Glycoconjugate Journal 17, 749–759, 2000

Structural studies of the glycopeptides of B-chain of cinnamomin – a type II ribosome-inactivating protein by nuclear magnetic resonance

Zheng Pu[1#], Yiming Li[2#], Fa-jian Hou[1], Fahu He[2], Naixia Zhang[2], Houming Wu[2*] and Wang-Yi Liu[1*]

[1]*Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yueyang Road, Shanghai 200031, China,* [2]*State Key Laboratory of Bioorganic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, 354 Fenglin Road, Shanghai 200032, China*

Cinnamomin is a plant type II ribosome-inactivating protein (RIP) isolated from the seeds of *Cinnamomum camphora*. It consists of two nonidentical polypeptide chains (A- and B-chain) held together through one disulfide linkage. Its A- and B-chain contain 0.3% and 3.9% sugars respectively. The B-chain of cinnamomin was digested by pronase E and then the liberated glycopeptides were separated from non-glycopeptides by gel filtration chromatography on a Bio-Gel P-4 column. Three crude glycopeptides were obtained by continuing chromatography over anion-exchange resin (AG1-X2) in the buffer of 2% pyridine-acetic acid (pH 8.3) with a polygradient elution system. Through further purification by the gel filtration chromatography and HPLC, three major glycopeptides, GP1, GP2 and GP3 were obtained. Mainly by two-dimensional Nuclear Magnetic Resonance (NMR) including TOCSY, DQF-COSY, NOESY, HMQC and HMBC, their primary structures were analyzed as: Manα1,3Manα1,6(Manα1,3)(Xylβ1,2)Manβ1,4GlcNAcβ1,4GlcNAcβ1-(Gly-)Asn-Asn-Thr(GP1), Manα1,6(Manα1,3)(Xylβ1,2)Manβ1,4GlcNAcβ1,4(Fucα1,3)GlcNAcβ1-Asn-Ala-Thr(GP2),Manα1,6(Manα1,3)Manα1,6(Manα 1,2 Manα1,3)Manβ1,4GlcNAcβ1,4GlcNAcβ1-(Ala-)Asn-Gly-Thr(GP3).

Keywords: **cinnamomin, glycopeptides, NMR, primary structure, ribosome-inactivating protein**

Introduction

Plant ribosome-inactivating proteins (RIPs) are a group of toxic proteins. There are two type RIPs: type I RIP consists of a single peptide chain with the RNA N-glycosidase activity and type II RIP is composed of two chains, A- and B-chain. The A-chain of type II RIP displays RNA N-glycosidase activity, and the B-chain is a galactose-specific lectin. Most RIPs are glycoproteins and their sugar contents are around 1–7% [1]. The primary structures of sugar chains of several RIPs known so far indicate that a "core portion" containing two N-acetyl-D-glucosamine residues and one mannose residue exists in these sugar chains in which one N-acetyl-D-glucosamine residue N-linked to the side chain of asparagine residue in protein chain. Most of the sugar chains are rich-mannose oligosaccharides and in some cases they contain one xylose and/or one fucose residue. The amino acid sequence around the glycosylation site of B-chain is Asn-X-Thr, where X represents any amino acid [2–4].

The primary structures of the sugar chains of several RIPs were determined by comparison of the elution positions of the sugar chains obtained by specific endoglycosidase digestion on the size-fractionation and reversed-phase HPLC with those of the standard sugar chains whose structures are known [2–5]. In this way, the carbohydrate composition of a sugar chain must be known before the sugar chain was digested by the specific endoglycosidases. Therefore, this method was time-consuming and inconvenient. Due to the development of the technology of Nuclear Magnetic Resonance (NMR), the determination of the primary structure of a new sugar chain by NMR became convenient [6]. However, the structure determination of sugar chain by NMR needs large amount of the sample and its purity must be high enough.

Cinnamomin is a new type II ribosome-inactivating protein purified in our laboratory from the seeds of *Cinnamomum Camphora*. It is a glycoprotein with the sugar contents of 0.3% and 3.9% in its A- and B-chain respectively [7]. Glycosylation frequently has a dramatic effect on protein structure, impacting

#Contributed equally to this work.
*To whom correspondence should be addressed: Wang-Yi Liu, Fax: 86-21-643 383 57; E-mail: liuwy@sunm.shcnc.ac.cn
Houming Wu, Fax: 86-21-641 661 28; E-mail: hmwu@pub.sioc.ac.cn

either overall structure or local conformation [8]. In some cases, the glycan chain can stabilize the structure of the fully folded protein by making specific, stabilizing contacts with the protein [9,10]. Although the physiological function of the sugar chains is unclear yet, sugar chains of cinnamomin were oxidized with periodate and then fluorescence-labeled with a fluorescent probe, both the RNA N-glycosidase activity of A-chain and the lectin activity of B-chain decreased three folds after the fluorescence-labeling [11]. In this paper, we report the purification of three major glycopeptides from cinnamomin B-chain by gel filtration chromatography, anion-exchange chromatography and HPLC and finally their primary structures were determined by NMR.

Experimental procedures

Materials

The A- and B-chain of cinnamomin were prepared according to the method as reported [12]. Polyacrylamide gel Bio-Gel P-4 and anion-exchange resin AG1-X2 were purchased from Bio-Rad Company. C_{18} reversed phase silica gel HPLC column was obtained from Alltech Company. Pronase E was the product of Serva Company. Other reagents were all of analytical grade.

Proteolysis of cinnamomin B-chain by Pronase E

Cinnamomin B-chain (1200 mg) was dissolved in 200 ml of the buffer A (0.1 M Tris-HCl, pH 7.8, 0.5% SDS) and then 100 mg of Pronase E was added. This reaction mixture was incubated at 37°C for 72 h, and during the period, another 200 mg of Pronase E were added at an interval of 24 h. Then the reaction mixture was boiled at 100°C for 20 min to stop the digestion and evaporated at 60°C for concentrating the mixture to 50 ml. The concentrated mixture was centrifuged (5000 g, 4°C) for 10 min to remove the insoluble materials appeared and the supernatant was stored at 4°C.

Separation of glycopeptides from nonglycopeptide by gel filtration chromatography

Five milligrams of the above supernatant was applied onto a Bio-Gel P-4 column (2 × 150 cm). The column was eluted with distilled water, each 5 ml of the eluate was collected and 0.35 ml of the sample was withdrawn for determining the sugar content by phenol-H_2SO_4 method [13]. From 50 ml of the applied sample, approximately 1000 ml of the sugar-containing fraction was collected, which was adjusted to pH 8.3 by 2% pyridine-acetic acid buffer.

Separation of the glycopeptides by anion-exchange chromatography

The acetate type of resin AG1-X2 was filled in a column (2 × 10 cm) and the column was then pre-equilibrated with 2% pyridine-acetic acid (pH 8.3). 100 ml of the glycopeptide-pyridine solution was loaded onto the column. After washing the column with 2% pyridine-acetic acid (pH 8.3), the glycopeptides were eluted with a polygradient elution system [14]. Each 5 ml of the eluate was collected and 0.35 ml of the sample was withdrawn to determine the sugar content by phenol-H_2SO_4 method. Three glycopeptides in peaks a, b and c were obtained and lyophilized.

Further purification of glycopeptides by gel filtration chromatography and HPLC

Three glycopeptides were dissolved respectively in 1 ml of re-distilled water and purified through Bio-Gel P-4 column (2 × 150 cm) using re-distilled water as the eluent. And then, three glycopeptides were further purified by reversed-phase HPLC on a C_{18} column (250 × 4.6 mm) using a gradient elution system composed of eluent A (0.1% TFA aqueous solution) and eluent B (40% CH_3CN and 0.1% TFA) at the flow rate of 1 ml/min. The peaks were monitored by UV absorption at 209 nm and the glycoside-containing peaks were detected by phenol-H_2SO_4 method (data not shown). Finally, three glycopeptides, GP1, GP2 and GP3, were obtained with high purity.

Glycopeptide analysis by Nuclear Magnetic Resonance spectroscopy

The glycopeptides of GP1, GP2 and GP3 were dissolved in D_2O or H_2O, respectively. All spectra (at 600 MHz) were performed on Varian Inova 600 at a probe temperature of 27°C. The measurements of two-dimensional spectroscopy such as DQF-COSY (Double quantum filter correlated spectroscopy), TOCSY (Total correlated spectroscopy) and NOESY (Nuclear Overhauser Effect spectroscopy) spectra

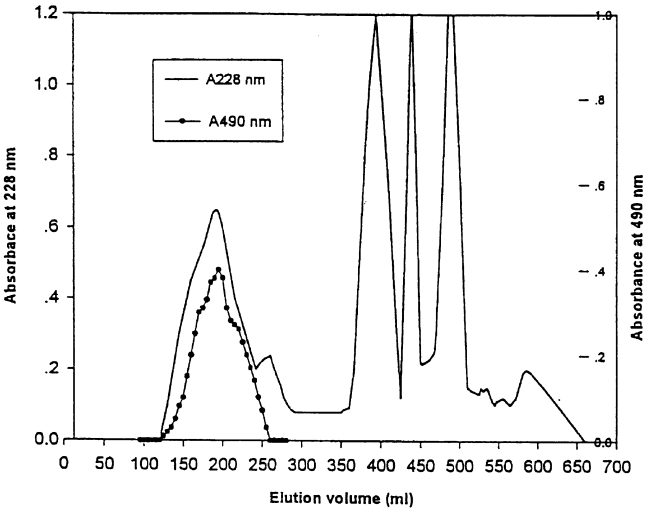

Figure 1. Separation of glycopeptides from non-glycopeptides by gel filtration of enzymatic digest of cinnamomin B-chain on Bio-Gel P-4 column. 5 ml of the concentrated digest of cinnamomin B-chain was applied by pronase E on the Bio-Gel P-4 column (2 × 150 cm) and eluted with distilled water. Each 5 ml fraction was collected and 0.35 ml of sample was analyzed by phenol-H_2SO_4 method (490 nm). Peptides were monitored at 228 nm.

were performed in phase sensitive mode using hyper-complex method. The quadrature detection was employed for both F2 and F1 dimension acquisition with the carrier frequency always maintained at the relaxation delay. The spectra were recorded with 1024 (for F2 dimension) $\times$ 512 (for F1 dimension) data points. The zero-filling was carried out in F1 dimension during data procession. A MLEV-17 composite pulse with the mixing time of 80 ms was used for spin lock in TOCSY experiment. NOESY experiments were performed with a mixing time of 300 ms. Inverse detected 2D-NMR experiments: HMQC (Heteronuclear Multi-quantum Coherence) and HMBC (Heteronuclear Multi-Bond Connection) were acquired using an inverse probe. A hmqc pulse sequence was used in HMQC and HMBC experiments. Mixing time 50 and 120 ms was used in HMBC experiments.

ESI-MS analysis of glycopeptides

The ESI-MS spectroscopy was recorded on a VG Quattro three stage mass spectrometer using methanol and water as matrix.

Results

Separation of glycopeptides from non-glycopeptides by gel filtration and anion-exchange chromatography

The exhaustive digestion of cinnamomin B-chain by Pronase E gave a set of small oligopeptides with similar size. Because glycopeptides contained sugar chains, their molecular size might be much greater than those of the non-glycopeptides. Size-fractionation chromatography was employed to separate the glycopeptides from non-glycopeptides with lower exclusion limit [15]. As shown in Figure 1, glycopeptides were separated effectively from non-glycopeptides by the gel filtration chromatography. Since different glycopeptides were composed of different amino acid and saccharide residues, their isoelectric points would be also different, implying that they could be separated by ion-exchange chromatography [2,16]. As indicated in Figure 2, three glycopeptides in peaks a, b and c were

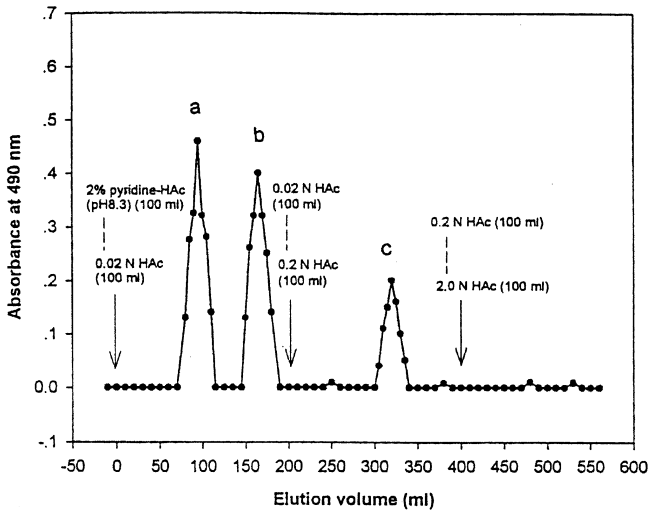

Figure 2. Fractionation of glycopeptides obtained from gel filtration of enzymatic digest of cinnamomin B-chain by AG1-X2 anion-exchange resin. 100 ml of the glycopeptides mixture in 2% pyridine-acetic acid buffer (pH 8.3) was loaded onto the AG1-X2 anion-exchange resin chromatography column (2 $\times$ 20 cm) pre-equilibrated with the same buffer (pH 8.3). The glycopeptides in peak a, b and c were obtained.

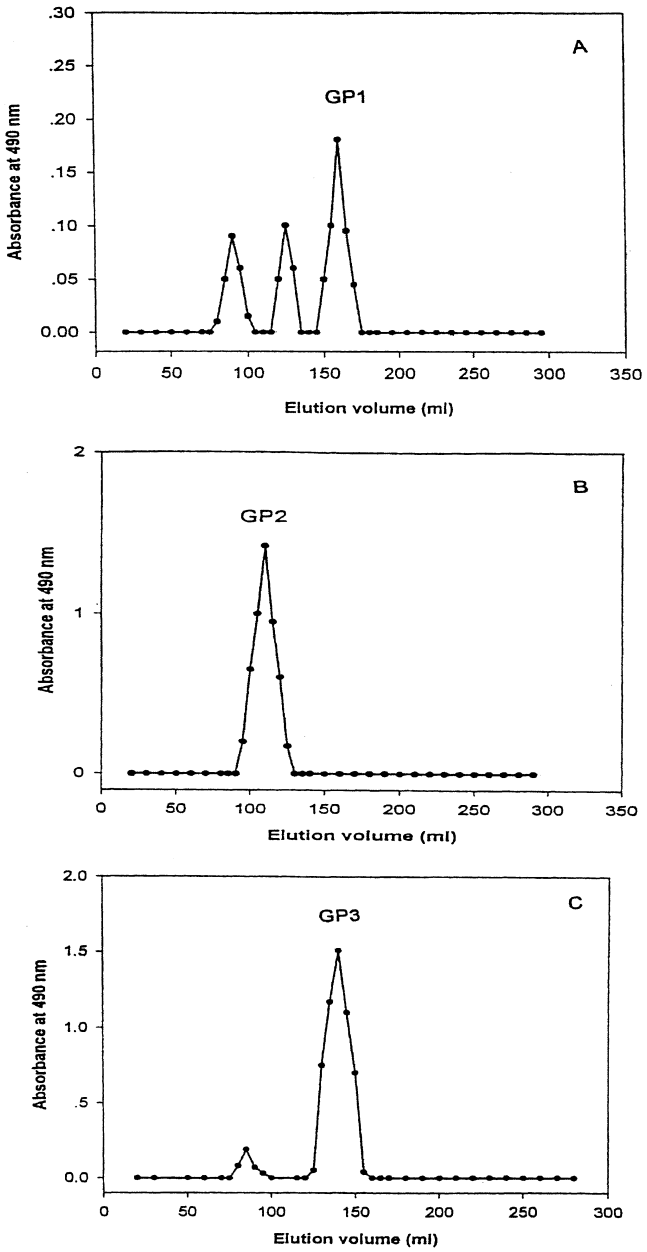

Figure 3. Further purification of the glycopeptide by Bio-Gel P-4 gel filtration chromatography column. Glycopeptide was dissolved in 1 ml of H_2O respectively and then loaded onto the Bio-Gel P-4 column. The column was eluted with H_2O. Each 5 ml of fraction was collected and 0.35 ml of each fraction was analyzed. (A), After further digestion of peak c, the glycopeptide GP1 was obtained and used for structural analysis; (B), Glycopeptide GP2 was from peak b; (C), Glycopeptide GP3 was from peak a.

separated by the anion-exchange chromatography. The peaks a, b and c contained approximately 15%, 40% and 45% of total sugar content of three peaks detected by phenol-H_2SO_4 method.

Further purification of the glycopeptides by anion-exchange chromatography and HPLC

The preliminary result examined by their 1D NMR spectra showed that three glycopeptides were not pure. In addition, the large volume of the glycopeptide eluate from anion-exchange resin (AG1-X2) chromatography made trace contamination from the pyridine-acetic acid buffer. Thus the glycopeptides were further separated from the impurity by size-fractionation chromatography of polyacrylamide gel (Figure 3) and HPLC. Three major glycopeptides, GP1, GP2 and GP3, were obtained with high purity for the NMR experiments.

Primary structure of GP1

The glycan chain of GP1 was extensively investigated by one and two-dimensional NMR in D_2O: DQF-COSY and TOCSY spectra were used for the spin system analysis of sugar residues, HMQC spectrum was recorded for the carbon resonance assignments, NOESY experiment was performed for the detection of intra- and inter-residue NOE correlations, and HMBC spectrum was used for revealing the long range connectivities ($^3J_{C\text{-}H}$) between proton and carbon across the glycosidic linkages. Anomeric proton signals of sugar residues were easily found in ^{1}H NMR at δ 5.14, 5.14, 5.06, 4.92, 4.88, 4.62 and 4.46, respectively. In the TOCSY spectrum (Figure 4) and the DQF-COSY spectrum, these Anomeric protons showed strong cross-peaks with their relevant saccharide ring proton at second position. These spin systems indicated that there were seven saccharide residues denoted as residue-A, B, C, D, E, F and G, respectively. In the TOCSY spectrum, H-1 of the residue-C, F and G showed complete set of correlation peaks throughout ring protons H-2, H-3, H-4 and H-5. This fact plus all big coupling constants ($>$7 Hz) among each pair of vicinal ring protons revealed that all these ring protons in residue-C, F and G were at axial bonds, which were confirmed by strong NOE cross-peaks between H_{C1}–H_{C3}, H_{C5}, H_{G1}–H_{G3}, H_{G5} and H_{F1}–H_{F3}, H_{F5} in the NOESY spectrum. Thus every proton in these residues was unambiguously assigned chemical shift. From these definite protons, the chemical

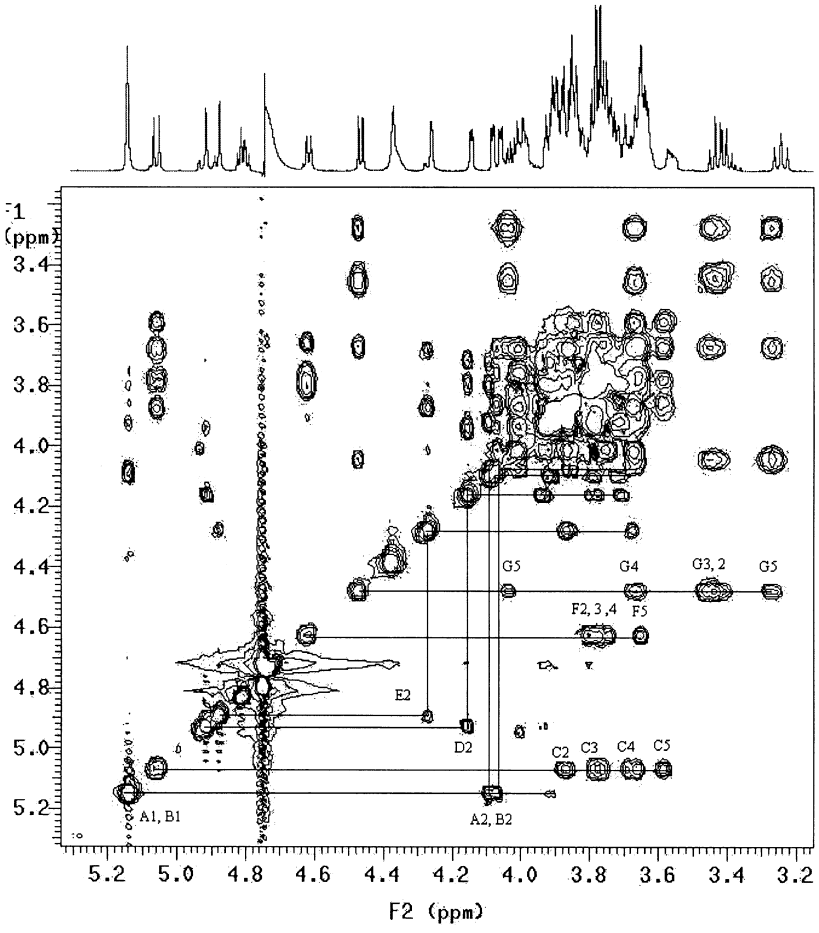

Figure 4. Carbohydrate part of the TOCSY spectrum of GP1. The spectrum was recorded in D_2O at 27°C with mixing time 80 ms. The proton spin systems for sugar residues were labeled by lines in the picture.

shifts of their relative carbons in residue-C, F and G were assigned by the HMQC spectrum furthermore. The high field chemical shift data of C_{C2} and C_{F2} at δ 53.80 and 55.18 suggested that residue-C and F were glucosamines. Therefore, on the evidences of spin system patterns, the NOE cross-peaks and the proton, as well as carbon chemical shifts, residue-C, F and G were attributed to β-D-glucosamine, β-D-glucosamine and β-D-xylose, respectively. In the TOCSY spectrum, the anomeric protons of remained residue-A, B, D and E had only cross-peaks to the relative H-2. From the H-2 of these residues, the cross-peaks and relay cross-peaks to their H-3, H-4 and H-5 could be observed. These facts plus the narrow line-shape of H-2 and wide line-shape of H-3, H-4 and H-5 revealed that H-2 must be at the equatorial bond and H-3, H-4 and H-5 at axial bonds of these four residues. Therefore, the residue-A, B, D and E were all mannoses. In the NOESY spectrum, H_{E1} showed the cross-peaks with H_{E3} and H_{E5}, indicating that H_{E1} was at axial bond and the configuration of residue-E was β-D-mannose. On the contrary, intra-residue NOE cross-peaks among H-1, H-3 and H-5 in residue-A, B and D were not observed, indicating residue-A, B and D were α-D-mannoses, which were further confirmed by the chemical shifts of C-5 of these residues.

The saccharide sequence of GP1 was established on the basis of inter-residue NOEs among the residues. In the NOESY spectrum (Figure 5), the anomeric H_{F1} had the cross-peak with H_{C4}, indicating that residue-F was connected to the fourth position of residue-C (β-D-glucosamine). Moreover, compared with the standard β-D-glucosamine [16], C_{C4} was downfielded 8.1 ppm, suggesting that the fourth position of residue-C was glycosylated. In the same way, NOEs between H_{A1} and H_{D3}, H_{B1} and H_{E3}, H_{D1} and H_{E6}, H_{G1} and H_{E2}, H_{E1} and H_{F2} were found respectively. At the same time, the glycosylated shifts of carbon C_{D3}, C_{E3} and C_{E6} to downfield were also found. Therefore, the sequence of saccharide residues in glycan chain of GP1 was determined.

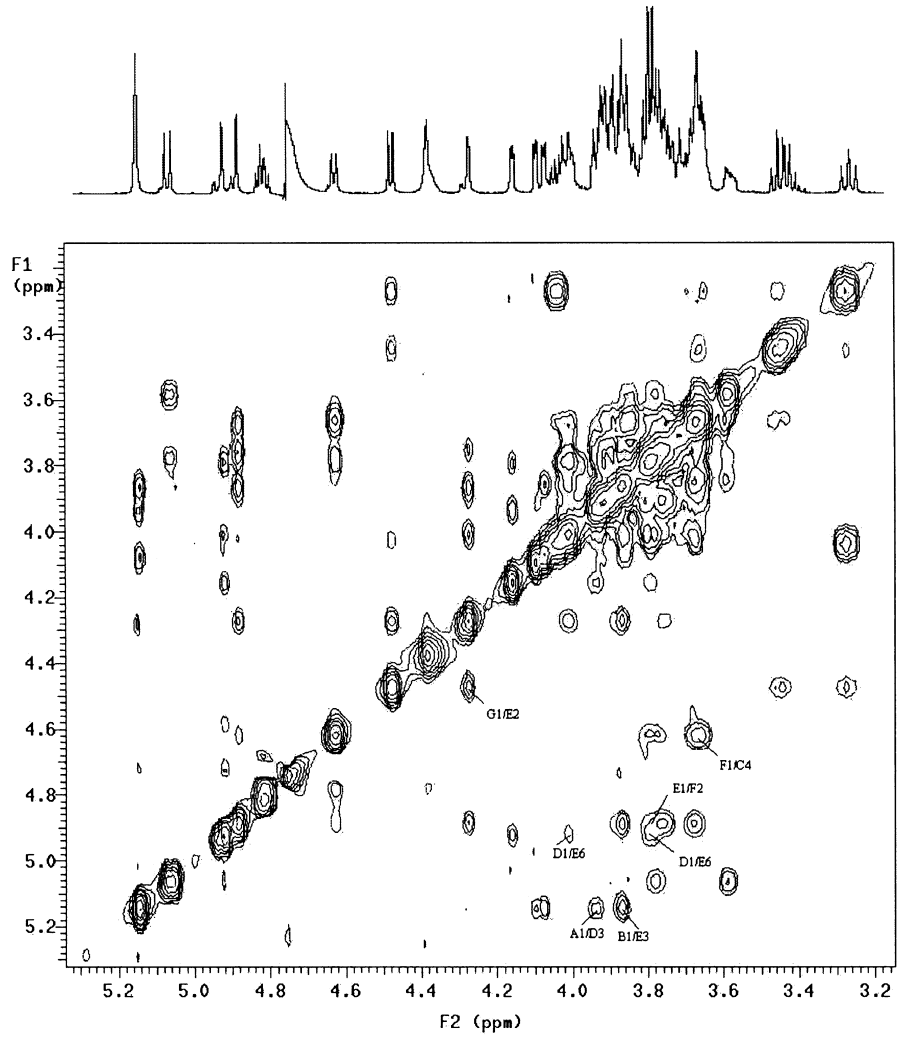

Figure 5. Carbohydrate part of the NOESY spectrum of GP1. The spectrum was recorded in D_2O at 27°C with mixing time 300 ms. The inter-residue NOE cross-peaks for the determination of sugar sequence were labeled in the picture.

The further evidences for the glycosidic linkages were obtained from the inter-residue long-range C−H correlation ($^3J_{C\text{-}H}$) either between the anomeric carbon and the aglycone proton or between the anomeric proton and the aglycone carbon in the HMBC spectra with two mixing times. Two acetyl groups linked to two glucosamine units were also assigned by the $^3J_{C\text{-}H}$ connectivity between the relevant H-2 of glucosamine residues and the carbonyl carbon of the acetyl groups.

The composition and sequence of amino acids in GP1 were established by NMR in H_2O using the similar way. The high field portion of the TOCSY spectrum showed proton spin systems as follows: one ABX_3 system at δ 4.37, 4.38 and 1.21 due to a threonine residue, two AMX-type spin systems at δ 4.81, 2.88 and 2.90 (each resonance corresponding two protons partly overlapped) assignable to two asparagine or aspartic acid residues, and one AB system at δ 3.86 and 3.87, which were correlated to a carbon at δ 40.53 in the HMQC spectrum, attributable to a glycine residue. At the lower field of the TOCSY spectrum (Figure 6), the spin systems arisen from amide protons were well agreement with the above assignments. The amide protons at δ 6.96 and 7.60 which were strongly correlated each other in DQF-COSY, showed the existence of an amino group of side chain of asparagine. In consideration of another amino group linked to the glycan, there were two asparagines in GP1. Mainly by the HMBC spectra which showed long range connectivity ($^2J_{C\text{-}H}$) between the amino proton in a residue and the acyl carbon of neighbor residue in Table 1, the sequence of the short peptide in GP1 was decided as Gly-Asn-Asn-Thr. No appearance of amide proton from glycine also demonstrated that glycine was at the N-terminus of the peptide. In the TOCSY spectrum, the side chain amide proton of Asn2 at δ 8.66 had cross-peak with H_{C1} at δ 5.06, indicating the glycan linked to Asn2, which was

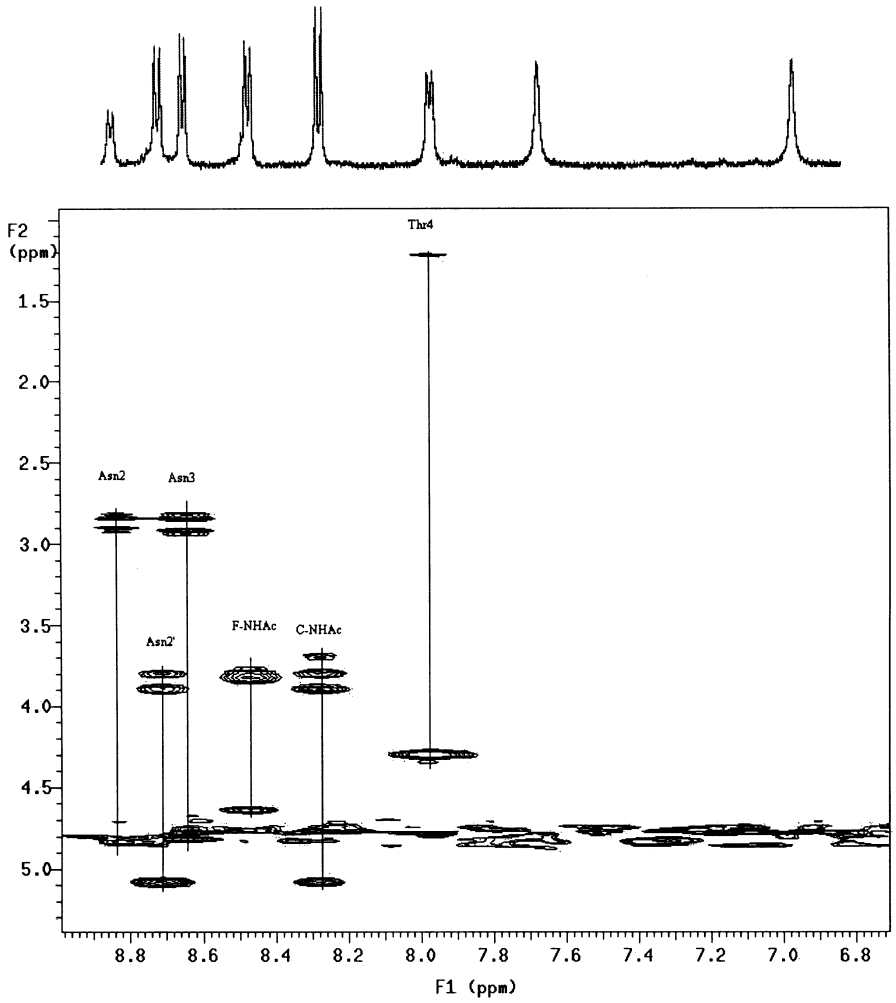

Figure 6. Amino acid part of the TOCSY spectrum of GP1. The spectrum was recorded in H_2O at 27°C with mixing time 80 ms. The Asn2, Asn3 and Thr4 indicated the spin systems from the amide protons of main chain of amino acid residues respectively. The Asn2′ indicated the spin system from the amino proton of side chain of Asn2. The F-NHAc and C-NHAc indicated the spin systems from the amide protons of glucosamine residues F and C.

Table 1. ^{1}H- and ^{13}C-chemical shifts of GP1[a]

Residue	*^{1}H*	*^{13}C*	*HMBC($^{3}J_{C\text{-}H}$)*
α-D-man (A)			
1	5.14	102.59	H_{D3}
2	4.09	70.33	
3	3.91	70.52	
4	3.79	66.00	
5	3.72	73.01	
6	3.92, 3.80	61.02	
α-D-man (B)			
1	5.14	102.59	
2	4.07	70.01	
3	3.85	70.52	
4	3.68	66.24	
5	3.76	72.23	
6	3.93, 3.75	61.32	
β-D-GluNAc (C)			
1	5.06(9.6)	78.43	
2	3.87	53.80	
3	3.77	72.30	
4	3.67	78.71	
5	3.58	76.31	
6	3.85, 3.67	60.00	
NHCO	8.22	174.72	
CH_3	2.04	22.12	
α-D-man (D)			
1	4.92(1.3)	99.73	H_{E6}
2	4.15	69.58	
3	3.93	78.80	H_{A1}
4	3.71	66.81	
5	3.79	73.19	
6	3.90, 3.80	60.91	
β-D-Man (E)			
1	4.88	100.89	H_{F4}
2	4.26	79.42	
3	3.86	77.68	H_{B1}
4	3.67	66.86	
5	3.68	74.52	
6	4.01, 3.79	65.58	
β-D-GluNAc (F)			
1	4.62(7.2)	101.51	H_{C4}
2	3.81	55.18	
3	3.78	72.92	
4	3.75	79.56	H_{F1}
5	3.65	74.63	
6	3.90, 3.76	60.19	
NHCO	8.41	174.55	
CH_3	2.08	22.23	
β-D-Xyl (G)			
1	4.46(7.2)	105.27	H_{E2}
2	3.42	73.48	
3	3.44	75.51	
4	3.66	69.26	
5	4.04, 3.27	65.01	
Gly1			
NH	–[b]		
α-CH	3.87	40.53	
C=O		167.11	NH-Asn2 ($^{2}J_{C\text{-}H}$)
Asn2			
NH	8.79		
α-CH	4.81	50.78	
β-CH	2.90, 2.88	36.57	
C=O		171.94	
CONH (side chain)	8.66	172.26	
Asn3			
NH	8.59		
α-CH	4.81	50.06	
β-CH	2.90, 2.88	36.06	
C=O		171.94	NH-Thr4 ($^{2}J_{C\text{-}H}$)
CONH (side chain)	7.62, 6.93	174.80	
Thr4			
NH	7.91		
α-CH	4.38	59.75	
β-CH	4.37	67.39	
γ-CH_3	1.21	18.99	
C=O		172.31	

[a] Water (δ_H 4.76) was used as the reference for proton chemical shifts and chloroform (δ_C 77.0) as the external reference for carbon chemical shifts.
[b] –, Not detected due to the exchange between proton and deuteron.

further confirmed by C_{C1} at higher field δ 78.43 comparing with the standard data (δ 95.9) of β-D-glucosamine.

Based on the above analysis, GP1 was identified as Manα1,3Manα1,6(Manα1,3)(Xylβ1,2)Manβ1,4GlcNAcβ1,4 GlcNAcβ1-(Gly-)Asn-Asn-Thr (Chart 1). The proton and carbon signals of amino acid residues were assigned as shown in Table 1.

Primary structure of GP2

^{1}H NMR spectrum of GP2 showed seven anomeric protons at δ 5.16, 5.15, 5.11, 4.94, 4.88, 4.59 and 4.49. Using the same procedures for analyzing the primary structure of GP1, the saccharide residues in GP2 were identified as two β-D-GlcNAc, two α-D-Man, one β-D-Xyl, one β-D-Man and one α-L-Fuc, and the amino acid residues were identified as one asparagine, one alanine and one threonine. The linkages among sugar residues and amino acids, as well as between the peptide and glycan, were solved by the NOESY and HMBC spectra. The primary structure of GP2 was deduced as Manα1,6(Manα1,3)(Xylβ1,2)Manβ1,4GlcNAcβ1,4(Fucα1,3)-GlcNAcβ1-Asn-Ala-Thr (Chart 1). The chemical shifts of protons and carbons were assigned as shown in Table 2.

Primary structure of GP3

^{1}H NMR spectrum of GP3 showed eight anomeric protons at δ 5.37, 5.12, 5.08, 5.07, 4.93, 4.89, 4.79 and 4.62, demonstrating that there were eight sugar residues in GP3. Using the same methods for analyzing GP1, the sugar residues of GP3 were identified by the TOCSY, DQF-COYS, NOESY, HMQC and HMBC spectra. Besides two β-D-GlcNAc and one β-D-mannose which were also found in GP1 and GP2, five

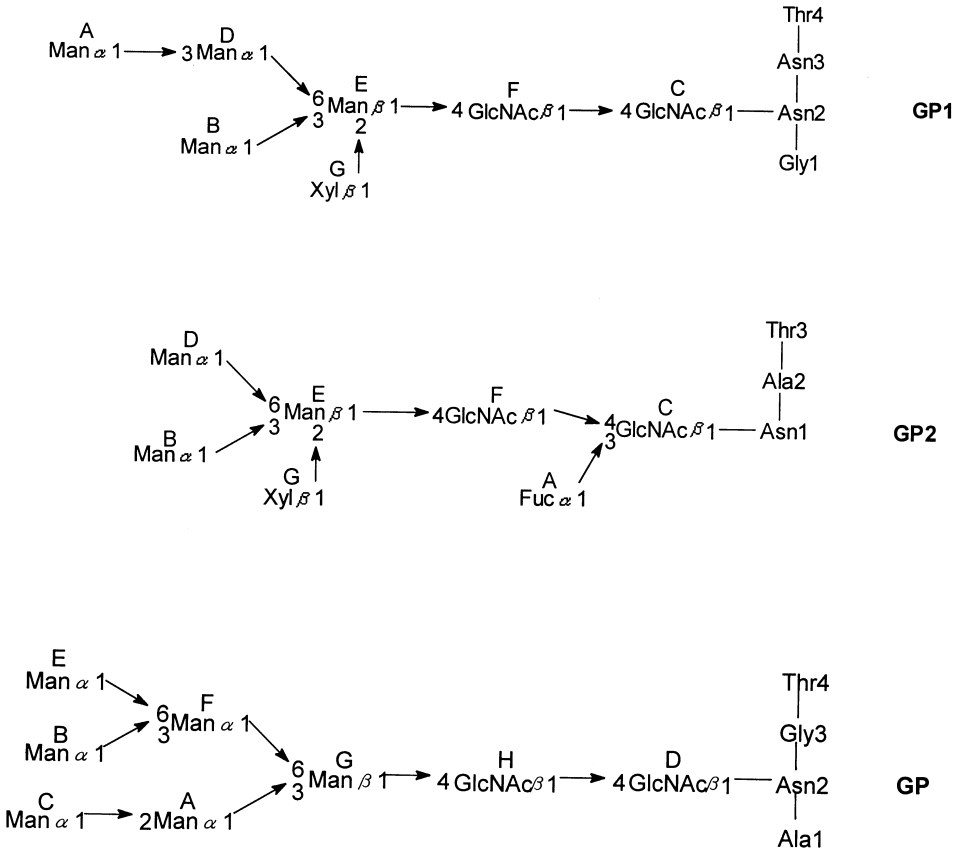

Chart 1. Schematic structure of glycopeptides GP1, GP2 and GP3. Gly1 is at N-terminus of the peptide in GP1. Asn1 is at N-terminus of the peptide in GP2. Ala1 is at N-terminus of the peptide in GP3.

remained sugar residues in GP3 were all α-D-mannoses. The amino acid residues of GP3 were elucidated as alanine, asparagine, glycine and threonine. The linkages among sugar residues and amino acids, as well as between the peptide and glycan, were solved by the NOESY and HMBC spectra. The primary structure of GP3 was deduced as Manα1,6(Manα1,3)-Manα1,6(Manα1,2 Manα1,3)Manβ1,4GlcNAcβ1,4GlcNAc-β1-(Ala-)Asn-Gly-Thr (Chart 1) The chemical shifts of protons and carbons were assigned as shown in Table 3.

ESI-MS analysis of GP1, GP2 and GP3

The ESI-MS spectrum of GP1 demonstrated quasi-molecular ion peaks at m/z 1593$[M + H]^+$ and 1615$[M + Na]^+$, indicating that its molecular weight was 1592 dalton. This value was consistent with that (1591.45 D) of the glycopeptide ($C_{59}H_{98}N_8O_{42}$) composed of seven sugar residues (2 GlcNAc, 4 Man and 1 Xyl) and four amino acid residues (1 Gly, 2 Asn and 1 Thr). The ESI-MS spectrum of GP2 showed quasi-molecular ion peaks at m/z 1476$[M + H]^+$ and 1498 $[M + Na]^+$, indicating that it had molecular weight of 1475 dalton. This value was consistent with that (1475.37 D) of the glycopeptide ($C_{56}H_{94}N_6O_{39}$) composed of seven sugar residues (2 GlcNAc, 3 Man, 1 Xyl and 1 Fuc) and three amino acid residues (1 Asn, 1 Ala and 1 Thr). The ESI-MS spectrum of GP3 exhibited quasi-molecular ion peaks at m/z 1742 $[M + H]^+$ and 1764$[M + Na]^+$, indicating that it molecular weight was 1741 dalton. This value was consistent with that (1740.60 D) of the glycopeptide ($C_{65}H_{109}N_7O_{47}$) composed of eight sugar residues (2 GlcNAc and 6 Man) and four amino acid residues (1 Ala, 1 Asn, 1 Gly and 1 Thr).

Discussion

After enzymatic digestion of cinnamomin B-chain, three major glycopeptides were obtained through three purification steps: gel filtration chromatography, anion-exchange chromatography and HPLC. These methods might be useful to purify other glycopeptides for studying their primary structure by NMR, which needed large amount of sample with the high purity.

In this paper, the primary structures of three glycopeptides GP1, GP2 and GP3 were determined by two dimensional NMR including TOCSY, DQF-COSY, NOESY, HMQC and HMBC spectral analysis and the ESI-MS data. One of advantages of 2D-NMR method was that it afforded complete assignments of ^1H and ^{13}C NMR signals of the molecule.

Table 2. ^{1}H- and ^{13}C-chemical shifts of GP2[a]

Sugar	*^{1}H*	*^{13}C*	*HMBC ($^{3}J_{C\text{-}H}$)*
α-L-Fuc (A)			
1	5.16(4.2)	98.42	H_{C3}
2	3.75	67.88	
3	4.00	69.42	
4	3.85	72.23	
5	4.73	66.94	
6	1.29	15.62	
α-D-man (B)			
1	5.15	102.33	H_{E3}
2	4.07	70.08	
3	3.85	70.57	
4	3.67	66.92	
5	3.65	72.81	
6	3.96, 3.79	61.14	
β-D-GluNAc (C)			
1	5.11(9.6)	78.13	
2	4.06	54.41	
3	3.95	75.24	
4	3.94	73.46	
5	3.61	77.13	
6	3.75, 3.91	59.88	
NHCO	8.20	174.42	
CH_3	2.02	22.30	
α-D-man (D)			
1	4.94	99.78	
2	4.00	69.90	
3	3.87	70.41	
4	3.66	66.93	
5	3.76	72.37	
6	3.92, 3.77	61.03	
β-D-man (E)			
1	4.88	100.67	H_{F4}
2	4.29	79.24	
3	3.87	77.53	
4	3.68	66.60	
5	3.68	74.61	
6	3.98, 3.80	65.49	
β-D-GluNAc (F)			
1	4.59(8.3)	100.40	
2	3.80	55.26	
3	3.75	72.82	
4	3.50	80.90	
5	3.60	74.81	
6	3.91, 3.70	61.15	
NHCO	8.38	174.78	
CH_3	2.08	22.30	
β-D-Xyl (G)			
1	4.49(7.7)	105.09	H_{F2}
2	3.40	73.51	
3	3.48	75.52	
4	3.66	69.45	
5	4.04, 3.29	65.03	
Asn1			
NH	–[b]		
α-CH	3.87	51.29	
β-CH	2.77, 2.61	39.67	
C=O		173.32	
CONH (side chain)	8.47	175.00	
Ala2			
NH	7.76		
α-CH	4.44	49.90	
β-CH	1.45	16.73	
C=O		174.42	NH-Thr3 ($^{2}J_{C\text{-}H}$)
Thr3			
NH	7.78		
α-CH	4.17	60.48	
β-CH	4.26	68.12	
γ-CH_3	1.21	19.31	
C=O		176.56	

[a] Water (δ_H 4.76) was used as the reference for proton chemical shifts and chloroform (δ_C 77.0) as the external reference for carbon chemical shifts.
[b] –, Not detected due to the exchange between proton and deuteron.

Furthermore, the spatial information between protons and $^{3}J_{C\text{-}H}$ coupling constants across glycoside linkages revealed by 2D-NMR spectra could provide crucial data for the conformation analysis of oligosaccharide chain such as inter-proton distances and dihedral angles across glycoside linkages. On the other hand, the primary structure of the short peptide chain linking to the carbohydrate moiety in the glycopeptides, at the same time, was also determined by this method. On the basis of the sequence of the peptide, the glycosylation site of each oligosaccharide chain in the original protein could be established. In this study, the amino acid sequences of GP1, GP2 and GP3 were determined as Asn-Asn-Thr, Asn-Ala-Thr and Ala-Asn-Gly-Thr, respectively, that could be used in determining the glycosylation sites of the oligosaccharide chains in cinnamomin B-chain.

When the sugar chain of glycopeptide links to the side chain of asparagine, it is called as N-glycan. Generally, N-glycans of plant RIPs can be classified into two main groups: complex type and high mannose type. Both types contain a "core portion" (reducing end) that is composed of two N-acetylglucosamine (GlcNAc) and three mannose residues. GP1 and GP2 belong to the complex type because the former contains a xylose and the latter contains a xylose and a fucose besides the "core portion"; while GP3 belongs to the high mannose type because it contains only mannose besides the "core portion". The "core portion" of N-glycans in GP1, GP2 and GP3 are identical with those of other plant RIPs. In addition, the same sugar chain structures of GP1, GP2 and GP3 are also found in several RIPs from different plants [17–21], suggesting that these glycoproteins may undergo a similar post-translation processing and transport. The possible processing pathways of the sugar chain synthesis of plant glycoprotein have been proposed [22,23].

The short peptides in GP1, GP2 and GP3 are very helpful for deciphering the glycosylation site. The peptide sequences of GP1 and GP2 were found in the complete amino acid sequence of cinnamomin B-chain obtained by cDNA method

Table 3. ^{1}H- and ^{13}C-chemical shifts of GP3[a]

Residue	*^{1}H*	*^{13}C*	*HMBC ($^{3}J_{C-H}$)*
α-D-man (A)			
1	5.37	100.89	H_{G3}
2	4.13	78.43	H_{C1}
3	4.02	70.12	
4	3.81	65.58	
5	3.71	72.75	
6	3.94, 3.78	61.16	
α-D-man (B)			
1	5.12	102.33	H_{F3}
2	4.08	70.00	
3	3.91	70.42	
4	3.74	72.83	
5	3.70	66.80	
6	3.85, 3.66	59.94	
α-D-man (C)			
1	5.08	102.28	H_{A2}
2	4.09	70.00	
3	3.87	70.37	
4	3.69	66.80	
5	3.76	73.29	
6	3.90, 3.76	60.09	
β-D-GluNAc (D)			
1	5.07 (9.6)	78.33	
2	3.87	53.69	
3	3.77	73.51	
4	3.67	78.72	H_{H1}
5	3.58	76.24	
6	3.86, 3.78	61.08	
NHCO	8.19	174.79	
CH_3	2.04	22.12	
α-D-man (E)			
1	4.93	99.35	H_{F6}
2	4.01	70.00	
3	3.86	70.64	
4	3.70	66.80	
5	3.72	72.75	
6	3.91, 3.79	61.01	
α-D-man (F)			
1	4.89	99.94	
2	4.16	69.52	
3	3.94	78.66	H_{B1}
4	3.88	65.74	
5	3.87	70.92	
6	4.01, 3.77	65.30	H_{E1}
β-D-Man (G)			
1	4.79	100.41	H_{H4}
2	4.25	70.19	
3	3.76	80.77	H_{A1}
4	3.65	74.46	
5	3.82	73.51	
6	3.97, 3.79	66.05	H_{F1}
β-D-GluNAc (H)			
1	4.62 (7.8)	101.42	H_{D4}
2	3.80	55.04	
3	3.79	72.02	
4	3.75	79.57	
5	3.62	74.54	
6	3.89, 3.81	61.08	
NHCO	8.36	174.49	
CH_3	2.08	22.27	
Ala1			
NH	–[b]		
α-CH	4.14	49.06	
β-CH	1.56	16.39	
C=O		170.88	
Asn2			
NH	8.87		
αC-H	4.77	50.38	
β-CH	2.94, 2.80	36.49	
C=O		172.36	NH-G1y3 ($^{2}J_{C-H}$)
CONH (side chain)	8.63	172.53	
Gly3			
NH	8.41		
α-CH	4.11, 3.98	42.65	
C=O		171.22	NH-Thr4 ($^{2}J_{C-H}$)
Thr4			
NH	7.79		
α-CH	4.16	62.24	
β-CH	4.30	67.68	
γ-CH_3	1.21	19.08	
C=O		172.53	

[a]Water (δ_H 4.76) was used as the reference for proton chemical shifts and chloroform (δ_C 77.0) as the external reference for carbon chemical shifts.
[b]–, Not detected due to the exchange between proton and deuteron.

(our unpublished data). Because GP1, GP2 and GP3 are the main glycopeptides, these sugar chains are composed of the main sugar content of cinnamomin B-chain.

Acknowledgement

This work is supported by two grants from National Natural Science Foundation of China, a grant (KSCx2-2-04) from Chinese Academy of Sciences and a grant from Shanghai Institutes for Biological Sciences. The authors of SIOC thank the State Ministry of Science and Technology for financial support.

References

1 Barbieri L, Battelli MG, Stirpe F, *Biochim Biophys Acta* **1154**, 237–82 (1993).
2 Kimura Y, Hase S, Ikenaka T, Funatsu G, *Biochim Biophys Acta* **966**, 150–9 (1988).
3 Kimura Y, Minami Y, Tokuda T, Nakajima S, Takagi S, Funatsu G, *Agric Biol Chem* **55**, 2031–6 (1991).
4 Kimura Y, Hase S, Kobayashi Y, Kyogoku Y, Ikenaka T, Funatsu G, *J Biochem* **103**, 944–9 (1988).
5 Kimura Y, Hase S, Ikenaka T, Funatsu G, *Biochim Biophys Acta* **966**, 160–7 (1988).
6 Agrawal PK, *Phytochemistry* **31**, 3307–30 (1992).

7 Ling J, Liu WY, Wang TP, *Biochim Biophys Acta* **1252**, 15–22 (1995).
8 O'Connor SE, Imperiali B, *Chem Biol* **3**, 803–12 (1996).
9 Dwek RA, Lellouch AC, Wormald MR, Glycobiology: *J Anatomy* **187**, 272–9 (1995).
10 Wyss DF, Choi JS, Li J, Knoppers MH, Willis KJ, Arulanandam ARN, Smolyar A, Reinherz EL, Wagner G, *Science* **269**, 1273–8 (1995).
11 Xie L, Hou FJ, Liu WY, Wang ED, *Protein Peptide Lett* **7**, 265–72 (2000).
12 Pu Z, Xie L, Wang ED, Liu WY, *Biol Chem* **379**, 1413–8 (1998).
13 Dubois D, Gilles KA, Hamilto JK, Smith F, *Anal Chem* **28**, 350–5 (1956).
14 Funatsu G, *Biochemistry* **3**, 1351–5 (1964).
15 Gerwig GJ, Kamerling JP, Vliegenthart JF, Morag E, Lamed R, Bayer EA, *J Bio Chem* **268**, 26956–60 (1993).
16 Nanno S, Ishiguro M, Funatsu G, Funatsu M, *Agric Biol Chem* **39**, 1645–50 (1975).
17 Hase S, Koyama S, Daiyasu H, Takemoto H, Hara S, Kobayashi Y, Kyogoku Y, Ikenaka T, *J Biochem* **100**, 1–10 (1986).
18 Takahashi N, Hotta T, Ishihara H, Mori M, Tejima S, Bligny R, Akazawa T, Endo S, Arata Y, *Biochemistry* **25**, 388–95 (1986).
19 Takayama S, Isogai A, Tsukamoto C, Ueda Y, Hinata K, Okazaki K, Suzuki A, *Agric Biol Chem* **50**, 1673–6 (1986).
20 Kitagawa-Ogawa H, Matsumoto I, Seno N, Takahashi N, Endo S, Arata Y, *Eur J Biochem* **161**, 779–85 (1986).
21 Kimura Y, Ohno A, Takagi S, *Biosci Biotech Biochem* **61**, 1866–71 (1997).
22 Funatsu G, Yoshitake S, Funatsu M, *Agric Biol Chem* **42**, 501–3 (1978).
23 Kimura Y, Hase S, Kobayashi Y, kyogoku Y, Funatsu G, Ikenaka T, *J Biochem* **101**, 1051–4 (1987).
24 Johnson KD, Chrispeels MJ, *Plant Physiol* **84**, 1301–8 (1987).

Received 4 May 2000, revised 20 December 2000, accepted 2 March 2001