

Investigation of Cationic Peanut Peroxidase Glycans by Electrospray Ionization Mass Spectrometry

Cunjie Zhang; Amanda Doherty-Kirby; Robert van Huystee; Gilles Lajoie;

University of Western Ontario, London, Canada

Introduction: Cationic isozyme peanut peroxidase (CP) is *N*-glycosylated at three sites at Asn60 (GPb), Asn144 (GPc), and Asn185 (GPa). It has been shown to be a stable enzyme requiring heme¹ and Ca²⁺ for optimal activity². Enzymatic deglycosylation of CP with PNGase F results in lost of activity and increased susceptibility to trypsinolysis³. These observations indicate that the glycans of CP are functionally important. The glycan chains have never been fully characterized. In this study, we used a combination of off-line RP HPLC and nano ESI MS/MS to determine the structure of glycans of the tryptic glycopeptides of CP. This approach analyzes the glycan structures and determines the glycosylation site simultaneously with minimal sample preparation.

Methods: CP (500 µg) isolated from peanut cell suspension culture medium was further purified by reversed phase RP-HPLC on an analytical C18 column. The molecular mass of intact CP was obtained using ESI-MS. CP was reduced with DTT and then alkylated with iodoacetamide (IAA) in solution. The sample was filtered to remove free IAA and DTT. Cysteine-modified CP was digested with trypsin (1:40 w/w) for 2 h at 37°C. Tryptic peptides were separated by off-line RP-HPLC on C18 column⁴. Glycopeptide-containing fractions were identified by LC-MS/MS (Q-TOF Micro, Micromass, UK) operating in the precursor ion discovery (PID) mode. Each identified glycopeptide group (GPa, GPb and GPc) was further analyzed by ESI-MS and ESI-MS/MS (QTOF-2, Micromass) with borosilicate tips. Data were interpreted manually.

Results: ESI-MS of the intact protein reveals the microheterogeneity of the glycans (Fig. 1). Tryptic digestion of CP gave a near complete sequence coverage by ESI-MS. The glycopeptides from the tryptic digestion were separated by RP HPLC and identified by ESI-MS (Fig. 2). The structures of the glycan chains were determined by ESI-MS/MS (Fig. 3). The glycans are large structures of up to 16 sugars, but most their non-reducing ends have been modified giving a mixture of shorter chains at each site. Several attempts have been made to elucidate the structure (Table 1) of the glycans of CP with this study being the most complete to date. It is clear that there is a similar amount of heterogeneity at each of the three sites of glycosylation of CP and that the glycans of GPc are generally larger than those from GPb and Gpa (Table 1). However, the structures for each set of glycans share several common features: they have the same core *N*-linked glycan structure; fucose attaches to the GlcNAc residue linked to Asn; the majority contains fucose and xylose. By analyzing the masses of the intact proteins and glycoforms at each site, the major species present in the MS spectrum of intact CP can be explained by various combinations of observed glycans (Table 2). With this data, it is now possible to examine the role of glycans in CP stability and activity by treating the protein with exo or endoglycosidases and comparing glycan structure with stability, and function of the enzyme.

Conclusions:

The main glycan structures of CP have been elucidated. Heterogeneity of glycans occurs at each *N*-glycosylation site of CP. ESI-MS/MS analysis of glycan structures at glycopeptide level not only reduces the sample processing time, but also determines the localization of the glycan. This study demonstrates the power of mass spectrometry to resolve complex mixtures of post-translationally modified proteins. The characterization of these modifications will be critical for the detailed understanding of protein structure and function in plants and other organisms.

References:

1. Chibbar, R.N., Cella, R., van Huystee, R.B., 1984. *Can. J. Biochem. Cell Biol.* 62, 1046–1050.
2. Rodriguez Maranon, M.J., Stillman, M.J., van Huystee, R.B., 1993. *Biochem. Biophys. Res. Commun.* 194, 326–332.
3. Hu, C.F., van Huystee, R.B., 1989. *Biochem. J.* 263, 129–135.
4. Lige, B., Ma, S., van Huystee, R.B., 2001. *Arch. Biochem. Biophys.* 386, 17–24

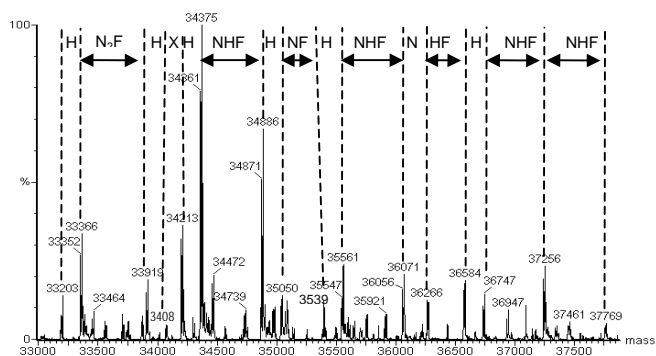


Fig. 1. MaxEnt1 processed ESI-MS data of CP. The masses between individual species correspond to various sugars (F=fucose, H=hexose, N=N-acetylglucosamine, X=xylose). Embedded is Base-line subtracted data.

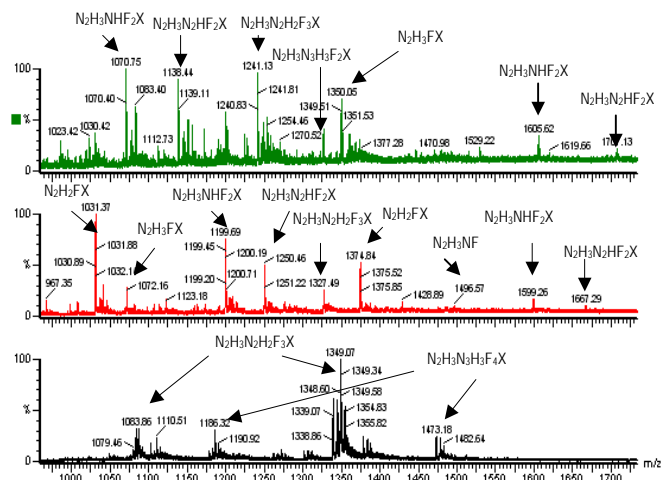


Fig. 2. ESI-MS spectrum of GPa (fraction 19), GPb (fraction 31 and 32) and GPC (fraction 45 and 46) with glycans labelled.

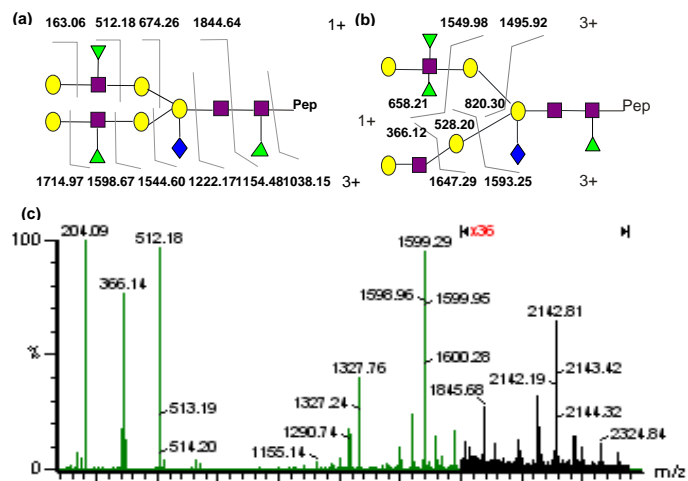


Fig. 3. Major deduced structures for the 1327 (4⁺) species of GPb (a,b) and the corresponding MS/MS spectrum. Hexose (H) ● Fucose (F) ▲ HexNAc (N) ■ Xylose (X) ◆

Table 2. Molecular masses of the intact protein forms from experimental MS data (Fig. 1) and calculated data

Glycoprotein Mass (Da)		GP a #	GP b #	GP c #
Observed Protein Mass (Figure 1) RI	Calculated Using Proposed Glycans (Table 1)			
33366**	33368	IV	III	
34082*	34083	IV	I	III
34213**	34215	III	III	III
34375***	34377	IV	III	III
34886***	34888	V III	III	III
35050*	35051	V III	IV	III
35398*	35400	V III	III	V II
35561**	35562	V III	V III	III
36071**	36073	V III(X)	X I(V III)	III
36276*	36277	V III(X)(X)	X I(X I)(X)	V III(III)
36584**	36585	V III	X I	V II
36747**	36747	V III	V III	X I
37256**	37258	X I(V III)	V III(V III)	X I(X III)
37769*	37770	X I	X I	X I

Table 1. Summary of proposed glycan structures for C

	Proposed glycan structure	Abbreviation	Glycan Mass Monoisotopic (Average)	Main species observed at m/z (charge states)		
				GP a	GP b	GP c
				m/z (CS) RI	m/z (CS) RI	m/z (CS) RI
I		N ₂ HF	714.27 (714.67)		958 (4+)** 967(4+, K)	
II		N ₂ H ₂ F	876.32 (876.82)		998 (4+)* 1007(4+, K)	
III		N ₂ H ₂ FX	1008.36 (1008.93)		1031 (4+)** 1374 (3+)	1390 (3+)** 1397 (3+, Na)
IV		N ₂ H ₃ FX	1170.41 (1171.08)	1350 (2+)**	1070 (4+)**	
V		N ₂ H ₂ NFX	1211.44 (1212.13)			1457(3+) 1465 (3+, Na)
VI		N ₂ H ₃ NFX	1373.49 (1374.27)	1451 (2+)* 1462 (2+, Na)	1122 (4+)* 1496 (3+)	1519 (3+, Na)*
VII		N ₂ H ₃ NF ₂ X	1519.54 (1520.41)			1568 (3+, Na)**
VIII		N ₂ H ₃ NHF ₂ X	1681.60 (1682.56)	1070 (3+)** 1605 (2+)	1199 (4+)** 1598 (3+)	
IX		N ₂ H ₃ N ₂ F ₂ X	1722.62 (1723.61)			1636 (3+, Na)**
X		N ₂ H ₃ N ₂ HF ₂ X	1884.68 (1885.75)	1139 (3+)** 1707 (2+)	1251 (4+)** 1666 (3+)	1689 (3+, Na)**
XI		N ₂ H ₃ N ₂ H ₂ F ₂ X	2192.79 (2194.04)	1241 (3+)**	1327 (4+)**	1339 (4+)** 1349 (4+, K) 1084 (5+, Na, K)
XII		N ₂ H ₃ N ₃ H ₂ F ₂ X	2411.86 (2413.23)	1326 (3+K)*		
XIII		N ₂ H ₃ N ₃ H ₃ F ₂ X	2703.97 (2705.52)			1472 (4+)** 1186 (5+, 3Na) 1963 (3+, Na)