Eur. J. Biochem. *204*, 313–316 (1992)

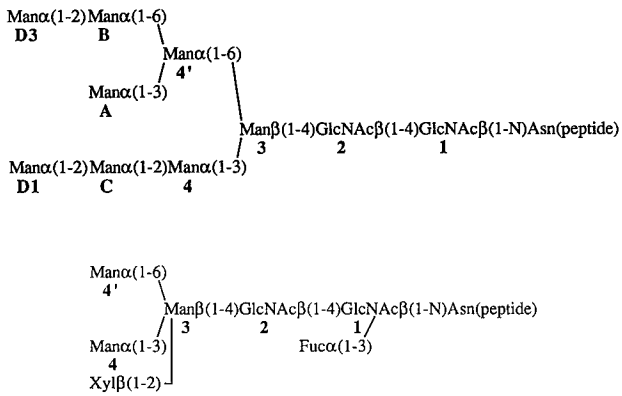
^{1}H-NMR structural determination of the N-linked carbohydrate chains on glycopeptides obtained from the bean lectin *phytohemagglutinin*

Arnd STURM[1], Aldert A. BERGWERFF[2] and Johannes F. G. VLIEGENTHART[2]

[1] Friedrich Miescher-Institute, Basel, Switzerland
[2] Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, The Netherlands

(Received September 5/November 11, 1991) – EJB 91 1186

Phytohemagglutinin, the lectin of the common bean *Phaseolus vulgaris*, is a N-linked glycoprotein with one high-mannose-type and one xylose-containing oligosaccharide side chain per polypeptide. The high-mannose-type glycan is attached to Asn12 and the complex-type glycan to Asn60 [Sturm, A. & Chrispeels, M. J. (1986) *Plant Physiol. 81*, 320–322]. The structures of the oligosaccharides were elucidated from two glycopeptides obtained from the lectin by Pronase digestion, affinity chromatography on concanavalin-A–Sepharose and gel-filtration chromatography on a column of BioGel P-4. The N-linked glycan structures were investigated by 500-MHz ^{1}H-NMR spectroscopy and were established to be:



Phytohemagglutinin (PHA), the lectin of the common bean *P. vulgaris*, accumulates in protein bodies of mature bean cotyledons and comprises about 2–4% of the total seed protein. Biosynthesis, post translational modifications and transport of PHA have been studied in considerable detail (see for example Chrispeels, 1983; Vitale et al., 1984; Vitale and Chrispeels, 1984). A possible role of PHA in plant defense was proposed by Chrispeels and Raikhel, 1991.

PHA isolated from mature seeds consists of two subunits with relative molecular masses of 34 kDa and 32 kDa named PHA-E and PHA-L, respectively. The PHA-E and PHA-L polypeptides are encoded by two tandemly linked genes, *dlec*1 and *dlec*2 (Hoffman and Donaldson, 1985). These genes, which are 90% similar at the nucleotide level, encode proteins that have 82% identity in their amino acid sequences. The derived amino acid sequences show the presence of three glycosylation sites (Asn-Xaa-Ser/Thr) in PHA-E (Asn12, Asn60 and Asn80) and two glycosylation sites in PHA-L (Asn12 and Asn60). The positions and sequences of the glycosylation sites at Asn12 and Asn60 are conserved in both polypeptides and only those sites are used. A high-mannose glycan was found attached to Asn12 and a complex glycan with xylose and fucose to Asn60 (Sturm and Chrispeels, 1986).

In the present paper, we report on the characterization of the N-linked glycans of the *P. vulgaris* lectin by 500-MHz ^{1}H-NMR spectroscopy.

Correspondence to A. Sturm, Friedrich Miescher-Institute, P.O. Box 2543, CH-4002 Basel, Switzerland

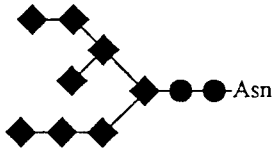
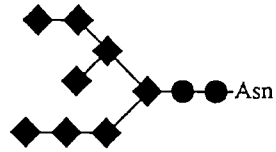
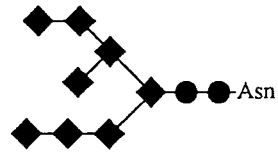
Abbreviations. ConA, concanavalin A; 2D, two dimensional; Fuc, L-fucose; PHA, phytohemaglutinin.

MATERIALS AND METHODS

Purification of PHA and isolation of glycopeptides

Seeds of *P. vulgaris* L. cv Greensleeves were obtained from plants grown in a greenhouse in San Diego, California, USA. PHA was isolated by affinity chromatography using the procedure of Feldsted et al., 1975. Generation of glycopeptides by Pronase digestion, purification of glycopeptides by affinity chromatography on concanavalin-A–Sepharose and gel-filtration chromatography on a column of BioGel P-4 were carried out as described by Sturm et al., 1987a.

Table 1. ^{1}H-NMR chemical-shift data of structural reporter-group protons of the constituent monosaccharides for glyco-asparagines derived from bean lectin phytohemagglutin (ConA+), reference 7S soybean glycoprotein (Neeser et al., 1985) and yeast external invertase (Trimble and Atkinson, 1986). Compounds are represented by the following short-hand symbolic notation: ◆, D-Man; ●, D-GlcNAc. n.d., not determined.

Reporter group	Residue	Chemical shift in		
		phytohemagglutin 	7S soybean glycoprotein 	yeast invertase 
		ppm		
H-1	GlcNAc-**1**	5.086	5.069	–
	GlcNAc-**2**	4.601	4.600	5.244[a]
	Man-**3**	n.d.[b]	n.d.[b]	n.d.[b]
	Man-**4**	5.338	5.336	5.340
	Man-**4′**	4.868	4.866	4.872
	Man-**A**	5.087	5.085	5.077α 5.104[b]
	Man-**B**	5.145	5.142	5.144
	Man-**C**	5.304	5.302	5.303
	Man-**D1**	5.042	5.041	5.044
	Man-**D3**	5.042	5.051	5.044
H-2	Man-**3**	4.232	n.d.[b]	4.236
	Man-**4**	4.088	n.d.[b]	4.091
	Man-**4′**	4.145	n.d.[b]	4.144
	Man-**A**	4.065	n.d.[b]	4.068
	Man-**B**	4.026	n.d.[b]	4.027
	Man-**C**	4.106	n.d.[b]	4.105
	Man-**D1**	4.074[c]	n.d.[b]	4.068
	Man-**D3**	4.065[c]	n.d.[b]	4.068
Nac	GlcNAc-**1**	2.006	2.011	–
	GlcNAc-**2**	2.061	2.061	2.043

[a] α configuration of GlcNAc-**2**.
[b] β configuration of GlcNAc-**1**.
[c] Values may be interchanged.

Determination of oligosaccharide structures by ^{1}H-NMR

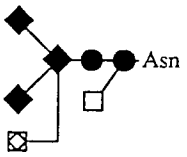
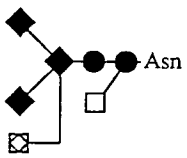
Prior to ^{1}H-NMR analysis, oligosaccharide samples were repeatedly exchanged in ^2H_2O (99.9% ^{2}H, MSD Isotopes) with intermediate lyophilization. The 500-MHz ^{1}H-NMR spectra were recorded on a Bruker AMX-500 spectrometer (Bijvoet Center, Department of NMR spectroscopy, Utrecht University) using a probe temperature of 27°C . Chemical shifts were measured relative to internal acetone (δ 2.225) in ^2H_2O, p^{2}H 7. The two-dimensional (2D) homonuclear Hartman Hahn spectrum (Bax and Davis, 1985; Marion et al., 1989) of fraction ConA+ (ConA, concanavalin A) was recorded on a Bruker AC-300 spectrometer (Department of Organic Chemistry, Utrecht University) operating at 300 MHz using a probe temperature of 27°C . For the 2D spectrum 474 experiments of 2048 datapoints representing a spectral width of 550 Hz were performed using MLEV-17 sequence of 100 ms for isotropic mixing, and a spin-lock field strength corresponding to a 90° pulse width of 28.4 μs. The carrier frequency was placed on 5534 Hz. The time-domain data were multiplied with a phase-shifted sine-bell function; phase-sensitive Fourier transformation was performed after zero filling to a 1024×4096 data matrix size. The 2D homonuclear Hartman Hahn spectrum of fraction ConA-was recorded on a Bruker AMX-500 spectrometer (Bijvoet Center, Department of NMR-spectroscopy, Utrecht University) operating at 500 MHz using a probe temperature of 27°C . For the 2D spectrum, 512 experiments of 2048 data points representing a spectral width of 3700 Hz, in both dimensions, was recorded using an MLEV-17 sequence of 100 ms for isotopic mixing, and a spin-lock field strength corresponding to a 90° pulse width of 26.5 μs. The carrier frequency was placed on 9030 Hz. The time-domain data were multiplied with a phase-shifted sine-bell function; phase-sensitive Fourier transformation was performed after zero filling to a 1024×4096 data matrix size.

RESULTS

Isolation of glycopeptides

PHA was isolated from mature seeds of the common bean *P. vulgaris* by affinity chromatography on thyroglobulin agarose. The purity of the lectin was confirmed by SDS/PAGE (data not shown). The purified protein was exhaustively digested with Pronase and glycopeptides carrying high-mannose-type glycans were separated from glycopeptides with xylose-containing glycans by affinity chromatography on concanavalin-A–Sepharose. The glycopeptides were further purified by gel-filtration chromatography on a calibrated column of BioGel P-4 (Sturm, 1991). The bulk of the

Table 2. ^{1}H-NMR chemical-shift data of structural reporter-group protons of the constituent monosaccharides for glyco-asparagines derived from bean lectin phytohemagglutin (ConA−) and reference seed lectin *S. japonica* (Capon et al., 1990). Compounds are represented by the following short-hand symbolic notation: ●, D-GlcNAc; ◆, D-Man; □, L-Fuc; ⊠, D-Xyl. n.d., not determined.

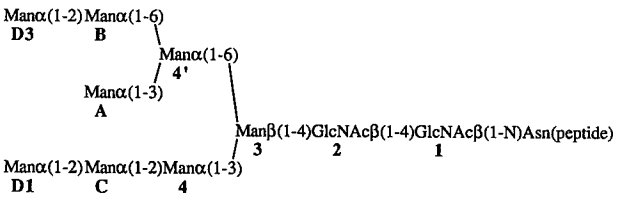
Residue	Reporter group	Chemical shift in	
		phytohemagglutin 	*S. japonica* seed lectin 
		ppm	
GlcNAc-**1**	H-**1**	5.075	5.042
	NAc	1.996	1.988
GlcNAc-**2**	H-**1**	4.566	4.561
	H-**4**	3.472	n.d.[a]
	NAc	2.050	2.050
Man-**3**	H-**1**	4.848	4.846
	H-**2**	4.266	4.263
Man-**4**	H-**1**	5.122	5.120
	H-**2**	4.038[a]	4.037
Man-**4′**	H-**1**	4.910	4.909
	H-**2**	3.973[a]	3.974
Fuc	H-**1**	5.128	5.128
	H-**5**	4.708	4.706
	CH_3	1.274	1.274
Xyl	H-**1**	4.463	4.462
	H-**2**	3.373	3.374
	H-**3**	3.449	3.447
	H-**5a**	3.257	3.255
	H-**5e**	4.014[a]	n.d.[a]

[a] Value determined from a 2D homonuclear Hartman Hahn experiment.

glycopeptides with high-mannose-type glycans eluted in one sharp peak corresponding to a glycopeptide with eight mannose residues (PHA/ConA+). Minor peaks (altogether less than 10% of the main peak) were observed corresponding to glycans with nine, seven or six mannose residues. The glycopeptide fraction with the xylose-containing glycan eluted in one sharp peak from the gel-filtration column (PHA/ConA-).

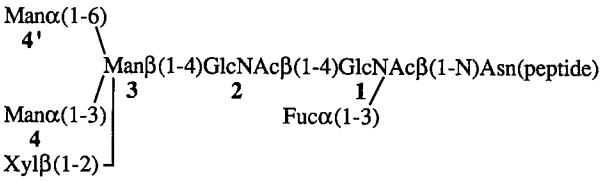
^{1}H-NMR structure determination

The ^{1}H-NMR spectrum of fraction PHA/ConA+ shows a homogeneous compound, namely:



The structural reporter-group signals coincide with those of $Man_8GlcNAc_2$-Asn derived from 7S soybean glycoprotein (Neeser et al., 1985) and yeast invertase (Trimble and Atkinson, 1986). Relevant ^{1}H-NMR parameters are given in Table 1. The H-2 signal of Man-**A** could be distinguished from the H-2 signals of Man-**D1** and Man-**D3** by a 2D homonuclear Hartman Hahn experiment (data not shown). On the H-1 track of Man-**A** a crosspeak with Man-**A** H-2 is found at δ 4.065. On the H-1 track of the coinciding H-1 atoms of Man-**D1** and **D3**, two partially overlapping crosspeaks are found at δ 4.065 and δ 4.074 with the H-2 atoms of these residues.

Close inspection of the ^{1}H-NMR spectrum reveals that the major component of fraction PHA/ConA− is:



The structural reporter-group signals (Table 2) are in good agreement with those of a major glycopeptide obtained from *Sophora japonica* seed lectin (Fournet et al., 1987; Capon et al., 1990). The xylose residue is recognized by a set of structural reporter-group signals, namely H-1, H-2, H-3 and H5a at δ 4.463, δ 3.373, δ 3.449 and δ 3.257, respectively (Van Kuik et al., 1985; Fournet et al., 1987).

DISCUSSION

In this study it has been demonstrated that the xylose-containing carbohydrate chain of the *P. vulgaris* lectin PHA

is a β(1-2)-xylosylated α(1-3)-fucosylated mannotriosyl *N,N'*-diacetylchitobiose glycan. This glycan structure has also been found attached to the lectins from *Clerodendron trichotomum* (Kitagaki-Ogawa et al., 1986), *Erythrina cristagalli* (Ashford et al., 1987), *S. japonica* (Fournet et al., 1987), *Ricinus communis* (Kimura et al., 1988), and *Artocarpus intergrifolia* (Capon et al., 1990), and to the protease inhibitor from *Caesalpinia pulcherrima* (Hase et al., 1986). Related xylose-containing glycan structures were found attached to the protease bromelain from *Ananas sativus* (Van Kuik et al., 1986), to the storage protein phaseolin from *P. vulgaris* (Sturm et al., 1987a) and to ascorbate oxidase from *Cucurbita pepo medullosa* (D'Andrea et al., 1988). PHA is unique among these proteins because it carries in addition to the xylose-containing glycan, a high-mannose glycan which has eight mannose residues and is similar to those of animal and yeast glycoproteins (reviewed by Kornfeld and Kornfeld, 1985), and to the high-mannose glycans already isolated and characterized from other plant glycoproteins (Dorland et al., 1981; Neeser et al., 1985; Sturm et al., 1987a).

Biosynthesis of the high-mannose and the xylose-containing glycan from PHA starts in the endoplasmic reticulum, where two high-mannose chains with nine mannose residues were found attached to Asn12 and Asn60 of the PHA polypeptides (Faye et al., 1986). From the endoplasmic reticulum, PHA is transported to the Golgi apparatus, where modification of the glycan at Asn60 takes place (Vitale and Chrispeels, 1984). Glycan modification includes the removal of mannose residues and the addition of *N*-acetylglucosamine, xylose and fucose. Glycosidases and glycosyltransferases required for PHA glycan processing act sequentially in a highly ordered fashion (Johnson and Chrispeels, 1987) and were localized in the Golgi apparatus (Sturm et al., 1987b). The modification of only one oligosaccharide side chain of PHA seems to be controlled by its accessibility to the Golgi-localized processing enzymes. The second glycan attached to Asn12 is shielded by the folded polypeptide and remains in its high-mannose form (Faye et al., 1986). Modified PHA is transported in vesicles from the Golgi apparatus to protein bodies (Chrispeels, 1983), where two terminal *N*-acetylglucosamine residues are removed (Vitale and Chrispeels, 1984) to give rise to the mature complex glycan structure (PHA/ConA−). The protein body is also the compartment in which the trimming of the high-mannose chain by one mannose residue occurs (Vitale et al., 1984; Sturm et al., 1987a).

The glycan structures of PHA seem to have no function in protein targeting, stability or protein oligomerization. PHA expressed in transgenic tobacco was faithfully glycosylated and transported into protein bodies and the mature protein isolated from tobacco seeds was indistinguishable from the lectin isolated from bean cotyledons (Sturm et al., 1988). PHA, from which the glycosylation sites were removed by site-directed mutagenesis, still accumulated in protein bodies of transgenic tobacco seeds although transport seemed to be less efficient than the transport of the glycosylated lectin (Voelker et al., 1989). The requirement of N-glycosylation for efficient intracellular protein transport was also shown for the bean storage protein phaseolin (Bustos et al., 1991) and yeast carboxypeptidase Y (Winther et al., 1991). N-glycosylation of PHA may also play a role in the dense packaging of the lectin into protein bodies. Such a function may be presumed because most proteins with similar complex glycan structures accumulate in vacuoles or in vacuole-derived protein bodies (Sturm, 1991).

Seeds of *P. vulgaris* L. cv. Greensleeves were a gift of Maarten J. Chrispeels of the University of California at San Diego, San Diego, California, USA. We thank Liliane Sticher and Pat King for critical reading of the manuscript.

REFERENCES

Ashford, D., Dwek, R. A., Welply, J. K., Amatayakul, S., Homans, S. W., Lis, H., Taylor, G. H., Sharon, N. & Rademacher, T. W. (1987) *Eur. J. Biochem. 166*, 311–320.

Bax, A. & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.

Bustos, M. M., Kalkan, F. A., Van den Bosch, K. A. & Hall, T. C. (1991) *Plant Mol. Biol. 16*, 381–395.

Capon, C., Piller, F., Wieruszeski, J. M., Leroy, Y. & Fournet, B. (1990) *Carbohydr. Res. 199*, 121–127.

Chrispeels, M. J. (1983) *Planta 158*, 140–151.

Chrispeels, M. J. & Raikhel, N. V. (1991) *Plant Cell 3*, 1–9.

D'Andrea, G., Bouwstra, J. B., Kamerling, J. P. & Vliegenthart, J. F. G. (1988) *Glycoconjugate J. 1*, 151–157.

Dorland, L., Van Halbeek, H., Vliegenthart, J. F. G., Lis, H. & Sharon, N. (1981) *J. Biol. Chem. 256*, 7708–7711.

Faye, L., Sturm, A., Bollini, R., Vitale, A. & Chrispeels, M. J. (1986) *Eur. J. Biochem. 158*, 655–661.

Feldsted, R. L., Laevitt, R. D. & Bachur, N. R. (1975) *Biochim. Biophys. Acta 405*, 72–81.

Fournet, B., Leroy, Y., Wiruszeski, J. M., Montreuil, J., Poretz, R. D. & Goldberg, R. (1987) *Eur. J. Biochem. 166*, 321–324.

Hase, S., Koyama, S., Daiyasu, H., Takemoto, H., Hara, S., Kobayashi, Y., Kyogoku, Y. & Ikenaka, T. (1986) *J. Biochem. 100*, 1–10.

Hoffman, L. M. & Donaldson, D. D. (1985) *EMBO J. 4*, 883–889.

Johnson, K. D. & Chrispeels, M. J. (1987) *Plant Physiol. 84*, 1301–1308.

Kimura, Y., Hase, S., Kobayashi, Y., Kyoguko, Y., Ikenaka, T. & Fanatsu, G. (1988) *J. Biochem. (Tokyo) 103*, 944–949.

Kitagaki-Ogawa, H., Matsumoto, I., Seno, N., Takahashi, N., Endo, S. & Arata, Y. (1986) *Eur. J. Biochem. 161*, 779–785.

Kornfeld, R. & Kornfeld, S. (1985) *Annu. Rev. Biochem. 54*, 631–664.

Marion, D., Ikura, M., Tschudin, R. & Bax, A. (1989) *J. Magn. Reson. 85*, 393–399.

Neeser, J. R., Del Vedovo, S., Mutsaer, J. H. G. M. & Vliegenthart, J. F. G. (1985) *Glycoconjugate J. 2*, 355–365.

Sturm, A. (1991) *Eur. J. Biochem. 199*, 169–179.

Sturm, A. & Chrispeels, M. J. (1986) *Plant Physiol. 81*, 320–322.

Sturm, A., Johnson, K. D., Szumilio, T., Elbein, A. D. & Chrispeels, M. J. (1987b) *Plant Physiol. 85*, 741–745.

Sturm, A., Van Kuik, J. A., Vliegenthart, J. F. G. & Chrispeels, M. J. (1987a) *J. Biol. Chem. 262*, 13392–13403.

Sturm, A., Voelker, T. A., Herman, E. M. & Chrispeels, M. J. (1988) *Planta 175*, 170–183.

Trimble, R. B. & Atkinson, P. H. (1986) *J. Biol. Chem. 261*, 9815–9824.

Van Kuik, J. A., Hoffmann, R. A., Mutsaers, J. H. G. M., Van Halbeek, H., Kamerling, J. P. & Vliegenthart, J. F. G. (1986) *Glycoconjugate J. 3*, 27–34.

Van Kuik, J. A., Van Halbeek, H., Kamerling, J. P. & Vliegenthart, J. F. G. (1985) *J. Biol. Chem. 260*, 13984–13988.

Vitale, A., Ceriotti, A., Bollini, R. & Chrispeels, M. J. (1984) *Eur. J. Biochem. 141*, 97–104.

Vitale, A. & Chrispeels, M. J. (1984) *J. Cell Biol. 99*, 133–140.

Voelker, T. A., Herman, E. M. & Chrispeels, M. J. (1989) *Plant Cell 1*, 95–104.

Winther, J. R., Stevens, T. H. & Kielland-Brandt, M. C. (1991) *Eur. J. Biochem. 197*, 681–689.