PERGAMON

Phytochemistry 53 (2000) 135–144

PHYTOCHEMISTRY

www.elsevier.com/locate/phytochem

Sequence specific analysis of the heterogeneous glycan chain from peanut peroxidase by ^{1}H-NMR spectroscopy

Gary S. Shaw[a], Yan Sun[b], Kathryn R. Barber[a], Robert B. van Huystee[b,*]

[a]*Departments of Biochemistry, The University of Western Ontario, London, Ont., Canada, N6A 5B7*
[b]*Departments of Plant Sciences, The University of Western Ontario, London, Ont., Canada, N6A 5B7*

Received 18 September 1998; received in revised form 13 July 1999

Abstract

The cationic peanut peroxidase is a complex enzyme consisting of a heme group, two calcium ions and three complex carbohydrate chains at positions Asn60, 144 and 185. Details of the heme and calcium ligation, necessary for oxidation, have recently been revealed from the three-dimensional structure of the peroxidase. However, the three glycans that may be important for the stability of the enzyme as well as its activity were not resolved. In order to determine the configuration of one of these glycans, PNGase A was used to cleave the glycan from the enzyme at Asn-144. This glycan was studied by two dimensional ^{1}H-NMR spectroscopy to identify the sugar linkages. The results indicated a glycan structure comprising a Manα1-6(Xylβ1-2)Manβ1-4GlcNAcβ1-4(Fucα1-3)GlcNAcβ core but with an additional Manα1-3 appendage linked to Man3. The glycan also appeared to be heterogeneous as was noted from a single terminating galactose being linked to approximately 20–25% glycan.

Keywords: Peanut peroxidase; Arachis hypogaea; Leguminosae; Glycan; NMR spectroscopy; Primary structure; Conformation

1. Introduction

Glycosylation of proteins represents one of the most important post-translational events because of the ubiquity of the phenomenon (Montreuil, 1995), yet much of glycoprotein structure and function, especially of plant glycoproteins, has yet to be examined (Sturm, 1995). The glycans of storage proteins have been examined (Lerouge, Cabanes-Macheteau, Rayon, Fichette-laine, Gomord & Faye, 1998), but information for extracellular glycoproteins is still lacking. While some protein structures of plant glycoproteins, as well as the terminal sugar (Heese-Peck & Raikhel, 1998) are known, little has been learned regarding the overall glycan structure, particularly of large complex glycans. Some glycan structures are known for animal glycoproteins (Montreuil, 1995); their composition and structure are altered in diseases (Montreuil, 1995), and may contribute to protein recognition.

Cationic peanut peroxidase (EC 1.11.1.7) (Arachis hypegaea, Leguminosae) secreted in cell suspension medium, is a heme and calcium containing glycoprotein. The locations and roles for the former two prosthetic groups in this enzyme (Schuller, Ban, van Huystee, McPherson & Poulos, 1996; Rodriguez Maranon & van Huystee, 1994) have been determined, but little is known about the structure and function of glycans (van Huystee & Wan, 1994). Three *N*-linkage sites have been mapped on the polypeptide chain, at Asn-60, -144 and -185 (Wan & van Huystee, 1993). Based on sugar composition and methylation assays, these glycans appear to be complex, containing GlcNAc, mannose, galactose, xylose and fucose sugars (van Huystee, Sesto & O'Donnell, 1992). The loss of terminal galactose from the antennae, brought about by the cosecreted β-galactosidase (Wan, Gijzen & van Huystee, 1994) does not appear to affect enzyme

* Corresponding author. Tel.: +1-519-679-2111; fax: +1-519-661-3935.

E-mail address: huystee@julian.uwo.ca (R.B. van Huystee).

0031-9422/00/$ - see front matter
PII: S0031-9422(99)00472-0

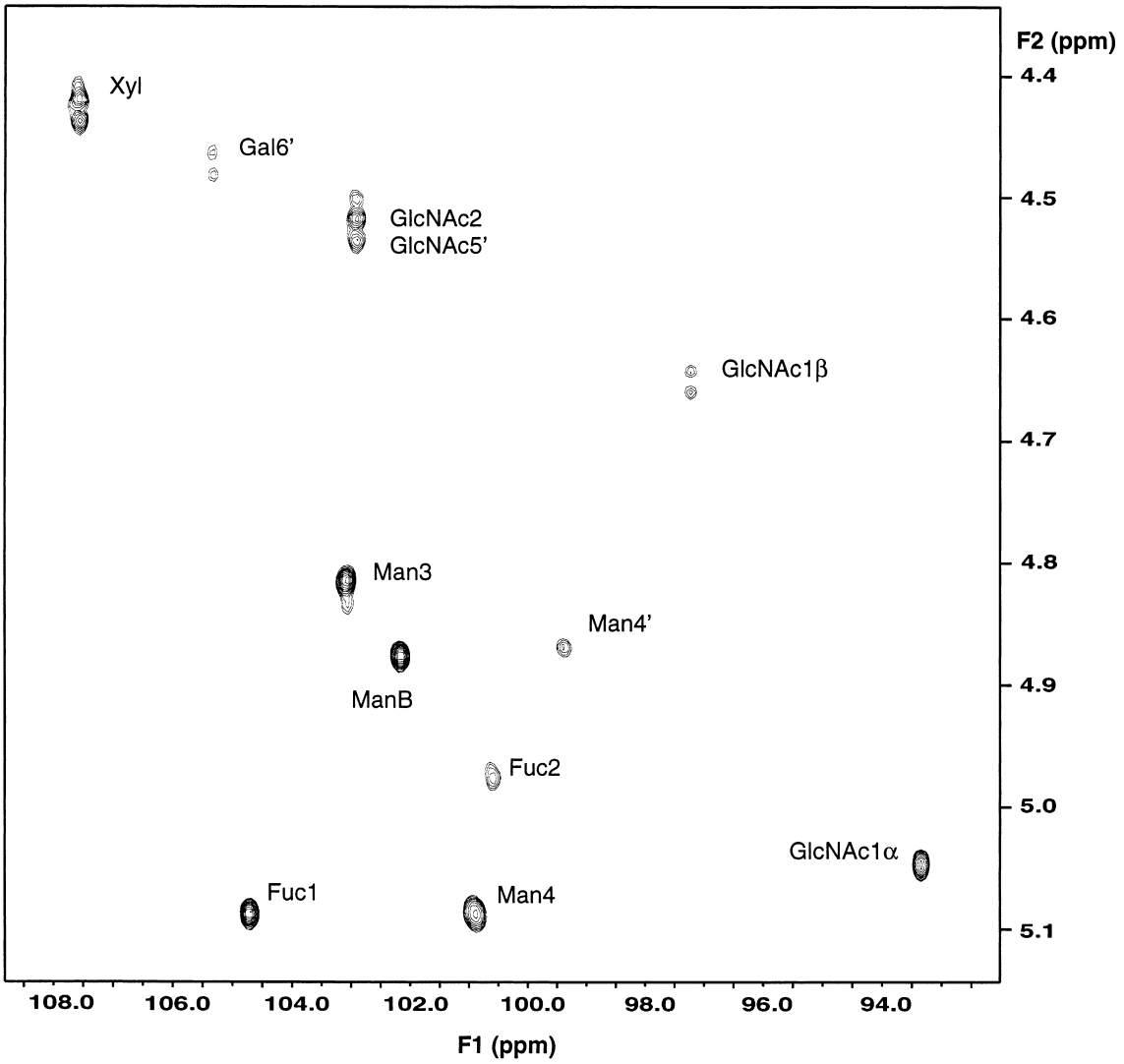

Fig. 1. Portion of the natural abundance ^{1}H–^{13}C HMQC spectrum of 1 mM glycan in D_2O at pH 6.3 and 25°C. The contour plot shows the anomeric region of the glycan with the correlations between ^{1}H and ^{1}C indicated by the assigned sugar position. For Xyl, Gal6′ GlcNAc2 and GlcNAc5′ and GlcNAc1β the β conformation can be identified from ^{1}H–^{1}H splitting in the F2 dimension where $^{3}J_{1,2} > \sim 7$ Hz.

activity (O'Donnell, Wan & van Huystee, 1992). However treatment of the peroxidase with PNGase F, an endo-glycopeptidase (EC 3.5.1.52), causes loss of peroxidative activity by the enzyme, and this loss of activity is not due to protease, nor the total loss of a glycan (van Huystee & Wan, 1994). Further, tunicamycin treatment of suspension cultured cells results in a loss of the glycans and a loss of secretion of peroxidase (Ravi, Hu, Reddi & van Huystee, 1986). A monoclonal antibody directed against one of the three glycans (Hu & van Huystee, 1988) inhibits enzyme activity (Hu, Carbonera & van Huystee, 1987). These results clearly indicate the glycans in peanut peroxidase are functionally important. However the lack of detailed structural information of the glycans makes a comprehensive understanding difficult.

The sequences and primary structure of very few plant glycans are known (van Huystee & Wan, 1994). For example the structure of a 6-sugar bromelain, a cellular pineapple protease, has been examined by two dimensional ^{1}H-NMR spectroscopy (Lommerse, Kroon-Batenburg, Kamerling & Vliegenthart, 1995). This *N*-glycan contains a Manα1-6(Xylβ1-2)Manβ1-4GlcNAcβ1-4(Fucα1-3)GlcNAcβ core structure, an observation consistent with one of several slightly modified core structures suggested to exist for plant glycoproteins (Sturm, 1991), although most are of a more complex structure. This information has been mainly derived from sugar compositional analysis which provides no direct information about the glycan primary structure or the linkages between sugars. The determination of glycan structure remains a major challenge that is complicated by the occurrence of

micro-heterogeneity of sugars and residual glycosidase activities in protein preparations (Wan et al., 1994).

For cationic peanut peroxidase, two glycoprotein forms are readily separated by concanavalin-A (Con-A) chromatography (O'Donnell et al., 1992). From sugar composition analyses (Sun, Lige & van Huystee, 1997) both forms are predicted to have a similar core structure but differing terminal sugars. In this work, the glycan linked at position Asn-144 was studied by NMR spectroscopy in an effort to elucidate its primary structure and linkages. For peanut peroxidase, NMR spectroscopy has proven valuable to understanding the enzyme's structure (Barber, Rodriguez Maranon, Shaw & van Huystee, 1995) in the absence of crystallographic data. It is anticipated that the information reported here can be combined with protein crystallographic studies of peanut peroxidase (Schuller et al., 1996) to provide a better understanding of the plant glycan core structure, and of the glycan-protected protein surface (Lerouge et al., 1998) and its role in enzyme function.

2. Results

2.1. Isolation of the glycan from peanut peroxidase

The three glycans were isolated from peanut peroxidase by trypsin digestion, gel-filtration and HPLC separation of the glyco-tryptic fragments. The first glycan, Asn-185 linked, was cleared from its tryptic fragment by pronase, whereas, the other two glycans at positions Asn-144 and Asn-60 were more resistant to cleavage. ^{1}H-NMR spectra of the cleavage product for the Asn-144 linked glycan, for example, revealed multiple doublet resonances near 1.2 ppm and singlets near 2.0 ppm. These peaks, corresponding to the CH_3 protons from fucose and *N*-acetyl groups from GlcNAc sugars, were more numerous than expected based on sugar composition, and exhibited large variations in signal intensity. This observation was most consistent with incomplete digestion by pronase yielding a heterogeneous mixture of different sized glycopeptides. When removal of the Asn-144 glycan from the protein was attempted by treatment with PNGase F (van Halbeek, 1994), the technique recommended for animal glycan isolation (Tarentino & Plummer, 1994), no cleavage was observed. However, treatment with PNGase A (EC 3.5.1.52) (Kurosaka et al., 1991) resulted in clean cleavage of the glycan from the peptide as judged from its ^{1}H-NMR spectrum, which showed two resolved doubletes of similar intensity near 1.2 ppm.

Sugar composition of the Asn-144 linked glycan was determined using trifluoroacetic acid hydrolysis, labelling of the sugars with *P*-aminobenzoic ethyl ester, and subsequent separation by reversed phase HPLC (Sun et al., 1997). The sugar composition data of the glycan indicated a ratio of fucose: *N*-acetyl glucosamine (GlcNAc): mannose: xylose: galactose of 2:4:4:1:1, for a total of 12 sugars. This observation was consistent with previous compositional work suggesting that the glycan comprises a maximum of twelve sugars (Wan et al., 1994).

2.2. ^{1}H-NMR assignment of glycan

Initial NMR analyses were aimed at identifying the spin systems corresponding to the 12 possible sugars determined from the compositional analysis. Our approach was to attempt to identify chemical shifts from the 2 Fuc, 4 GlcNAc, 1 Gal (galactose) and 1 Xyl (xylose) sugar residues since these have some distinguishing properties allowing unequivocal assignment. The remaining four sugars could then be assigned to mannose (Man).

The natural abundance ^{1}H–^{13}C HMQC (heteronuclear multiple quantum correlation) experiment shown in Fig. 1 indicates correlations between the anomeric H1 of each sugar and its attached ^{13}C. The figure indicates several peaks of similar intensity as well as some less intense correlations. This indicated that some microheterogeneity, commonly observed in other complex carbohydrates, was present in the glycan (vide infra). For ^{1}H-NMR assignment purposes the most intense resonances corresponded to the major glycan species present and were of the greatest utility during the analysis. From this data it was clear that 11 resolved correlations were observed, although it was expected that 12 correlations would be present based on sugar composition analysis. These 11 correlations were used for the assignment of the glycan as described below. In general the ^{13}C shift range for most of the C1 chemical shifts of the glycan (~99–106 ppm) were in agreement with previous observations for other carbohydrates (van Halbeek, 1994). There were, however, two exceptions, firstly, the anchoring GlcNAc sugar residue, originally linked to Asn-144, was identified by its unique ^{13}C chemical shift (93.4 ppm), nearly 10 ppm upfield of the other resonances. This peak position results from the free hydroxyl group at C1 in GlcNAc1 (Wyss et al., 1995). Secondly, the single Xyl residue had a unique ^{13}C chemical shift for the C1 (107.4 ppm) downfield of all other C1 resonances, and a unique H1 resonance (4.45 ppm) upfield of all other anomeric H1 resonances. Though few xylose ^{13}C chemical shifts in complex glycans have been reported, this ^{1}H shift in the glycan was nearly identical to that found in bromelain (Lommerse et al., 1995).

The ^{1}H–^{13}C HMQC spectrum provided additional information about the linkages involved at the anomeric carbon, based on the observed residual ^{1}H–^{1}H split-

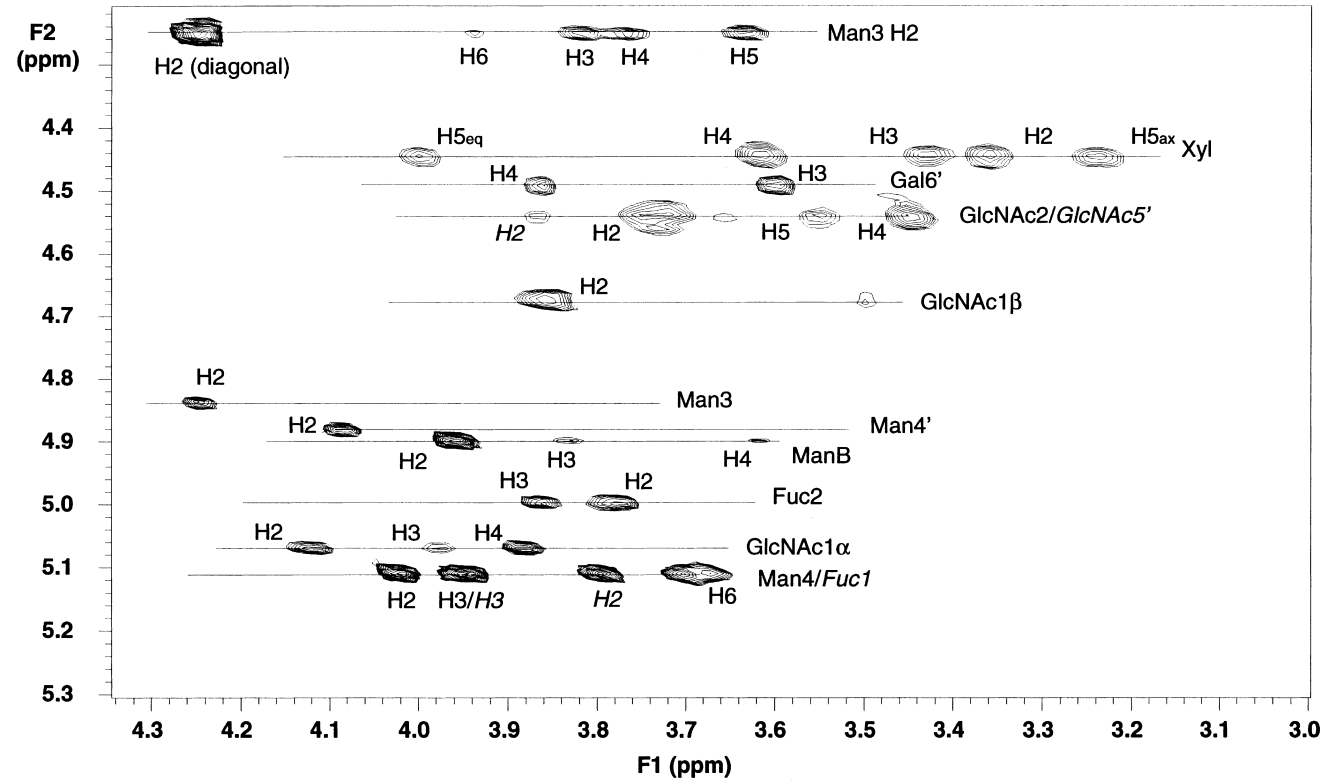

Fig. 2. Part of the contour plot for the 125 msec ^{1}H TOCSY spectrum of 1 mM glycan in D_2O at pH 6.3 at 25°C. In each case the intra-sugar connectivities are indicated by a line passing through the sugar resonances. Sugars are labelled according to the complete assignment. In cases where two sugars have degenerate H1 resonances assignments, one of the sugars is indicated in italics (Fuc1, GlcNAc5′).

tings in the spectrum. For GlcNAc1 the ^{1}H–^{13}C correlation appears as a single peak due to $^{3}J_{1,2} \sim 3.5$ Hz corresponding to the α-anomer of this sugar. Similarily, xylose appears as a doublet with $^{3}J_{1,2} \sim 7.5$ Hz due to a *β*-linkage at its anomeric carbon. One interesting observation was the presence of a weak correlation near a ^{13}C resonance at 8 97.2 ppm. The location of this resonance agreed well with that previously observed for the β-anomer of an initiating GlcNAc residue (Agrawal, 1992; van Rooijen, Kamerling & Vliegenthart, 1998). Further a $^{3}J_{1,2}$ coupling of 8.5 Hz was consistent with the β-configuration. Compared to the GlcNAc1α anomer, the magnitude of this correlation indicated the β-anomer was present in about 10% of the glycan.

Using the ^{1}H–^{13}C HMQC spectrum as a base for the location of the anomeric ^{1}H resonances, the ^{1}H assignment of the glycan was carried out from the

Table 1
Chemical shift data[a] for glycan at 25°C

Residue	H1($^{3}J_{1,2}$) C1 ($^{1}J_{C1,H1}$)	H2	H3	H4	H5	H6, H6′
GlcNAc1	5.07 (3.5) 93.42 (170.6)	4.13	3.98	3.89	3.84	3.77, 3.69
GlcNAc2	4.54 (8.2) 103.08 (169.2)	3.76	3.87	3.45	3.56	3.67, 3.72
Fuc1	5.11 (1.6) 104.94 (171.8)	3.80	3.96	3.69	4.71	1.25
Man3	4.84 (2.1) 103.23 (159.0)	4.26	3.83	3.78	3.64	3.78, 3.95
Man4	5.11 (2.1) 101.05 (170.6)	4.03	3.97	3.82	3.89	3.72, 3.63
Man4′	4.89 (1.7) 99.50	4.09	3.86	3.54	3.74	3.96, 3.73
Xyl	4.44 (7.5) 107.75 (157.5)	3.36	3.43	3.62	4.00, 3.24	
GlcNAc5′	4.55 (8.5) 103.08	3.88	4.05	3.73	3.51	3.97, 3.73
ManB	4.90 (1.8) 102.30 (169.3)	3.97	3.85	3.75	3.63	n.d.[b]
Gal6′	4.49 (7.7) 105.57 (153.9)	3.47	3.60	3.87	3.56	n.d.[b]
Fuc2	5.00 (3.8) 100.74	3.79	3.86	3.78	4.86	1.16

[a] Chemical shifts are reported in ppm and referenced to DSS at 0.00 ppm. Coupling constants reported in Hz.
[b] n.d. = not determined.

TOCSY spectra, measured with spin-locking times of 55 and 125 ms (see Fig. 2). All chemical shifts are listed in Table 1. For several sugars, including GlcNAc1, Fuc2, Man3, Xyl and Gal, the anomeric proton was completely resolved, leading to complete unambiguity in the assignment. The initiating sugar, GlcNAc1, was confirmed in Fig. 2 from strong correlations between H1 and H2, H3 and H4 in the hexose ring, and a weaker correlation with H5. This resulted from efficient magnetization transfer from H2 through to H5 via the relatively large scalar couplings (6–9 Hz) in the sugar ring structure (Inagak, Shimada, Kohda, Suzuki & Bax, 1989). Further identification was made through a TOCSY spectrum in 90% H_2O/10% D_2O (data not shown) which showed correlations from the sidechain NH resonance to protons at positions H1, H2, H3 and H4. The weak correlation observed in the ^1H–^{13}C HMQC was confirmed as the GlcNAc1β anomer through its H1 and H2 resonances, which are shifted 0.4 and 0.25 ppm (respectively) upfield compared to the α-anomer (Agrawal, 1992).

Identification of the anomeric protons for the two fucose sugars was aided from weak TOCSY correlations of H6 CH_3 groups at 1.25 (Fuc1) and 1.17 ppm (Fuc2) to the H1 protons. The chemical shifts for H1 of Fuc1 (5.11 ppm) and Fuc2 (5.00 ppm) were in agreement with values reported in the literature for α-linked fucose rings (Lommerse et al., 1995; Kurosaka, Yano, Itoh, Kuroda, Nakagawa & Kawasaki, 1991; Agrawal, 1992; van Rooijen, Kamerling & Vliegenthart, 1998; Takahashi et al., 1986). In addition, the CH_3 groups provided an efficient means of identifying H5 and H4 protons of the fucose rings in TOCSY spectra, since crosspeaks resulting from H1 to H3 and H4 were very weak even at long spin-locking times (Fig. 2). The observation of the H1–H6 crosspeaks was likely a result of the increased peak intensity afforded from the three methyl protons. In both cases identification of H2 and H5 was confirmed by observation of the coupling between H1–H2 and H5–H6 (CH_3) in DQF-COSY spectra. The characteristic downfield-shifted H5 protons in fucose between 4.7–4.9 ppm were used to complete the assignment, this proved especially useful for Fuc1, where its anomeric proton resonance was nearly degenerate with that of Man4.

The single xylose assignment was confirmed by the observation of a spin system at 4.44 ppm in TOCSY spectra (Fig. 2) which had significantly upfield shifted non-anomeric protons (H2, H3 and $H5_{ax}$) very near

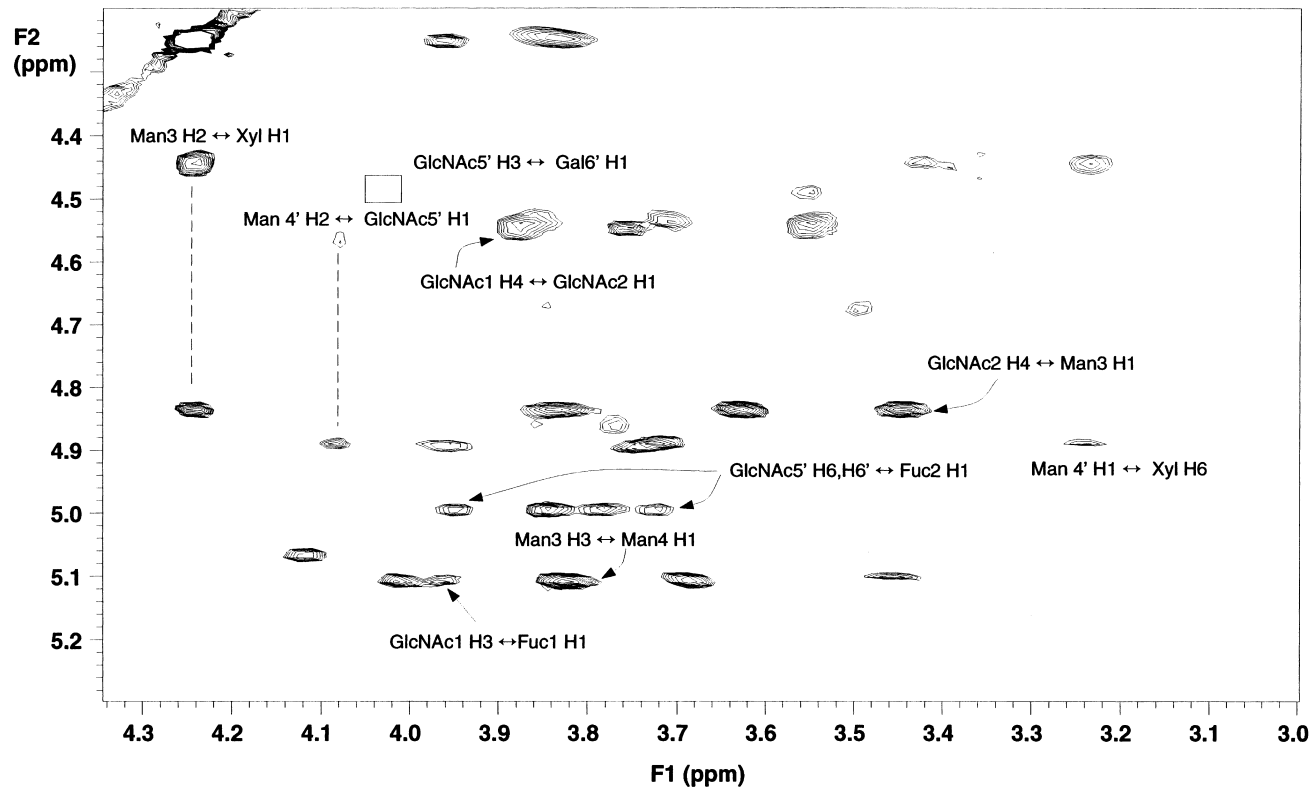

Fig. 3. Portion of the contour plot of the NOESY spectrum of 1 mM glycan recorded using a 300 ms mixing time. For clarity only nOes which correspond to an intersugar interaction are indicated. A box is used to show the position of an nOe between GlcNAc5′ H3 and Gal6′ H1 which is present at a lower contour level. All other conditions are as in Figs. 1 and 2.

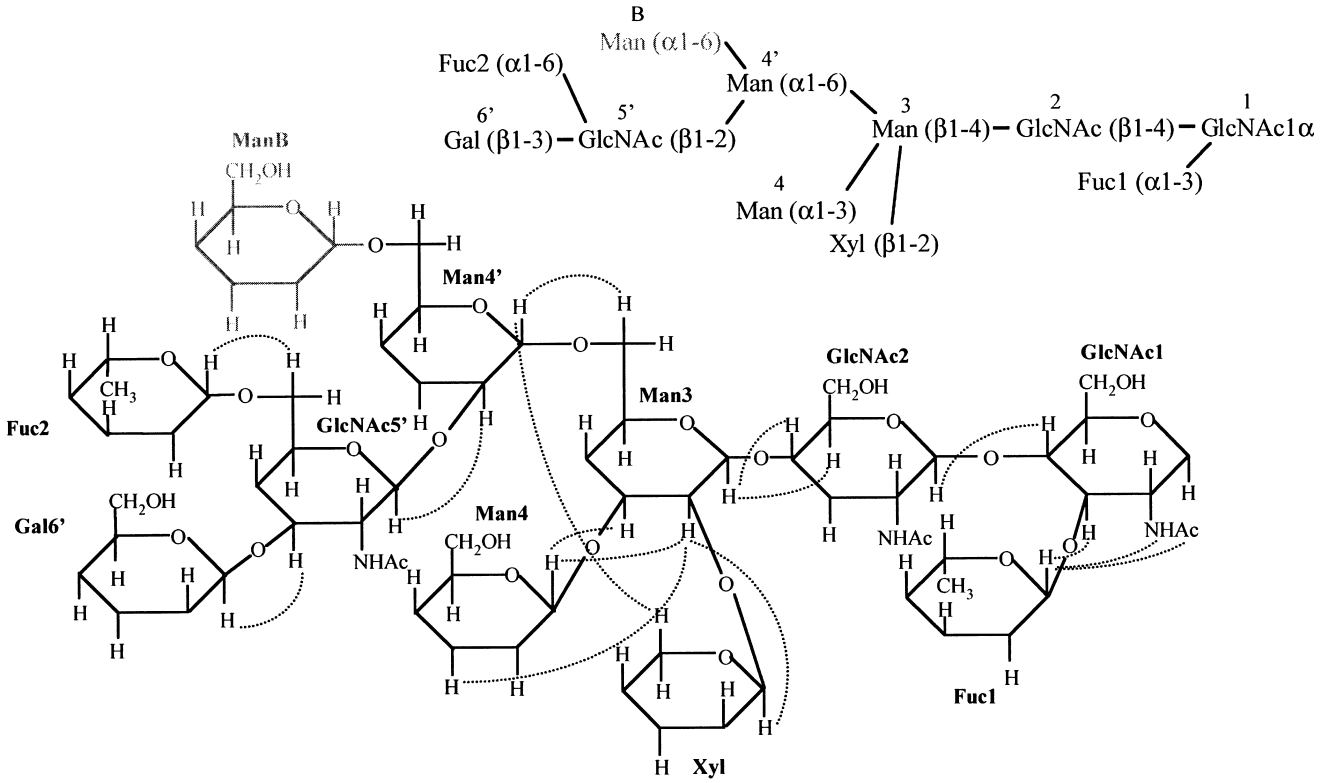

Fig. 4. Summary of the interactions observed in the glycan based on observed nOes and ^{1}H–^{13}C HMBC data. The figure shows only contacts identified from unambiguous nOes. Many more than this might be expected but could not be resolved due to overlap of H2, H3, H4, H5 and H6 resonances. The schematic arrangement of the glycan is indicated in the upper right corner. In both cases sugars where unambiguous linkages could not be identified from nOe data (ManB) are shown in a lighter shade. A linkage involving Gal6′ is shown although it was apparent from NMR integration that this represented about 20–25% of the total species.

those reported in the literature for core xylose protons (Lommerse et al., 1995; Agrawal, 1992; Takahashi et al., 1986). This sugar also yielded the most complete TOCSY spectrum, with intense correlations from H1 through to H6 over a relatively broad chemical shift range. The last unique monosaccharide, Gal, was identified on the basis of chemical shift comparison with literature values, for example, the position of H1 (4.49 ppm) was nearly identical to several values reported in the literature (Agrawal, 1992; van Rooijen et al., 1998; Takahashi et al., 1986; Tomiya et al., 1988), as was that of C1 (105.57 ppm) for a β-linked galactose ring (Agrawal, 1992).

The TOCSY and the HMQC spectra indicated near degenerate resonances for H1 of Fuc1 and Man4. Confirmation of the assignment for Man4 was made using spectra collected at 45°C where its anomeric proton was shifted 0.02 ppm downfield from that of H1 in Fucl. In addition, the carbon chemical shift for Man4 was very similar to that previously observed for Man4 in an *N*-linked glycan from CD2 (Wyss et al., 1995).

With six sugar spin systems identified (1 GlcNAc, 1 Man, 2 Fuc, 1 Xyl, 1 Gal), the remaining correlations in the HMQC and the TOCSY spectra were assessed. In general the remaining correlations for up to six anomeric carbons had chemical shifts in a tight range of 99–103 ppm. These sugars were clustered into two distinct groups based on ^{1}H chemical shift, one group had anomeric proton resonances between 4.8–4.9 ppm, consistent with other reports for high Man containing glycans (Wyss et al., 1995). Analysis of the TOCSY spectra for these sugars revealed a pattern nearly identical to that of the high mannose glycan from human CD2 (Wyss et al., 1995). In particular, Man3 and Man4′ exhibited correlations only between H1 and H2 at 25°C (Fig. 2). At 45°C, a further correlation from H1 to H3 was noted for Man4′ but not for Man3. However ManB, named based on the linkage pattern and similarity to CD2, showed correlations from H1 to H2, H3 and H4.

The last remaining glycan resonances observed were located 4.55 ppm (^{1}H) and 103 ppm (^{13}C) in the HMQC spectrum. These resonances clearly resulted from 2 GlcNAc residues (GlcNAc2, GlcNAc5′) based on examination of the 1D NMR spectrum, where the sugars gave rise to an apparent triplet representing two overlapping doublet patterns. The key to distinguishing these two sugars was the identification of H3 for GlcNAc5′ near 4.05 ppm which could be correlated with H2, H4 and H5 in the TOCSY spectrum (Fig. 2).

2.3. Identification of sugar linkages in the glycan

Sequential assignment and linkages of the sugars in the glycan was done using the initiating GlcNAc, 2 fucose, 1 xylose, 1 galactose and 1 mannose residues as starting points and identifying ^{1}H–^{1}H interactions in the 300 msec NOESY spectrum (Fig. 3). Further information about the linkage was obtained from a ^{13}C-decoupled ^{1}H–^{13}C HMQC (Fig. 1) and a ^{1}H–^{13}C HMBC (data not shown) which yielded long range ^{1}H–^{13}C correlations (Agrawal, 1992). A coupled ^{1}H–^{13}C HMQC experiment was used to measure $^{1}J_{C1,\,H1}$ coupling constants. This experiment allowed determination of the anomer configuration, as the $^{1}J_{C1,H1}$ coupling constant is generally near 170 Hz for the α-anomer but closer to 160 Hz in the β-anomer (Bock, Lundt & Pedersen, 1973). A summary of the observed inter-saccharide connectivities is shown in Fig. 4.

The linkage for Fuc1 was established from the observation of nOes between H1 of Fuc1 and H3 of the initiating GlcNAc1 residue. The ^{1}H–^{13}C HMQC spectrum displayed a small ^{1}H–^{1}H splitting for Fuc1 indicating a α1–3 link between these two sugars. The measured $^{1}J_{C1,\,H1}$ coupling constant for Fuc1 (171.8 Hz) was consistent with the α-anomer. Further, nOes were observed between Fuc1 H1 and GlcNAc1 HN and *N*-acetyl. Also appended to the GlcNAc1 sugar was GlcNAc2, established from the nOes between H1 of GlcNAc2 and H4 of GlcNAc1. Observation of a $^{3}J_{1,\,2} = 8.2$ Hz and $^{1}J_{C1}$, H1 = 160.3 Hz indicated the β-anomer was present for GlcNAc2, and suggested a β1–4 linkage between these two sugars.

Analysis of the NOESY and ^{1}H–^{13}C HMBC spectra showed that Man3 was the key central branching sugar in Glc. Several nOes were observed between GlcNAc2 and Man3, including Man3 H1-GlcNAc2 H4 and H5. Further, an unambiguous correlation between Man 3 C1 and GlcNAc2 H4 was observed in the ^{1}H–^{13}C HMBC. NOes were also observed from Man3 H2 to Xyl H1 and from Man3 H3 and H2 (not visible at contour level shown in Fig. 3) to Man4 H1. These interactions were confirmed by observations of a Man3 C2–Xyl H1 and Man C3–Man4 H1 correlations in the ^{1}H–^{13}C HMBC spectrum. The ^{1}H–^{13}C HMQC spectrum (Fig. 1) was less diagnostic of the anomeric conformations for the mannose residues in the glycan since $^{3}J_{1,\,2}$ coupling constants for Man3 (2.1 Hz), Man4 (2.1 Hz) and Man4′ (1.7 Hz) could be consistent for either the α- or β-anomer. The anomeric conformations of these sugars was clear from the coupled ^{1}H–^{13}C HMQC spectrum which revealed $^{1}J_{C1}$, H1 = 159.0 Hz for Man3 and 170.6 Hz for Man4 indicative of the β and α anomers respectively. The large $^{3}J_{1,\,2}$ coupling constant for Xyl (7.5 Hz) and $^{1}J_{C1}$, H1 = 157.5 Hz confirmed a β-linkage involving this sugar. The conformation for Man4′ could not be unequivocally determined from the $^{1}J_{C1,\,H1}$ coupling, as peaks were not readily visible in the spectrum for this residue. However, the chemical shift of H1 and H2 were nearly identical to those reported for the α conformation in human CD2 (Wyss et al., 1995). Additionally, nOes were observed from Man4′ H1 to H2 but not to H3. This latter nOe would be expected for the β-anomeric configuration only. Together these observations point to Man3 as the key branching sugar, bearing the β-linked Xyl residue at O2, the α-linked Man4 residue at O3 and the α-linked Man4′ residue at O6. These observations were consistent with previous NMR studies for the plant glycans from laccase (Takahashi et al., 1986) and pineapple stem bromelain (Lommerse et al., 1995).

The remaining GlcNAc residue, GlcNAc5′, is linked to Man4′ by a β1-2 linkage. The β-linkage was obvious from the splitting pattern in the ^{1}H–^{13}C HMQC (Fig. 1), and an nOe observed from GlcNAc5′ H1 to Man4′ H2. The single galactose (Gal6′) was linked to GlcNAc5′ via a β1-3 linkage as was evident from a weak nOe between the Gal6′ H1 and GlcNAc5′ H3 protons. The β-anomer for Gal6′ was confirmed from $^{1}J_{C1}$, H1 = 153.9 Hz. Finally, a single nOe correlation was noted between Fuc2 H1 and GlcNAc5′ H6. Together with the observed small $^{3}J_{1,\,2}$ coupling observed for Fuc2 in Fig. 1 this indicated an α1–6 linkage. The chemical shift data for H1 of Fuc2 (5.00 ppm) agrees well with that reported by Takahashi and co-workers for a terminating fucose ring having a α1–6 linkage to GlcNAc (Takahashi et al., 1986).

The linkage involving ManB could not be resolved. However some interpretation was made based on a comparison to literature chemical shift data. ManB exhibited chemical shifts for H1, H2 and H3 (4.90, 3.97 and 3.85 ppm respectively) that were all within 0.03 ppm of a terminal ManB in the high mannose glycan huCD2 (Wyss et al., 1995). A one-bond $^{1}J_{C1,\,H1}$ coupling constant of 169.3 Hz was measured for this residue, consistent with the α-anomeric conformation. These observations tentatively indicate ManB is α-linked to Man4′ at O6.

2.4. Microheterogeneity of the glycan

During the NMR investigation of the glycan an inconsistency between the number of sugars identified by NMR methods and those predicted by sugar compositional analysis was identified. Specifically, the NMR analysis determined that only three GlcNAc residues were present in the glycan while the sugar analysis indicated four such sugars. To determine whether this difference was a result of either overlapping resonances or incomplete assignment of the NMR spectra, several one-dimensional spectra of the glycan

were collected. The anomeric regions between 4.4–5.2 ppm and the H2 resonance for Man3 at 4.26 ppm were integrated to determine the relative abundance of each sugar residue. This approach lead to resolved integrated ^{1}H regions for Man4, Fuc1, GlcNAc1, Fuc2, ManB, Man 4′, Xyl, Gal6′ and Man3 in ratios 1.2, 1.2, 0.8, 0.7, 0.7, 0.7, 0.9, 0.22 and 1.2, respectively. Further analysis of the methyl region for GlcNAc residues between 2.0 and 2.1 ppm revealed a relative integral of 3.4. These observations indicate that the glycan was comprised of 3 GlcNAc, 2 fucose, 4 mannose and 1 xylose sugars based on NMR integration. The observation that only 3 GlcNAc residues were present in the glycan is in contrast to the 4 GlcNAc residues suggested from compositional analysis. However the interpretation of 3 GlcNAc residues is consistent with two-dimensional NMR analyses where we were unable to locate a spin system or nOe crosspeaks from a fourth GlcNAc. From the integration of the NMR spectra it was apparent that Gal6′ was found in a much lower proportion than other sugars having a relative ratio of only 0.22. This would indicate that Gal6′ is present in only a minor amount (20–25%) of the glycan studied.

3. Discussion

Determinations of the primary structures of plant glycoproteins are made difficult by the low abundance of the species and their heterogeneity. As a result, relatively few plant glycans have been examined by ^{1}H-NMR spectroscopy. The few that have been studied, bromelain (Lommerse et al., 1995) and laccase (Takahashi et al., 1986), have had partial NMR assignment done but have fallen short of complete analysis. Thus, the 11-residue glycan from this work, Glc, represents one of the largest and most complex plant glycans to be analyzed by NMR spectroscopy.

The primary structure of the glycan determined by NMR spectroscopy shows a core structure that is similar to many other complex carbohydrate chains, from both plant (Lommerse et al., 1995; Takahashi et al., 1986) and mammalian (Wyss et al., 1995; van Rooijen et al., 1998) sources. It comprises the same Manα1-6(Xylβ1-2)Manβ1-4GlcNAcβ1-4(Fucα1-3)GlcNAcβ core structure found in the pineapple stem glycoprotein bromelain. The observation that only PNGase A is able to cleave the glycan from the protein supports the finding of a α1–3 fucose linked to the proximal *N*-acetyl glucosamine (Kurosaka et al., 1991). All 3 glycans of this peanut peroxidase likely have this linkage (van Huystee et al., 1992) since PNGase F treatment has little effect (Hu & van Huystee, 1989). However, this glycan exhibits a more complex arrangement than bromelain (Lommerse et al., 1995) containing a third chain stemming from Man3 via a Manα1–3 linkage. Similar patterns are found in the high mannose glycoprotein huCD2 (Wyss et al., 1995) and the sycamore cell glycoprotein, laccase (Takahashi et al., 1986) which also contains Fuc2α1-6,GlcNAc5′α1,2 linked to Man4′.

Primary structure analysis of glycoproteins can be complicated by micro-heterogeneity, especially near the terminal sugars. For example the hu CD2 glycoprotein has at least two different forms where two terminal mannose sugars are removed (Wyss et al., 1995). This observation in plants is likely the result of cell wall glycosidases, which have been shown to be active towards the glycan in peanut cells (Wan et al., 1994), tomato (Morvan & Lhernould, 1996) and strawberry (Faugeron, Lhernould, Lemoine, Costa & Morvan, 1997). However in this glycan, this heterogeneity does not seem to affect peroxidase activity. Further evidence for microheterogeneity arises from observations of a residual ^{1}H–^{13}C resonance indicative of the β-form of the initiating GlcNAc sugar. The magnitude of this resonance compared to the majority α-form indicates that the α-form is the predominant species in dynamic equilibrium with the β-anomer. The appearance of these two species after PNGase A cleavage is consistent with other reports (Vliegenthart, Dorland & van Halbeek, 1983). The major source of heterogeneity was derived from the intensity of the Gal6′ resonances, which appeared to be about 20–25% of other comparable resonances in 1D NMR spectra and in the ^{1}H–^{13}C HMQC. Gal6′ is a terminal sugar in this glycan and is perhaps most prone to glycosidase activity in the cell.

Many questions remain about the roles of glycans in peanut peroxidase and other plant glycoproteins. One suggestion is that the glycans serve to protect the protein (Olden, Parent & White, 1982; Dwek, 1996), but structural information to verify this proposal has been difficult to come by due to the lack of three dimensional structures of intact glycoproteins. For peanut peroxidase it is anticipated that modelling of the NMR-derived primary structures to the three dimensional structure (Schuller et al., 1996) may reveal insight into the glycan role in enzyme activity.

4. Experimental

4.1. Glycopeptide isolation

Medium from peanut cells grown for 14 days was separated by vacuum filtration (Sun et al., 1997) and the enzyme separated from the medium by CM ion exchange, Con-A affinity, and gel filtration chromatography (Sesto & van Huystee, 1989). Enzyme purity was assayed by SDS–PAGE using from 5 to 40 μg peroxidase (Sesto & van Huystee, 1989). The glycopep-

tides were digested by TPCK (1-tosylamido-2-phenylethylchloromethyl ketone)-trypsin treatment (Wan & van Huystee, 1993) and then separated from peptides by filtration by Biogel P-6. Next, the three individual glycopeptides were separated by reversed phase HPLC equipped with a Beckman Gold system and a C-18 column (Ultrasphere 4.6 × 150 mm Beckman). This separation, as well as the sugar analysis on each glycan, was done as described (Sun et al., 1997).

4.2. Isolation of the glycan from Asn-144 peptide

Initial assays were done with pronase as reported (Wan & van Huystee, 1993). The PNGase F assays were carried out as suggested (Kurosaka et al., 1991). The digestion with PNGase A was carried out with 100 μl 1 mM glycopeptide (4 mg) and 150 μl 3 mU PNGase A (Cedarlane Lab.) at pH 5 in a McIlvaine buffer at 37°C for 48 h. The glycan was retrieved by passing the 250 μl sample again through the HPLC column as before (1 ml min^{-1} flow rate). The buffer salts were removed from the glycans by a small column (1 × 5 cm) of AG 50W-X8 and Dowex-1. The sample was washed through the column with 5 volumes of HPLC grade H_2O, then lyophilized before being resuspended in D_2O in preparation for ^1H-NMR.

4.3. NMR spectroscopy

Purified, freeze-dried glycan (1.2 mg) was dissolved in 1 ml D_2O and lyophilized three times to remove traces of H_2O. After the final lyophilization the sample was dissolved in 500 μl 99.9995% D_2O and the pH adjusted to 6.3 using a dilute solution of DCl in D_2O. This resulted in a final glycan concentration of ~1 mM which was used for most NMR experiments. For the TOCSY spectrum in H_2O, the above sample was lyophilized and redissolved in 90% H_2O/10% D_2O.

NMR spectra were acquired on a Varian Unity 500 spectrometer in the phase-sensitive mode. Experiments to assign sugar resonance were carried out at 25°C, although spectra were also recorded at 45°C to resolve overlapping peaks. In all experiments, water suppression was accomplished by a weak presaturation pulse of 2 s duration. Typically the series of experiments include ^1H DQF-COSY (Piantini, Sorensen & Ernst, 1982), TOCSY (Bax & Davis, 1985) and NOESY experiments and natural abundance ^1H–^{13}C HMQC and HMBC experiments. TOCSY experiments were collected using MLEV-17 spin lock mixing periods of 55 and 125 ms. NOESY spectra were collected by the hypercomplex method (States, Haberhorn & Rubin, 1982) using a 300 ms mixing time. Typical data sets included 32 or 64 transients for 256 complex t1 increments. Spectral widths of 4000 Hz in D_2O and 6000 Hz in 90% H_2O were used. For natural abundance ^1H–^{13}C HMQC and HMBC experiments data sets included 256 transients for 256 complex t1 increments. Spectral widths of 4000 Hz for ^1H and 10,000 Hz for ^{13}C were used keeping the ^1H transmitter centered at 4.73 ppm and ^{13}C centered at 74.2 ppm. 2D ^1H spectra were processed using a $\pi/8$ shifted sinebell in F1 and F2 with baseline correction in F2. For ^1H–^{13}C HMQC and HMBC, linear prediction was used to double the number of points in F1 prior to fourier transformation. Data was analyzed using Vnmr, NMRPipe (Delaglio, Greesiek, Vaister, Zhu, Pfeifer & Bax, 1995) and Pipp/Stapp (Garrett, Powers, Gronenborn & Clore, 1991).

Acknowledgements

We thank Frank Delaglio and Dan Garrett (NIH) for providing NMRPipe and Pipp for NMR processing and analysis. This research was supported in part by grants from the Natural Sciences and Engineering Research Council of Canada (RBvH) and Medical Research Council of Canada (GSS). Funding for the NMR spectrometer in the McLaughlin Macromolecular Structure Facility was provided by grants from the Medical Research Council of Canada, a generous gift from the R. Samuel McLaughlin Foundation, the London Life Insurance Co. of Canada, and an Academic Development Fund of the University of Western Ontario.

References

Agrawal, P. (1992). *Phytochemistry*, *31*, 3307.
Barber, K. R., Rodriguez Maranon, M. J., Shaw, G. S., & van Huystee, R. B. (1995). *Eur. J. Biochem*, *232*, 825.
Bax, A., & Davis, D. G. (1985). *J. Mag. Res*, *65*, 355.
Bock, K., Lundt, I., & Pedersen, C. (1973). *Tet. Lett*, *13*, 1037.
Delaglio, F., Greesiek, S., Vaister, G., Zhu, G., Pfeifer, J., & Bax, A. (1995). *J. Biomol. NMR*, *6*, 277.
Dwek, R. A. (1996). *Chem. Rev*, *96*, 683.
Faugeron, C., Lhernould, S., Lemoine, J., Costa, G., & Morvan, H. (1997). *Plant Physiol. Biochem*, *35*, 891.
Garrett, D. A., Powers, R., Gronenborn, A. M., & Clore, G. M. (1991). *J. Mag. Res*, *95*, 214.
Heese-Peck, A., & Raikhel, N. V. (1998). *The Plant Cell*, *10*, 599.
Hu, C., & van Huystee, R. B. (1988). Biochem. Biophys. Res. Comm., *156*, 500.
Hu, C., & van Huystee, R. B. (1989). *Biochem. J*, *263*, 129.
Hu, C., Carbonera, D., & van Huystee, R. B. (1987). *Plant Physiol*, *85*, 299.
Inagak, F., Shimada, I., Kohda, D., Suzuki, A., & Bax, A. (1989). *J. Mag. Res*, *81*, 186.
Kurosaka, A., Yano, A., Itoh, N., Kuroda, Y., Nakagawa, T., & Kawasaki, T. (1991). *J. Biol. Chem*, *266*, 4168.
Lerouge, P., Cabanes-Macheteau, M., Rayon, C., Fichette-laine, A.-C., Gomord, V., & Faye, L. (1998). *Plant Mol. Biol*, *38*, 31.

Lommerse, J. P. M., Kroon-Batenburg, L. M. J., Kamerling, J. P., & Vliegenthart, J. (1995). *Biochemistry*, *34*, 8196.
Montreuil, J. (1995). *New Compr. Biochem*, *29a*, 1.
Morvan, H., & Lhernould, S. (1996). *Plant Physiol. Biochem*, *34*, 335.
O'Donnell, J. P., Wan, L., & van Huystee, R. B. (1992). *Biochem. Cell Biol*, *70*, 166.
Olden, K., Parent, J. B., & White, S. L. (1982). *Biochem. Biophys. Acta*, *650*, 209.
Piantini, U., Sorensen, O. W., & Ernst, R. R. (1982). *J. Am. Chem. Soc*, *104*, 6800.
Ravi, K., Hu, C., Reddi, P. S., & van Huystee, R. B. (1986). *J. Exptl. B*, *37*, 1708.
Rodriguez Maranon, M. J., & van Huystee, R. B. (1994). *Phytochemistry*, *27*, 1217.
Schuller, D. J., Ban, N., van Huystee, R. B., McPherson, A., & Poulos, T. L. (1996). *Structure*, *4*, 311.
Sesto, P. A., & van Huystee, R. B. (1989). *Plant Sci*, *61*, 163.
States, D. J., Haberhorn, R. A., & Rubin, D. J. (1982). *J. Mag. Res*, *48*, 286.
Sturm, A. (1991). *Eur. J. Biochem*, *199*, 169.
Sturm, A. (1995). *New Compr. Biochem*, *29*, 521.
Sun, Y., Lige, B., & van Huystee, R. B. (1997). *J. Agric. Food Chem*, *45*, 4196.
Takahashi, N., Hotta, T., Ishihara, H., Mori, M., Tejima, S., Bligny, R., Akazawa, T., Endo, S., & Arata, Y. (1986). *Biochemistry*, *25*, 388.
Tarentino, A. L., & Plummer, T. H. (1994). *Methods in Enzymology*, *230*, 44.
Tomiya, N., Awaya, J., Kurono, M., Endo, S., Arata, Y., & Takahashi, N. (1998). *Anal. Biochem.*, *171*, 73.
van Huystee, R. B., & Wan, L. (1994). *C. R. Acad. Sci. Paris*, *317*, 789.
van Huystee, R. B., Sesto, P. A., & O'Donnell, J. P. (1992). *Plant Physiol. Biochem*, *30*, 147.
van Rooijen, J. J. M., Kamerling, J. P., & Vliegenthart, J. F. G. (1998). *Eur. J. Biochem*, *256*, 471.
Vliegenthart, J. F. G., Dorland, L., & van Halbeek, H. (1983). *Adv. Carbohydr. Chem. Biochem*, *41*, 209.
Wan, L., & van Huystee, R. B. (1993). *Biochem. Brophys. Res. Comm*, *194*, 1398.
Wan, L., Gijzen, M., & van Huystee, R. B. (1994). *Biochem. Cell Biol*, *72*, 411.
Wyss, D. F., Choi, J. S., & Wagner, G. (1995). *Biochemistry*, *34*, 1622.