Solution structure of antifreeze glycopeptides

Determination of the major conformers of the glycosidic linkages

S.W. Homans, A.L. DeVries[+] and S.B. Parker*[°]

Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, England, [+]Department of Physiology and Biophysics, 524 Burrill Hall, 407 South Goodwin, University of Illinois, Urbana, IL 61801, USA and *[°]Department of Inorganic Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QR, England*

Received 24 January 1985

Using data obtained from ^{1}H-^{1}H correlated spectroscopy and one- and two-dimensional NOE measurements, we have determined the preferential solution conformation of the disaccharide unit Gal β(1-3)GalNAc in the antifreeze glycopeptide (Galβ(1-3)GalNAcα1-Thr-Ala-Ala)$_n$, derived from the blood of the antarctic fish *Trematomas borchgrevinki*. The disaccharide, which is a rigid unit, exists with a restricted orientation about the GalNAc-α1-Thr linkage. Preliminary evidence indicates that the backbone (Ala-Ala-Thr)$_n$ lies in an extended helical conformation.

Antifreeze glycopeptide *^{1}H NMR* *Glycosidic linkage* *Dissaccharide structure* *Nuclear Overhauser effect*

1. INTRODUCTION

Certain marine teleost fish species have evolved the ability to live in polar waters with a winter temperature of -1.9°C, over 1°C lower than the normal freezing point of water in a fish in temperate waters (-0.8°C) [1]. These fish have an increased level of electrolytes, but the colligative contribution to the depression of the freezing point of the body fluids remains inadequate by about 0.5°C to prevent ice formation and subsequent death [2]. Over half the effect has been shown to be due to a series of macromolecules with molecular masses between 2400 and 34000 Da [3]. In most antarctic fish and many arctic species these are glycopeptides, although some arctic and northern temperate species have peptides. They have been shown to depress the freezing point 200–300-times more than their molality would predict on a colligative basis alone, although they have no known effect on the melting point of ice [4].

Antifreeze glycopeptides were first isolated from the blood of *Trematomus* spp. [5] and shown to contain 8 fractions; AFGP 1–5 having a repeating sequence of the tripeptide alanyl-alanyl-threonine with the disaccharide β-D-galactopyranosyl-(1→3)-2-acetamido-2-deoxy-α-D-galactopyranose glycosidically linked to the threonine [6]. These AFGPs differ only in molecular mass; fractions 6–8 contain proline in place of the first amino acid of each tripeptide from alanine residue 7 [7]. This renders the fractions inactive in in vitro experiments, although a cooperative effect in vivo has been postulated [8]. The AFGPs are present at a concentration of about 3.5% in the blood, urine and cytosolic fluid of many species of antarctic fish and also some, otherwise unrelated, arctic species.

* To whom correspondence should be addressed at (present address): Unilever Research, Colworth Laboratory, Colworth House, Sharnbrook, Bedford MK44 1LQ, England

Abbreviations: NOE, nuclear Overhauser effect; DSS, sodium 2,2-dimethyl-2-silapentane-5-sulphonate; GalNAc, *N*-acetylgalactosamine; COSY, correlated spectroscopy; AFGP, antifreeze glycopeptide

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/85/$3.30

We have noted the work of Bush et al. [9] on the solution conformation of the AFGPs based upon an NMR study of fraction 8. The spectrum of fractions 3–5 is significantly simpler than that of fraction 8, since the former is homogeneous. This has allowed the assignment of the majority of proton resonances by one- and two-dimensional spectroscopy and NOE measurements for the first time.

Below we deduce from first principles, the solution structures of the saccharide moieties of the major fractions of AFGPs from antarctic fish.

2. EXPERIMENTAL

2.1. *Preparation of glycoprotein samples*

The sample of AFGP was isolated from *Trematomas borchgrevinki* by the method of DeVries et al. [10] and consisted of fractions 3, 4 and 5, differing only in molecular mass, which was 23.5, 17.5 and 10.5 kDa, respectively. The sample was exhaustively dialysed, dissolved in and lyophilised from ^2H_2O (99.96% Aldrich) to exchange adventitiously bound water and labile protons. The sample was finally dissolved in ^2H_2O to form a 35 $mg \cdot ml^{-1}$ solution.

2.2. *^1H NMR spectrometers and methods*

Two-dimensional FT ^1H NMR was performed at 500 MHz at a probe temperature of 300 K using an Oxford Instruments superconducting magnet with the Oxford Enzyme Group Spectrometer, employing a pulsed deuterium lock, a Nicolet 1180 microcomputer and a Nicolet 293 B pulse programmer. Chemical shifts are given relative to DSS.

2D-^1H homonuclear correlation experiments [11] were conducted with the basic pulse cycle $90°(+x)-t_1-\theta(\phi)$-acquisition (t_2), where $\phi = +x, +y, -x, -y$ and $\theta = 90°$. For each t_1 value 64 transients were collected. Coherence transfer echoes [12] were selectively detected by the subtraction of alternate scans of the form $90°-t_1-\theta°(\phi)-t_2$ and $90°-t_1-\theta(\phi+90°)-t_2$, where $\phi = 0$ or $180°$. 1024×2 K spectra (sweep width = ± 900 Hz) were collected with a minimum evolution time (t_1) of 5 μs and an increment of 555 μs to t_1 between spectra. Prior to Fourier transformation, the time domain matrix was multiplied in both the t_1 and t_2 dimensions with phase-shifted sine-bell functions. The use of phase cycling to achieve quadrature phase detection in both dimensions allowed the transmitter frequency to be set to the middle of the spectrum and thus reduced the size of the data matrix.

One-dimensional NOEs were obtained at 300 MHz as described [13]. 2D NOE measurements were performed at 500 MHz at a probe temperature of 298 K using the following pulse sequence:

$$90_\phi-t_1-90_\phi-\tau_m-90_\phi-t_2$$

The phases of the first two pulses were cycled in 90° steps to cancel magnetisation arising from double quantum coherence. In addition, the phases of the first and third pulses were cycled in 90° steps to cancel both axial magnetisation and transverse interference arising from relaxation during the mixing period, τ_m. Finally, the phases of the second and third pulses were cycled in 90° steps to achieve quadrature phase detection in ω_1. In total, 32 transients (sweep width ± 1000 Hz, 2048 data points) were collected for each of 256 increments in t_1 (500 μs). Prior to Fourier transformation, time domain signals were weighted with phase-shifted sine-bell functions.

3. RESULTS

The 500 MHz ^1H NMR spectrum of the antifreeze glycopeptides, consisting of fractions 3, 4 and 5 (see section 2) is shown in fig.1. Typically, the majority of the carbohydrate proton resonances can be found in the region

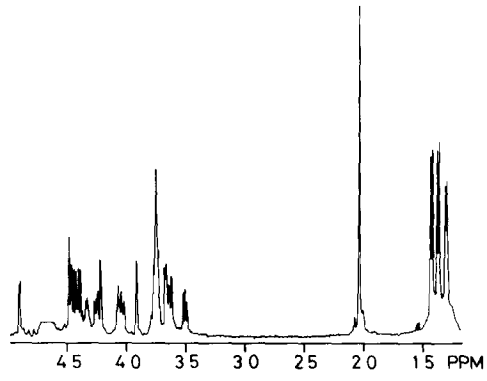

Fig.1. 500 MHz ^1H NMR spectrum of antifreeze glycopeptides (fractions 3–5; see text). The strong HOD resonance (~4.7 ppm) and the Tris resonance (~3.7 ppm) have been truncated.

3.5–4.5 ppm but the anomeric (H1) protons are found at low field. Thus, the well resolved resonance at 4.92 ppm is assigned to GalNAc α H1, by virtue of its chemical shift position and multiplicity (determined by $^3J_{1,2}$). These are characteristic of GalNAc in α-1-*O* linkage [14]. A general strategy in the assignment of oligosaccharide proton resonances is to correlate sequentially the shift positions of all the ring protons from the anomeric proton resonance using COSY [15]. While in general, this strategy requires support from other spectral editing techniques, such as multi-step relayed correlation spectroscopy [16] or double-quantum NMR spectroscopy in the case of oligosaccharide mixtures [17], the chemical shift dispersion of the GalNAc ring protons were found to be sufficient to obtain the majority of assignments using COSY alone. Thus, by the usual correlation of crosspeak vs diagonal peak shift positions in the COSY spectrum of the glycopeptides (not shown) the H1-H5 proton resonance assignments were determined. These are shown in table 1. Assignments of alanine αCH proton resonances and threonine α and β CH proton resonances were determined by correlation via the resolved methyl resonances of these residues at ~1.4 ppm. Low-intensity doublets were observable in the region ~1.4 ppm. These probably correspond to the methyl proton resonances of terminal Ala and Thr residues. These have not been interpreted in the present study. Assignment of Galβ proton resonances was made by straightforward elimination of the crosspeaks belonging to the residues described above, and by sequential correlation of remaining crosspeaks using the low-field GalβH1 proton resonance (4.48 ppm) as a basis. It should be noted that it was not possible to assign the H6 protons due to the complexity of the spectrum in this region, together with the overlapping Tris resonance at 3.73 ppm. The assignments obtained using the above procedures are listed in table 1. Despite heterogeneity in glycopeptide molecular mass, the resonance positions of similar protons in each population are degenerate. The inference therefore, is that the magnetic environments of the protons in each population are very similar if not the same. Furthermore, with the exception of terminal Ala and Thr, intra-population protons at similar locations within the repeating sequence of the glycopeptide have degenerate chemical shifts. These observations are of importance in the determination of the overall conformation of the molecule (vide infra).

Table 1

500 MHz ^{1}H chemical shifts (δ, ppm; relative to acetone, δ = 2.225 ppm at 25°C) of the antifreeze glycopeptides from *T. Borchgrevinki*

Gal	H1	H2	H3	H4	H5	
	4.48	3.52	3.63	3.93	~3.67	
GalNAc	H1	H2	H3	H4	H5	CH_3
	4.92	4.26	4.04	4.23	4.08	2.01
Thr	αCH	βCH	γCH_3			
	4.34	4.49	1.30			
Ala	αCH	βCH_3				
	4.40	1.39				
Ala	αCH	βCH_3				
	4.44	1.42				

(~) Approximate value due to overlap with Tris resonance

Due to partial overlap of both disaccharide and peptide backbone protons, it was not possible to irradiate selectively a particular resonance in many cases in one-dimensional NOE studies. Accordingly, a two-dimensional NOE spectrum was recorded (not shown), and thus a qualitative map of proton through-space connectivities was available with which to screen the observed resonances in one-dimensional NOE difference spectra for non-specific NOEs.

Irradiation of the H1 resonance of αGal H1 shows small intra-residue NOEs to both Gal H3 and Gal H5. In addition, a strong inter-residue NOE is observed between Gal H1 and GalNAc H3, together with a weak NOE to GalNAc H4. The relative magnitudes of those inter-residue and intra-residue NOEs implies that the preferred solution conformation of the β1-3 linkage results in close proximity (~3 Å) of Gal H1 and GalNAc H3. Irradiation of GalNAc H1, shows a strong intra-residue NOE to GalNAc H2, together with strong inter-residue NOEs to Thr αCH, βCH and γCH_3. Since there is no single conformation about the GalNAc α-Thr linkage which could simultaneous-

ly generate all 3 intra-residue NOEs, it is probable that there is restricted motional averaging about this linkage or that the conformations of the distinct glycosidic linkages differ slightly. Finally, we have observed NOEs which can only be satisfactorily explained as inter-residue NOEs between distinct disaccharide units. These, together with NOE connectivities within the peptide backbone, will be of importance in the determination of the overall solution conformation of the molecule, and are presently under investigation.

4. DISCUSSION

Determination of the solution structure of the antifreeze glycoproteins from antarctic fish has been attempted by several techniques [18,20] and, most recently by ^{1}H NMR [9]. The nature of the glycoproteins isolated from fish means that the choice of sample is vital for meaningful interpretation. The choice of fractions 3–5 for this study produces a solution with glycoproteins of no more than 23.5 kDa, which is important for the resolution of the spectrum and also to minimise the spread of correlation times for an NOE study. The fraction used also contains no AFGPs with proline, such as fraction 8. The information from fraction 8 is limited by the heterogeneity of the sample; the prolines fit into different positions in the backbone and the resulting mixture of identically sized glycopeptides cannot be separated. The study of fractions 3–5 has enabled us to assign the spectrum from first principles including the methyl groups from the 3 different amino acid positions and their corresponding α-CHs. These unambiguous assignments have allowed us to use NOE measurements with which to determine the preferred conformation of the saccharide units.

The NOE results at room temperature do not provide an unambiguous conformation of the disaccharide unit relative to the peptide backbone; preliminary results of NOEs at physiological temperatures indicate that one conformation is preferred, producing a structure in which the alanines have different exchange times distinct from the other residues, implying that they play little part in the antifreeze function of the glycoprotein.

The conformation of the β(1-3) linkage can be determined by the strength of the NOE between βGal-H1 and GalNAc-H3, which may be affected by conformational averaging, relative to that of intra-residue NOEs from βGal-H1 to βGal-H3 and -H5; the fact that the inter-residue NOE is larger than intra-residue effect implies that there is little or no averaging and that the two protons are adjacent (within 3 Å).

The ^{1}H NMR spectrum of fractions 3–5 implies that the total structure of the glycopeptides is the sum of the solution structure of its individual units. This indicates that the AFGPs exist in an extended conformation in solution. Furthermore, this conformation is possibly stabilised by inter-disaccharide interactions, as evidenced by our observation of inter-disaccharide NOEs. This would imply a loose helical arrangement to align the disaccharide units. Confirmation of this postulate will require careful measurement of NOEs within the Ala-Ala-Thr backbone in H_2O. These measurements are currently in progress.

ACKNOWLEDGEMENTS

The authors wish to thank Drs R.A. Dwek, T.W. Rademacher and Professor R.J.P. Williams for encouragement and support. S.W.H. is a member of the Oxford Enzyme Group. This work was supported in part by a grant from Monsanto Co., USA.

REFERENCES

[1] Black, V.S. (1951) Univ. Toronto Biol. Ser. 59, 53–89.
[2] O'Grady, S.M. and DeVries, A.L. (1982) J. Exp. Marine Biol. Ecol. 57, 219–228.
[3] Feeney, R.E. and Yeh, Y. (1978) Adv. Protein Chem. 32, 191–282.
[4] DeVries, A.L. (1971) Science 172, 1152–1155.
[5] DeVries, A.L. and Wohlschlag, D.E. (1967) Science 163, 1074–1075.
[6] DeVries, A.L., Vandenheede, J. and Feeney, R.E. (1971) J. Biol. Chem. 246, 305–308.
[7] Lin, Y., Duman, J.G. and DeVries, A.L. (1972) Biochem. Biophys. Res. Commun. 46, 87–92.
[8] Mulvihill, D.M., Geoghegan, K.F., Yeh, Y., DeRemer, K., Osuga, D.T., Ward, F.C. and Feeney, R.E. (1980) J. Biol. Chem. 255, 659–662.
[9] Bush, C.A., Ralapati, S., Matson, G.M., Yamasaki, R.B., Osuga, D.T., Yeh, Y. and Feeney, R.E. (1984) Arch. Biochem. Biophys. 232, 624–631.

[10] DeVries, A.L., Komatsu, S.K. and Feeney, R.E. (1970) J. Biol. Chem. 245, 2901–2908.
[11] Aue, W.P., Bartholdi, E. and Ernst, R.R. (1976) J. Chem. Phys. 64, 2229–2246.
[12] Maudsley, A.A., Wokaun, A. and Ernst, R.R. (1978) Chem. Phys. Lett. 55, 9–14.
[13] Homans, S.W., Dwek, R.A., Fernandes, D.L. and Rademacher, T.W. (1982) FEBS Lett. 150, 503–506.
[14] Gejyo, F., Chang, J.-L., Burgi, W., Schmid, K., Offner, G.D., Van Halbeck, H., Dorland, L., Gerwig, G.J. and Vliegenthart, J.F.G. (1983) J. Biol. Chem. 258, 4966–4971.
[15] Homans, S.W., Dwek, R.A., Fernandes, D.L. and Rademacher, T.W. (1983) Biochim. Biophys. Acta 760, 256–261.
[16] Homans, S.W., Dwek, R.A., Fernandes, D.L. and Rademacher, T.W. (1984) Proc. Natl. Acad. Sci. USA 81, 6286–6289.
[17] Homans, S.W., Dwek, R.A., Fernandes, D.L. and Rademacher, T.W. (1984) Biochim. Biophys. Acta 798, 78–83.
[18] Ahmed, A.I., Feeney, R.E., Osuga, D.T. and Yeh, Y. (1975) J. Biol. Chem. 250, 3355–3347.
[19] Franks, F. and Morris, E.R. (1978) Biochim. Biophys. Acta 540, 346–356.
[20] Berman, E., Allerhand, A. and DeVries, A.L. (1980) J. Biol. Chem. 255, 4407–4410.