-L, Broseta D, Gueron M. 1985. Proton exchange and base-pair kinetics of poly(rA).poly(rU) and poly(rI).poly(rC). *J Mol Biol 184*:165–178.

Lipari A, Szabo G. 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.

Mimura M, Yamamoto Y, Inoue Y, Chujo R. 1992. N.M.R. study of interaction between sugar and peptide moieties in mucin-type model glycopeptides. *Int J Biol Macromole 14*:242–248.

Nicholson LK, Kay LE, Torchia DA. 1996. Protein dynamics as studied by solution NMR techniques. In Sarkar SS, ed. *NMR spectroscopy and its application to biomedical research*. Amsterdam: Elsevier. pp 241–279.

Nieto PM, Birdsall BB, Morgan WD, Frenkiel TA, Gargaro AR, Feeney J. 1997. Correlated bond rotations in interactions of arginine residues with ligand carboxylate groups in protein ligand complexes. *FEBS Lett 405*:16–20.

Piotto M, Saudek V, Sklenar V. 1992. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol Str 2*:661–665.

Poppe L, van Halbeek H. 1994. NMR spectroscopy of hydroxyl protons in supercooled carbohydrates. *Nat Struct Biol 1*:215–216.

Rao BNN, Bush CA. 1987. Comparison by $^1H$ NMR spectroscopy of the conformation of the 2600 dalton antifreeze glycopeptide of polar cod with that of the high molecular weight antifreeze glycoprotein. *Biopolymers 26*:1227–1244.

Rubinsky B, Arav A, DeVries AL. 1991. Cryopreservation of oocytes using directional cooling and antifreeze glycoproteins. *Cryo Lett 7*:93–106.

Schwalbe H, Fiebig KM, Buck M, Jones JA, Grimshaw SB, Spencer A, Glaser SJ, Smith LJ, Dobson CJ. 1997. Structural and dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulants of lysozyme in 8 M urea. *Biochemistry 36*:8977–8991.

States DJ, Haberkorn RA, Ruben DJ. 1982. A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants. *J Magn Reson 48*:286–292.

Tablin F, Oliver AE, Walker NJ, Crowe LM, Crowe JH. 1996. Membrane phase transition of intact human platelets: Correlation with cold-induced activation. *J Cell Physiol 168*:305–313.

van Halbeek H, Dorland L, Vliegenthart JFG, Kochetkov NK, Arbatsky NP, Derevitskaya VA. 1982. Characterization of the primary structure and the microheterogeneity of the carbohydrate chains of porcine blood-group H substance by 500 MHz $^1H$ NMR spectroscopy. *Eur J Biochem 127*: 21–29.

Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD. 1995. $^1H$, $^{13}C$ and $^{15}N$ random coil NMR shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J Biomolec NMR 6*:67–81.

Wüthrich K. 1986. *NMR of proteins and nucleic acids*. New York: Wiley.

Yeh Y, Feeney RE. 1996. Antifreeze proteins: Structures and mechanisms of function. *Chem Rev 96*:601–617.