Structural determination of the N-glycans of a lepidopteran arylphorin reveals the presence of a monoglucosylated oligosaccharide in the storage protein

Soo hyun Kim, Soo kyung Hwang, Raymond A. Dwek, Pauline M. Rudd, Yeong Hee Ahn, Eun-Hee Kim, Chaejoon Cheong, Seung Il Kim, Nam Sook Park, and Sang Mong Lee

2 Proteome Analysis Team, Korea Basic Science Institute, Daejeon 305-806, Korea; 3 Glycobiology Institute, Department of Biochemistry, University of Oxford, Oxford, OX1 3QU, UK; 4 Magnetic Resonance Team, Korea Basic Science Institute, Daejeon 305-806, Korea; and 5 Department of Sericultural and Entomological Biology, Myriyang National University, Myriyang 627-702, Korea

Received on May 2, 2002; revised on October 8, 2002; accepted on October 8, 2002

The structures of the oligosaccharides attached to arylphorin from Chinese oak silkworm, Antheraea pernyi, have been determined. Arylphorin, a storage protein present in fifth larval hemolymph, contained 4.8% (w/w) of carbohydrate. This glycan structure is not normally present in secreted mammalian glycoproteins; however, it has been identified in avian species, Anser anser, and was characterized by normal-phase HPLC, mass spectrometry oligosaccharide profiling confirmed that arylphorin contained mainly oligomannose-type glycans as well as truncated manno-type structures with or without fucosylation. Interestingly, the most abundant oligosaccharide was monoglucosylated Man9-GlcNAc2, which was characterized by normal-phase HPLC, mass spectrometry, Aspergillus saitoi α-mannosidase digestion, and 1H 600 MHz NMR spectroscopy. This glycan structure is not normally present in secreted mammalian glycoproteins; however, it has been identified in avian species. The Glc3Man9GlcNAc2 structure was present only in arylphorin, whereas other hemolymph proteins contained only oligomannose and truncated oligosaccharides. The oligosaccharide was also detected in the arylphorin of another silkworm, Bombyx mori, suggesting a specific function for the Glc3Man9GlcNAc2 glycan. There were no processed glucosylated oligosaccharides such as Glc3Man5,6-GlcNAc2. Furthermore, Glc1Man9GlcNAc2 was not released from arylphorin by PNGase F under nondenaturing conditions, suggesting that the N-glycosidic linkage to Asn is protected by the protein. Glc1Man9GlcNAc2 may play a role in the folding of arylphorin or in the assembly of hexamers.

Key words: Antheraea pernyi storage protein/arylphorin glycosylation/monoglucosylated oligomannose

Introduction

The study of insect glycosylation is very fragmentary compared to that of mammalian systems. It was, however, established that Drosophila Kc cells contained lipid-linked oligosaccharide precursors as well as dolichol-linked Glc or Man (Sagami and Lennarz, 1987), and D. melanogaster had the same molecular size as the lipid-linked precursor Glc1Man3GlcNAc2 (Parker et al., 1991), indicating common N-glycosylation in insects as well as in mammalian cells. An early study of insect glycosylation showed that complex N-glycans of vertebrate cell-derived Sindbis virus were replaced by Man3GlcNAc2 when the same protein was expressed in mosquito C6/36 cells (Hsieh and Robbins, 1984). Although the structural analysis was incomplete, these data suggested that insect N-linked oligosaccharides were mainly of the oligomannose type. Subsequently, paucimannose- and oligomannose-type oligosaccharides ranging from Man2GlcNAc2 to Man9GlcNAc2 with or without α-6-linked Fuc were described in D. melanogaster larval membrane glycoproteins and larval serum protein (Williams et al., 1991).

Unusual GlcNAc-terminated glycans with carbohydrate-linked 2-aminoethylphosphonate (Hard et al., 1993), α1-3-monofucosylated or α1-3(α1-6)difucosylated oligosaccharides, and terminal trisaccharide GalNAcβ1-4[Fucα1-3]GlcNAcβ1- on the core α1-3 mannosyl residue (Kubelka et al., 1993, 1995) have been also described in insects. Recent studies with nonengineered insect cells Trichoplusia ni (Rudd et al., 2000; Ailor et al., 2000) or Spodoptera frugiperda (Davidson et al., 1990) have revealed hybrid and complex-type oligosaccharide structures, albeit with limited amounts. These may indicate that insects have the potential to synthesize diverse glycan structures, although the majority of the N-linked oligosaccharides of insects found to date are of the oligomannose type.

Storage proteins, such as arylphorin, are synthesized from the fat body of a wide range of lepidopteran and dipteran larvae and also in other insect orders. They are the major proteins in the hemolymph. Storage proteins are taken up by the fat body shortly before pupation and stored in protein granules. The central function of the proteins seems to be a storage pool of the amino acid resources for complete development of the adult. There are at least two kinds of storage proteins in lepidoptera. They form loose clusters on a bivariate plot of the proportion of aromatic amino acids (tyrosine plus phenylalanine) versus methionine (Telfer and Kunkel, 1991). Arylphorin is rich in aromatic amino acids, and the second protein is referred to as methionine-rich storage protein. All have molecular masses of nearly 500 kDa and are hexamers composed of approximately 80 kDa subunits.
Antheraea pernyi is a lepidopteran and an economically important wild silk moth found in Europe and Far Eastern Asia. There is no report to describe oligosaccharide structure or function in A. pernyi. The work described here is the most detailed study of oligosaccharides in the lepidopteran storage protein arylphorin of A. pernyi, so far. Unexpectedly, arylphorin had monoglucosylated Glc1Man9GlcNAc2 in the oligosaccharide pool of A. pernyi as well as in that of Bombyx mori. This sugar may play a role in the folding of arylphorin or in the assembly of hexamer.

Results

Monosaccharide composition analysis

Monosaccharide analysis was carried out on the purified arylphorin of A. pernyi. The monosaccharide composition of the oligosaccharides was determined by high-pH anion exchange chromatography (HPAEC) following acid hydrolysis (Table I). The arylphorin of A. pernyi had 4.8% (w/w) carbohydrate based on composition analysis. GlcNAc and Man were the main sugar components of the arylphorin. Four moles GlcNAc per mole protein suggested that it had two N-glycosylation sites. The high proportion of Man compared to GlcNAc suggested that the major glycans of the arylphorin were of the oligomannose type. The analysis provided evidence of Fuc, which is commonly found in glycans attached to insects or plant glycoproteins. There was about 1 mole of glucose. It was unlikely the protein had mucin type O-linked oligosaccharides because there was no detectable Gal and GalNAc.

Storage proteins (SP1 and SP2) of B. mori were obtained by simple gel filtration chromatography. SP1 (methionine-rich) and SP2 (arylphorin) were identified by N-terminal amino acid sequencing (data not shown). These storage proteins also contained a moderate amount of sugars: 6.9% (w/w). The monosaccharide composition (Table I) was very similar to that of A. pernyi except for the presence of GalNAc, which is a strong indicator of O-glycosylation. It was not clear, however, that GalNAc was released from methionine-rich protein or arylphorin.

Oligosaccharide profile

As a final step, A. pernyi arylphorin used for oligosaccharide analysis was purified by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) to remove other possible contaminating glycoproteins or glycans, even though it was estimated to be already more than 95% pure. An aliquot of the fluorescent-labeled glycans released by in-gel digestion with peptide-N-glycosidase F (PNGase F) was analyzed by normal-phase (NP) high-performance liquid chromatography (HPLC). The profile showed a clearly separated set of peaks, suggesting a series of oligosaccharides each differing by one monosaccharide unit (Figure 1A). Individual peaks were assigned glucose unit (GU) values by comparison with the elution positions of a standard 2-aminobenzamide (AB)-labeled dextran hydrolysate ladder. Then, each peak was assigned a preliminary structure using the elution positions and predetermined incremental values for monosaccharide residues (Guile et al., 1996). These data indicated that the major glycans were probably of the nonfucosylated oligomannose type; however, the predominant species was contained in the peak 11.

<table>
<thead>
<tr>
<th>Carbohydrate content (%, w/w)</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aarylphorin (A. pernyi)</td>
<td>4.8</td>
</tr>
<tr>
<td>SP1 and 2 (B. mori)</td>
<td>6.9</td>
</tr>
</tbody>
</table>

The proteins were hydrolyzed with acids and the amounts of monosaccharides were measured by HPAEC as described in Materials and methods.

![Fig. 1. HPLC profiles of the glycan population of A. pernyi arylphorin simultaneously digested with exoglycosidases. The HPLC analysis of the total glycan pool (A) and the products resulting from the digestion of five aliquots of the total A. pernyi arylphorin glycan pool with exoglycosidases, jack bean α-mannosidase (JB mannosidase) (B); JB mannosidase plus almond meal α-fucosidase (AM fucosidase) (C); JB mannosidase plus bovine kidney α-fucosidase (BK fucosidase) (D); A. saitoi α-mannosidase (AS mannosidase) (E); and X. manihotis α-mannosidase (XM mannosidase) (F) indicated in each panel. The glucose unit value of each peak was calculated (Table II) by comparison with the glucose ladder shown at the top of the figure. Structures were assigned from the glucose unit values, previously determined incremental values for monosaccharide residues (Guile et al., 1996) and the known specificity of the exoglycosidase enzymes.](https://academic.oup.com/glycob/article-abstract/13/3/147/608960)
(10.20 GU) with higher retention time than Man$_{3}$GlcNAc$_{2}$ (Figure 1A). Arylphorin of *A. pernyi* was well separated from other hemolymph proteins by gel-filtration chromatography, and the protein recovered in this way had the same oligosaccharide profile as glycans released from arylphorin in SDS-PAGE gel bands (data not shown). Other experiments, including the hydrazinolysis described later, were carried out with arylphorin purified by gel filtration.

The glycans were also prepared by automatic hydrazinolysis using the GlycoPrep 1000 operated in $N + O$ mode to determine if there were $N$-glycans that could not be released by PNGase F because of the presence of a Fuc residue linked $\alpha$1-3 to the core GlcNAc residues and whether or not arylphorin contained any $O$-glycans. The chromatographic profile was identical to that generated by PNGase F (data not shown). It indicated that the protein contained only $N$-glycans, consistent with the monosaccharide composition analysis (Table I).

### Exoglycosidase sequencing of the oligosaccharides

The results strongly suggested that *A. pernyi* arylphorin contained mainly oligomannose-type oligosaccharides. To confirm this, aliquots of the glycan pool were digested with exoglycosidases alone or in combination. The specific activity of each enzyme was confirmed by digestion of authentic standards. The digestion mixtures were resolved by NP HPLC (Figure 1B-F). The assignment of each peak in the glycan pool was checked by following its predicted elution position through each of the enzyme digests and by comparison with authentic standards. As expected, with the exception of three minor peaks, all of the structures collapsed into Man$_{1}$GlcNAc$_{2}$ (peak 12) following incubation with jack bean $\alpha$-mannosidase (Figure 1B), which removes $\alpha$-mannose (Man$_{1}$-2, 3, 6Man). The minor peaks were not digested with a combination of $\alpha$-mannosidase and almond meal $\alpha$-fucosidase (Figure 1C). However, simultaneous treatment of the glycan with the $\alpha$-mannosidase and bovine kidney $\alpha$-fucosidase removed peak 13 from the total glycan pool, suggesting that Fuc was attached to proximal GlcNAc via an $\alpha$1-6 linkage (Figure 1D).

Two undigested peaks (14 and 15) remained after jack bean $\alpha$-mannosidase and bovine kidney $\alpha$-fucosidase treatment (Figure 1D). These structures were resistant to a range of exoglycosidases either alone or in combinations (*A. ureafaciens* sialidase, bovine testes $\beta$-galactosidase, jack bean $\beta$-$N$-acyethylhexosaminidase, bovine kidney $\alpha$-fucosidase, almond meal $\alpha$-fucosidase, or bovine testes $\beta$-galactosidase and jack bean $\beta$-$N$-acyethylhexosaminidase with jack bean $\alpha$-mannosidase) (data not shown). There was no peak corresponding to peak 14 (6.00 GU) in the total glycan pools, which indicated that the peak was generated by the $\alpha$-mannosidase digestion. Peak 15 (6.89 GU) eluted close to peak 7 (6.95 GU), corresponding to Man$_{3}$GlcNAc$_{2}$. Peak 7 decreased significantly after digestion with jack bean $\alpha$-mannosidase, however peak 15 remained, suggesting that the two glycans are of the same size but of different composition.

When the glycan pool was subjected under *A. saitoi* $\alpha$-mannosidase treatment that is specific for Man$_{1}$-2Man, positions and patterns of peaks 1-4 were not changed, indicating no terminal Man linked via $\alpha$1-2 to these glycans (Figure 1E). Peak 5 that seemed to be Man$_{3}$GlcNAc$_{2}$ compared to standard oligosaccharide was increased, whereas peaks 6-11 disappeared with emerging peak 16 (8.60 GU) around peak 8 (8.67 GU). It was likely that peak 16 was moved from peak 11, because the amount of peak 16 fraction (25.5%) was only comparable to that of peak 11 (28.3%). Furthermore, an extra aliquot of the exoglycosidase did not change the chromatogram, suggesting that peak 16 was the final product of *A. saitoi* $\alpha$-mannosidase digestion. The jack bean and *A. saitoi* $\alpha$-mannosidase digestion results together with the profiling of the total oligosaccharide pool demonstrated that, with the exception of peak 11, the major glycan pool of arylphorin might consist of typical oligomannosetype structures. *Xanthomonas manihotis* $\alpha$-mannosidase that is specific for unbranched Man$_{3}$-6Man was used to determine if there was a Man with other than $\alpha$1-2 linkage. Definitely, peaks 1 and 2 were moved to peaks 12 and 13, respectively, but other glycans were not changed at all on the chromatogram (Figure 1F). It demonstrated that peaks 1 and 2 had a terminal $\alpha$1-6-linked Man on $\beta$-mannosyl chitobiose and core-fucosylated one, respectively.

To confirm or determine the structures of all the glycans, each major glycan peak larger than 6.0 GU was fractionated from NP HPLC runs and digested with jack bean $\alpha$-mannosidase, individually. Purified peak 11 split into two peaks (14 and 15) (Figure 2), and other oligosaccharide fractions were fully digested to Man$_{1}$GlcNAc$_{2}$ (data not shown). The proportion of peak 14 increased; although peak 15 decreased it was still detectable even when incubation was carried out with a second aliquot or higher enzyme concentration (Figure 2B, C). This suggested that peak 14
was the final digestion product of the peak 11 by jack bean \(\alpha\)-mannosidase but that peak 15 was very resistant to the enzyme digestion. Moreover, matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (MS) showed clearly that the two peaks were composed of 2GlcNAc and 4 or 5 hexoses (Figure 2D). We postulated that peaks 11, 14, and 15 corresponded to Hex\(_1\)Man\(_9\)GlcNAc\(_2\), Hex\(_1\)Man\(_5\)GlcNAc\(_2\), and Hex\(_1\)Man\(_3\)GlcNAc\(_2\), respectively (Figure 2E), and peak 11 produced two peaks by jack bean \(\alpha\)-mannosidase treatment. These data are consistent with the presence of Glc\(_1\)Man\(_9\)GlcNAc\(_2\). In structure 16 (Figure 2E), the \(\alpha\)1,6-linked Man is particularly resistant to digestion with jackbean \(\alpha\)-mannosidase when the \(\alpha\)1,3 arm is substituted (Cacan et al., 2001).

**MS**

To determine the size and monosaccharide composition of the glycans for further characterization, the pool of 2-AB-labeled oligosaccharides was subjected to MALDI MS analysis (Figure 3). MALDI MS signals indicated that arylphorin contained at least 10 glycan structures. These included truncated glycans with or without fucosylation and oligomannose with compositions (Table II). The mass value (m/z 1540.7) of isolated peak 6 (6.76 GU), which appeared around peak 7 (m/z 1540.4, 6.95 GU), suggested that the peak was an isomer of Man\(_9\)GlcNAc\(_2\). The largest structure detected was consistent with an assignment of Hex\(_{10}\)GlcNAc\(_2\).

**Analysis of peak 11**

If an isoform of Hex\(_1\)Man\(_9\)GlcNAc\(_2\) (where Hex is not Man) is digested by \(A.\) *saitoi* \(\alpha\)-mannosidase, the digested oligosaccharide would be Hex\(_1\)Man\(_6\)GlcNAc\(_2\) or Hex\(_1\)Man\(_7\)GlcNAc\(_2\) according to the Hex position as shown in Figure 4A. The specificity of the \(A.\) *saitoi* \(\alpha\)-mannosidase was confirmed by digestion of a standard, Man\(_9\)GlcNAc\(_2\) (Figure 4B, D). The \(\alpha\)-mannosidase digestion shifted peak 11 (10.20 GU) to near Man\(_8\)GlcNAc\(_2\) position (8.67 GU, Figure 4C, E). It was consistent with the shift of peak 11 to peak 16 by \(A.\) *saitoi* \(\alpha\)-mannosidase treatment in arylphorin glycan pool (Figure 1E). There is about 0.1 GU difference when Glc or Man on the same backbone oligosaccharide is separated by NP HPLC. The results shown in Figure 4 therefore suggested strongly that the Hex is linked to D1 arm of Man\(_9\)GlcNAc\(_2\). When a 2-AB-labeled standard oligosaccharide, Glc\(_1\)Man\(_9\)GlcNAc\(_2\), where Glc is linked to Man of D1 arm via \(\alpha\)-3, was digested by jack bean \(\alpha\)-mannosidase, the digested fragments migrated on the NP HPLC in positions matching those of peak 11.

**Table II.** Details of the glycans of *A. pernyi* arylphorin (2-AB, [M+Na\(^+\)])

<table>
<thead>
<tr>
<th>Peak(^a)</th>
<th>GU(^b)</th>
<th>Measured</th>
<th>Calculated</th>
<th>% Total(^d)</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.54</td>
<td>891.9</td>
<td>892.0</td>
<td>2.9</td>
<td>2 2 0</td>
</tr>
<tr>
<td>2</td>
<td>3.98</td>
<td>1038.0</td>
<td>1038.0</td>
<td>13.2</td>
<td>2 2 1</td>
</tr>
<tr>
<td>3</td>
<td>4.43</td>
<td>1054.0</td>
<td>1054.0</td>
<td>3.4</td>
<td>3 2 0</td>
</tr>
<tr>
<td>4</td>
<td>4.93</td>
<td>1200.1</td>
<td>1200.1</td>
<td>5.4</td>
<td>3 2 1</td>
</tr>
<tr>
<td>5</td>
<td>6.20</td>
<td>1378.2</td>
<td>1378.2</td>
<td>12.9</td>
<td>5 2 0</td>
</tr>
<tr>
<td>6</td>
<td>6.76</td>
<td>1540.7</td>
<td>1540.2</td>
<td>3.1</td>
<td>6 2 0</td>
</tr>
<tr>
<td>7</td>
<td>6.95</td>
<td>1540.4</td>
<td>1540.2</td>
<td>19.5</td>
<td>6 2 0</td>
</tr>
<tr>
<td>8</td>
<td>7.82</td>
<td>1702.5</td>
<td>1702.3</td>
<td>1.9</td>
<td>7 2 0</td>
</tr>
<tr>
<td>9</td>
<td>8.67</td>
<td>1864.6</td>
<td>1864.3</td>
<td>2.2</td>
<td>8 2 0</td>
</tr>
<tr>
<td>10</td>
<td>9.50</td>
<td>2026.7</td>
<td>2026.4</td>
<td>7.2</td>
<td>9 2 0</td>
</tr>
<tr>
<td>11</td>
<td>10.20</td>
<td>2188.8</td>
<td>2188.5</td>
<td>28.3</td>
<td>10 2 0</td>
</tr>
</tbody>
</table>

\(^a\)Peak number in Figure 1.

\(^b\)Glucose unit.

\(^c\)Monoisotopic mass of the [M+Na\(^+\)]\(^+\) ion.

\(^d\)Amount of glycan expressed as a percentage of the total glycans measured by NP HPLC.
Oligosaccharides for nuclear magnetic resonance (NMR) analysis were released from *A. pernyi* arylphorin by PNGase F and fractionated by selective precipitation and preparative HPAEC as described under Materials and methods. The HPAEC chromatogram was equivalent to that of the total glycan pool labeled with 2-AB (Figure 5A). One of the fractions coeluted with Man$_9$GlcNAc$_2$ (peak 10), which was located after Man$_9$GlcNAc$_2$ (peak 10). The peak 11 oligosaccharide fractionated by HPAEC was labeled with 2-AB and analyzed by NP HPLC. The elution position (10.20 GU) corresponded to that of peak 11 (data not shown).

The peak 11' fraction was subjected to 600 MHz $^1$H NMR spectroscopy to complete structural determination. The signals for the oligosaccharide were assigned by comparing the data with the spectrum of Glc$_1$Man$_9$GlcNAc$_2$, reported by Matsuura *et al.* (1993). The values of the $^1$H chemical shifts for the oligosaccharide were exactly the same as those obtained from standard Glc$_1$Man$_9$GlcNAc$_2$ purified from hen albumen. The presence of an $\alpha$-Glc residue in the fraction manifested itself by the characteristic signal for $^1$H of Glc-$\alpha$ ($\delta = 5.245$) and by upfield shift of Man-$\alpha$ ($\delta = 5.036$). The results, together with the monosaccharide composition analysis and $\alpha$-mannosidase digestions just described, indicated that oligosaccharide contained an additional Glc residue linked to the nonreducing terminal Man-$\alpha$ of Man$_9$GlcNAc$_2$ via $\alpha$1-3 as shown in Table III and Scheme 1.

Fig. 4. Determination of hexose position by *Aspergillus saitoi* $\alpha$-mannosidase digestion. (A) D$_1$ type of Hex$_1$Man$_5$GlcNAc$_2$ isoform generates Hex$_1$Man$_7$GlcNAc$_2$ unlike Hex$_1$Man$_9$GlcNAc$_2$ from D$_2$ or D$_3$ isoforms by *A. saitoi* (AS) $\alpha$-mannosidase. (B) A standard Man$_9$GlcNAc$_2$ oligosaccharide was used to confirm AS mannosidase specificity, and (C) it produced Man$_8$GlcNAc$_2$. (D) Peak 11 was digested to identify the Hex position with the enzyme, and (E) it shifted to 10.2 GU consistent with the peak assignment of Glc$_1$Man$_9$GlcNAc$_2$.

Fig. 5. Purification of the proposed Glc$_1$Man$_9$GlcNAc$_2$ target oligosaccharide by HPAEC. Oligosaccharides were extracted from 10 mg of denatured arylphorin as described under Materials and methods. (A) Released oligosaccharides were separated with 0-0.25 M sodium acetate gradient on CarboPac PA1 (9 x 250 mm) and monitored with pulsed amperometric detection. Peak 10' was coeluted with a standard Man$_9$GlcNAc$_2$ oligosaccharide. (B) A part of the isolated target oligosaccharide fraction was rerun to confirm purity. Peak 11' in each panel indicates target and isolated oligosaccharide.

### 1H NMR spectroscopy

Oligosaccharides for nuclear magnetic resonance (NMR) analysis were released from *A. pernyi* arylphorin by PNGase F and fractionated by selective precipitation and preparative HPAEC as described under Materials and methods. The HPAEC chromatogram was equivalent to that of the total glycan pool labeled with 2-AB (Figure 5A). One of the fractions coeluted with Glc$_1$Man$_9$GlcNAc$_2$ (peak 11'), which was located after Man$_9$GlcNAc$_2$ (peak 10'). The peak 11' oligosaccharide fractionated by HPAEC was labeled with 2-AB and analyzed by NP HPLC. The elution position (10.20 GU) corresponded to that of peak 11 (data not shown).

The peak 11' fraction was subjected to 600 MHz $^1$H NMR spectroscopy to complete structural determination. The signals for the oligosaccharide were assigned by comparing the data with the spectrum of Glc$_1$Man$_9$GlcNAc$_2$, reported by Matsuura *et al.* (1993). The values of the $^1$H chemical shifts for the oligosaccharide were exactly the same as those obtained from standard Glc$_1$Man$_9$GlcNAc$_2$ purified from hen albumen. The presence of an $\alpha$-Glc residue in the fraction manifested itself by the characteristic signal for $^1$H of Glc-$\alpha$ ($\delta = 5.245$) and by upfield shift of Man-$\alpha$ ($\delta = 5.036$). The results, together with the monosaccharide composition analysis and $\alpha$-mannosidase digestions just described, indicated that oligosaccharide contained an additional Glc residue linked to the nonreducing terminal Man-$\alpha$ of Man$_9$GlcNAc$_2$ via $\alpha$1-3 as shown in Table III and Scheme 1.

The structural data indicate that *A. pernyi* arylphorin contains (1) simple oligomannose-type glycans from Man$_5$- to Man$_9$GlcNAc$_2$, (2) truncated Man$_2$- and Man$_3$GlcNAc$_2$ oligosaccharides with or without $\alpha$1-6Fuc to the proximal core GlcNAc, and (3) monoglucosylated Man$_9$GlcNAc$_2$. Structural formulae proposed for oligosaccharides are summarized in Table IV.

**Oligosaccharides of nonarylphorin hemolymph proteins**

The oligosaccharide profile of nonarylphorin hemolymph proteins was compared with that of arylphorin to see if the glucosylated Man$_9$GlcNAc$_2$ was common to other proteins of *A. pernyi* hemolymph. There was no detectable Glc$_1$Man$_9$GlcNAc$_2$ oligosaccharide (Figure 6A). Oligosaccharides were extracted and 2-AB-labeled from arylphorin (SP2) of another silkworm, *B. mori*, to see whether the monoglucosylated Man$_9$GlcNAc$_2$ was also present on this protein. SP2 had significant amounts of the peak comigrating with the Glc$_1$Man$_9$GlcNAc$_2$ and with oligomannose structures (Figure 6B). These data strongly suggest that arylphorin from both *B. mori* and *A. pernyi* contained the identical Glc$_1$Man$_9$GlcNAc$_2$ structure and therefore that this glycan might have a specific function in the storage proteins.
PNGase F digestion of native arylphorin

Non-denatured arylphorin was subjected to PNGase F treatment, and the recovered oligosaccharides were 2-AB-labeled. The NP HPLC profile of labeled oligosaccharides was compared with the one from denatured arylphorin to determine which oligosaccharides were exposed on the surface of the protein. The glycan pool from the native protein was composed of mainly Man5GlcNAc2, Man6GlcNAc2, and small amounts of Man7±9GlcNAc2, but only a little Glc1Man9GlcNAc2 (Figure 7A). In contrast, denatured arylphorin treated with PNGase F reproduced a glycan pattern with monoglucosylated Man9GlcNAc2 (Figure 7B). The profile was not changed significantly with additional enzyme treatment (data not shown). It demonstrated that the \( N \)-glycosidic linkage on the glucosylated oligosaccharide was almost inaccessible to PNGase F and might be protected by protein itself. These data suggest that the glucosylated glycoform of arylphorin may be folded differently from the other glycoforms.

**Discussion**

Insect arylphorins, which are storage proteins, are known to be glycosylated (Haunerland, 1996). Arylphorins have 2–10% (w/w) carbohydrate content consisting of mainly GlcNAc and Man in a 1:4–5 molar ratio (Telfer et al., 1983; Tojo and Yoshiga, 1993; Ancsin and Wyatt, 1996). The *A. pernyi* arylphorin has 4.8% (w/w) of carbohydrates (Table I). Recently, it was found that the methionine-rich storage protein of common cutworm also had moderate amounts of carbohydrates and that the molar ratio of GlcNAc to Man was 1:1–2 (Tojo and Yoshiga, 1993), suggesting that arylphorins contained both truncated and full-length oligomannose-type glycans. *A. pernyi* arylphorin had typical oligomannose type \( N \)-glycans ranging from Man5- to Man10GlcNAc2 as well as truncated glycans, some of which were fucosylated. These data are in agreement with the analysis of other insect storage proteins, in particular larval serum protein (LSP2) of *D. melanogaster* (Williams et al., 1991). Also, using 250 MHz NMR spectroscopy, Ryan et al. (1985) demonstrated that arylphorin from *Manduca sexta* contained Man10GlcNAc2.

Here the calculated numbers of each monosaccharide from oligosaccharide sequencing (Glc:Man:GlcNAc: Fuc = 1:12:4:0.4) were quite well matched with the result of sugar composition analysis. Four moles of GlcNAc in oligomannose oligosaccharides suggests two \( N \)-glycosylation sites in arylphorin, and this is consistent
with two conserved \(N\)-glycosylation sites in the \(N\)-terminal region of arylphorins of \(A.\) \(perryni\), \(B.\) \(mori\), \(M.\) \(sexta\), and \(Spodoptera\) \(litura\) (Fujii \textit{et al}., 1989; Wu \textit{et al}., 1996; Zheng \textit{et al}., 2000).

The presence of Man\(_3\)GlcNAc\(_2\) may indicate that the \(A.\) \(perryni\) glycan processing pathway is, in this respect, the same as that of higher animals in that complete mannose trimming is achieved (März \textit{et al}., 1995). Recently, \(\beta(1,2)\)-\(N\)-acetylglucosaminyltransferase I (Velardo \textit{et al}., 1993), \(\beta(1,2)\)-\(N\)-acetylglucosaminyltransferase II (Altmann \textit{et al}., 1993), and \(\beta(1,4)\)-galactosaminyltransferase (Van Die \textit{et al}., 1996) activities have been detected in insect cells. In addition, \(D.\) \(melanogaster\) UDP-GlcNAc\(\alpha\)-3-D-mannoside \(\beta(1,2)\)-\(N\)-acetylglucosaminyltransferase I (Sakar and Schachter, 2001) was cloned and expressed, which suggested that insects have the potential to make glycans other than oligomannose type. However, in the storage proteins we found only oligomannose-type oligosaccharides. There is no report to date describing glycans carrying Gal or GlcNAc residues from an insect storage protein. Fucosylation and processing of Man\(_3\)GlcNAc\(_2\) to Man\(_1\) GlcNAc\(_2\) require prior addition of an outer-arm GlcNAc by GlcNAc transferase I and subsequent removal of the GlcNAc (Marchal \textit{et al}., 1999). Although we found no

<table>
<thead>
<tr>
<th>Peak</th>
<th>Oligosaccharide structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Man((\alpha)-6)Man((\beta)-1(\alpha)-4)GlcNAc((\beta)-1(\alpha)-4)GlcNAc</td>
</tr>
<tr>
<td>2</td>
<td>Man((\alpha)-6)Man((\beta)-1(\alpha)-4)GlcNAc((\beta)-1(\alpha)-4)GlcNAcFuc((\alpha)-1(\alpha))</td>
</tr>
<tr>
<td>3</td>
<td>Man((\alpha)-6)Man((\beta)-1(\alpha)-4)GlcNAc((\beta)-1(\alpha)-4)GlcNAc</td>
</tr>
<tr>
<td>4</td>
<td>Man((\alpha)-6)Man((\beta)-1(\alpha)-4)GlcNAc((\beta)-1(\alpha)-4)GlcNAcFuc((\alpha)-1(\alpha))</td>
</tr>
<tr>
<td>5</td>
<td>Man((\alpha)-6)Man((\alpha)-3)Man((\beta)-1(\alpha)-4)GlcNAc((\beta)-1(\alpha)-4)GlcNAc</td>
</tr>
<tr>
<td>6 &amp; 7</td>
<td>({)Man((\alpha)-6)Man((\alpha)-3)Man((\beta)-1(\alpha)-4)GlcNAc((\beta)-1(\alpha)-4)GlcNAc</td>
</tr>
<tr>
<td>8</td>
<td>2Man((\alpha)-6)Man((\alpha)-3)Man((\beta)-1(\alpha)-4)GlcNAc((\beta)-1(\alpha)-4)GlcNAc</td>
</tr>
<tr>
<td>9</td>
<td>3Man((\alpha)-6)Man((\alpha)-3)Man((\beta)-1(\alpha)-4)GlcNAc((\beta)-1(\alpha)-4)GlcNAc</td>
</tr>
<tr>
<td>10</td>
<td>Man((\alpha)-6)Man((\alpha)-3)Man((\alpha)-6)Man((\alpha)-1(\alpha)-2)Man((\alpha)-1(\alpha)-3)Man((\beta)-1(\alpha)-4)GlcNAc((\beta)-1(\alpha)-4)GlcNAc</td>
</tr>
<tr>
<td>11</td>
<td>Glic((\alpha)-1(\alpha)-3)Man((\alpha)-2)Man((\alpha)-1(\alpha)-2)Man((\alpha)-1(\alpha)-3)Man((\beta)-1(\alpha)-4)GlcNAc((\beta)-1(\alpha)-4)GlcNAc</td>
</tr>
</tbody>
</table>

The oligosaccharide structures of peaks in Figure 1 are illustrated.
evidence of GlcNAc-terminated chains, the fucosylation on the truncated oligosaccharides suggested GlcNAc transferase I activity in this silk moth, A. pernyi. No difucosylated glycans, such as those on honeybee venom phospholipase A2 (Staudacher et al., 1991), were detected in this study. All glycans studied were neutral. The most abundant peak was monoglucosylated Man₉GlcNAc₂, and this assignment was consistent with the truncated oligosaccharides suggested GlcNAc transferase-GlcNAc₂ makes up some 20–40% of the total glycans in α-mannosidase extracted from jackbean that is a plant storage tissue (Kimura et al., 1999). Glc₁Man₉GlcNAc₂ was not present in other hemolymph proteins (Figure 6A) but was present in arylphorin (SP2) of another silkworm, B. mori (Figure 6B). These data strongly suggest that Glc₁Man₉GlcNAc₂ might play a specific role in storage proteins.

We could not describe the exact function of the oligosaccharides on storage protein, arylphorin, of A. pernyi. It is, however, worth pointing out that the Glc₁Man₉GlcNAc₂ structure was not released efficiently by PNGase F (Figure 7A), although a significant amount of the glycan was reproducibly released from denatured arylphorin. If the oligosaccharide is exposed on the surface of arylphorin, we may expect to find Glc₁Man₉GlcNAc₂, since the lepidopteran insect cell line has class I and II α-mannosidase, which can convert Man₉GlcNAc₂ to Man₂GlcNAc₂ in the Golgi apparatus (Kawar et al., 2000, 2001). Considering that there are no detectable intermediate glucosylated oligomannosides, arylphorin, which is a hexameric protein, may be assembled with the Glc₁Man₉GlcNAc₂ oligosaccharide in a location that restricts further processing of the glycan in the ER and Golgi. This speculation is supported by the evidence that the carbohydrate moity of and deglucosylation by glucosidase I and II (Bergeron et al., 1994; Helenius, 1994) and reglucosylation by UDP-Glc:glycoprotein glucosyltransferase in the ER (Parodi, 2000). The reaction is, however, transient, and the oligosaccharides on protein are finally deglucosylated and further processed after acquisition of their native tertiary and quaternary structures. Monoglucosylated oligosaccharide on mature glycoproteins was found in Leishmania spp. promastigote (Funk et al., 1997). The glycan was one of two predominant species and a common feature of the cell surface glycoproteins. The authors concluded that the Glc₁Man₉GlcNAc₂ structure was present on mature glycoproteins when organisms either lack glucosidase II enzymes or have a low level of glucosidase II activity. This may explain the presence of glucosylated glycans; however, the oligosaccharide is presented only on arylphorins, not on overall hemolymph proteins of A. pernyi (Figure 6), suggesting that if the explanation given by Funk et al. (1997) is correct, the monoglucosylated glycans on arylphorin are somewhat inaccessible. Characterization of putative ER and Golgi processing enzymes from A. pernyi’s fat body, which is a biosynthetic place of arylphorin, would be necessary to address how this oligosaccharide comes to remain on arylphorins.

Even though it is very rare to find Glc₁Man₉GlcNAc₂ in native mature glycoproteins, it has been described that normal storage proteins or tissues have the oligosaccharide as a major N-glycan. First of all, the standard Glc₁Man₉GlcNAc₂ used in this study was purified from hen albumen, in which the structure was determined from immunoglobulins of egg yolk of hen or Japanese quail (Ohta et al., 1991; Matsuura et al., 1993). A major glycoprotein present in the egg jelly coat of a starfish, Asterias amurensis, and vitellogenic substances from the ovary of A. rubens (L.) have the same glycan structure (Endo et al., 1987; De Waard et al., 1987). Moreover, Glc₁Man₉GlcNAc₂ makes up some 20–40% of the total glycans in α-mannosidase extracted from jackbean that is a plant storage tissue (Kimura et al., 1999). Glc₁Man₉GlcNAc₂ was not present in other hemolymph proteins (Figure 6A) but was present in arylphorin (SP2) of another silkworm, B. mori (Figure 6B). These data strongly suggest that Glc₁Man₉GlcNAc₂ might play a specific role in storage proteins.
housefly. Musca domestica, arylphorin was not involved in its uptake by the fat bodies and integument (Marinotti et al., 1988). In addition, removal of glucose-containing oligomannose-type oligosaccharides resulted in difficulty with the assembly of enzyme subunits to form the active quaternary structure for jackbean α-mannosidase. The authors concluded that the oligomannose type glycan of α-mannosidase might be necessary for the assembly of protein subunits rather than folding of the protein molecule (Kimura et al., 1999). Taken together with these observations, it suggests that A. pernyi arylphorin requires Glc1Man9GlcNAc2 oligosaccharide for assembly of the hexamer as well as protein folding.

Materials and methods

Materials
Most of the materials were obtained from Sigma (St. Louis, MO), except 50% (w/v) NaOH and HPLC solvents, which were purchased from Fisher (Pittsburgh, PA). All exoglycosidases, PNGase F, and standards were purchased from Glyko (Novato, CA), except Glc1Man9GlcNAc2, which was from Sigma.

Animals
The Chinese oak silkworms, A. pernyi, were raised throughout all the stages in the oak tree field established in the National Institute of Agricultural Science and Technology, RDA, Korea. The A. pernyi and B. mori were reared on fresh mulberry leaves.

Hemolymph collection
The abdominal legs of the corresponding silmoth larvae were injured with sharpened scissors and hemolymph bled from the wound. It was directly collected into a precooled test tube with a few crystals of 1-phenyl-2-thiourea to inhibit phenol oxidase. The hemolymph was centrifuged at 5000 g for 15 min to remove hemocytes and other debris. The hemolymph prepared was kept at −70°C until analysis.

Purification of storage proteins
Storage proteins were separated by simple native-gel electrophoresis described previously (Shimada et al., 1987) or gel filtration with Superdex 200 pg (16 × 60 cm, Pharmacia, Upsala, Sweden). The solvent was 50 mM sodium phosphate, pH 7.5, containing 0.1 M sodium chloride. Flow rate was 1.0 ml/min, and the eluting proteins were monitored at 280 nm. Storage proteins were eluted just after the void volume of the column.

Monosaccharide composition analysis
Each 0.1 mg of the protein was subjected to different conditions of acid hydrolysis. For analysis of amino or neutral sugars, the protein was hydrolyzed in 6 N HCl or 2 M trifluoroacetic acid at 100°C for 4 h, respectively. The hydrolyzate was evaporated to dryness using a SpeedVac (Savant Instruments, Holbrook, NY), resuspended in distilled water, and injected into a Bio-LC DX-300 (Dionex, Sunnyvale, CA) HPAEC with pulsed amperometric detection. Sixteen millimolar NaOH at a flow rate of 1 ml/min was used to separate monosaccharides on a CarboPac PA-1 column (Dionex, 4 × 250 mm) and a guard column (4 × 50 mm).

Extraction and fluorescence labeling of oligosaccharides
Approximately 5 μg of a protein was separated on a 10% SDS–PAGE. Protein bands visualized by Coomassie Brilliant Blue staining were excised from the gel. Proteins were alkylated, digested with PNGase F, and extracted as described previously (Küster et al., 1997). All extracts were combined and dried in a SpeedVac. Automated glycan preparation by hydrazinolysis with arylphorin that was isolated by gel-filtration chromatography was performed using a GlycoPrep1000 (Oxford GlycoSciences). The dried samples were labeled with 2-AB using the Glyko Signal labeling kit (Bigge et al., 1995).

Exoglycosidase sequencing with NP HPLC
The 2-AB-labeled glycans were dried and incubated with several exoglycosidases at 37°C for 16–24 h in 100 mM sodium acetate buffer, pH 5.0, containing 2 mM Zn2+. Conditions for individual enzymes in the array were as follows: jack bean α-mannosidase, 100 U/ml; A. saitoi α-mannosidase, 2 U/ml; X. manihotis α-mannosidase, 0.2 U/ml; bovine kidney α-fucosidase, 0.2 U/ml; almond meal α-fucosidase, 0.2 U/ml; A. ureafaciens sialidase, 1 U/ml; jack bean β-N-acetylhexosaminidase, 10 U/ml; bovine testes β-galactosidase, 2 U/ml. Each reaction mixture was cleaned up by filtration through a cellulose nitrate centrifugal filter with 0.45 μm pore size. The filter was prequileubricated with each 20 μl water for 30 min and 5% (v/v) acetonitrile for 5 min. After brief centrifugation, the mixture was applied on the top of the filter followed by incubation for 30 min. The filter was washed twice with 5% acetonitrile. The filtrates were combined, dried, and analyzed by HPLC.

The 2-AB-labeled sugars were separated on a 4.6 × 250 mm GlycoSep N column (Glyko, Milford, MA) using two Waters 510 pumps, 717 autosampler, and 474 fluorescence detector. The initial conditions were 20% solvent A, 50 mM ammonium formate, pH 4.4, and 80% solvent B, acetonitrile at a flow rate of 0.4 ml/min, followed by a linear gradient of 20–58% A over 152 min, followed by 58–100% A over the next 3 min. The total run time was 180 min, and column temperature was 30°C.

MALDI MS
All mass spectra were acquired on Micromass (Altrincham, Cheshire, UK) AutoSpec-FPDQ magnetic sector instrument fitted with a pulsed nitrogen laser (337 nm) and an array detector. Oligosaccharide pool or each peak of total 2-AB-labeled glycan on the HPLC chromatogram was collected to determine mass. Prior to MALDI MS, collections were dried in a SpeedVac and resuspended in 20 μl distilled water. Samples were applied to Ziptip C18 (Millipore, Bedford, MA) equilibrated with acetonitrile. The Ziptip was washed with water, and labeled glycans were eluted with twice 20 μl of 50% (v/v) acetonitrile.
The eluate was dried in vacuo. Sugar samples were loaded onto the mass spectrometer target in 1 μl water, mixed with 1 μl dihydroxybenzoic acid (10 mg/ml in acetonitrile), and allowed to dry. The array detector was set to the high-resolution position, and the mass range was set to be appropriate to the sample being examined. For data acquisition, the laser was operated at full power, and the laser beam was moved manually over the sample to compensate for sample depletion under the laser beam.

**Glycan preparation for 1H NMR analysis**

N-linked oligosaccharides were released from hemolymph proteins (10–20 mg) by treatment with PNGase F according to the standard deglycosylation protocol (Tarentino et al., 1989). Oligosaccharides were extracted by selective organic procedures (Verostek et al., 2000). Briefly, the pH of the mixture was adjusted to around pH 5.5 using 1 N phosphoric acid after the deglycosylation. Proteins and oligosaccharides were precipitated by adding 4 volumes of ice-cold acetone followed by incubation at −20°C for 30 min. The precipitate was collected by centrifugation at 14,000×g for 10 min, and the pellet was extracted with 0.25 ml of ice-cold 60% (v/v) aqueous methanol by sonication and 1 h incubation at −20°C. A second 60% methanol extract of the pellet was combined with the first and dried. The dried material was dissolved in a minimal volume of water. Any residual SDS and peptides were removed from the oligosaccharide pools by passage through a Sep-Pak C18 cartridge (Waters). The run-through fraction containing oligosaccharides was concentrated with a SpeedVac or injected directly into HPAEC with CarboPac PA1 (9 × 250 mm, Dionex). The oligosaccharides were eluted with a linear gradient of 0–0.2 M sodium acetate in 0.25 M NaOH in 30 min at a flow rate of 4.5 ml/min, and the eluant was monitored with pulsed amperometric detection. Oligosaccharide fractions were collected and desalted by passage through cation-exchange resin, AG50X-12 (H+ form, Bio-Rad). The glycan pools were dried in vacuo.

**1H NMR spectrometry**

Oligosaccharides were exchanged several times by rotary evaporation from 99.9% D2O to 0.25 ml of ice-cold 60% (v/v) aqueous methanol by sonication and 1 h incubation at −20°C. A second 60% methanol extract of the pellet was combined with the first and dried. The dried material was dissolved in a minimal volume of water. Any residual SDS and peptides were removed from the oligosaccharide pools by passage through a Sep-Pak C18 cartridge (Waters). The run-through fraction containing oligosaccharides was concentrated with a SpeedVac or injected directly into HPAEC with CarboPac PA1 (9 × 250 mm, Dionex). The oligosaccharides were eluted with a linear gradient of 0–0.2 M sodium acetate in 0.25 M NaOH in 30 min at a flow rate of 4.5 ml/min, and the eluant was monitored with pulsed amperometric detection. Oligosaccharide fractions were collected and desalted by passage through cation-exchange resin, AG50X-12 (H+ form, Bio-Rad). The glycan pools were dried in vacuo.

**N-terminal amino acid sequencing**

Amino acid sequences were determined using a 491A protein sequencer (Perkin-Elmer, Foster City, CA). The N-terminal sequence obtained was used for protein identification by a BLAST search of NCBI.

**PNGase F digestion of native protein**

Ten micrograms each of purified arylphorin was subjected under nondenaturing digestion for 18 h with 10 mU PNGase F in 10 μl 50 mM sodium phosphate, pH 7.5. An extra aliquot of PNGase F was added to complete the reaction. Released oligosaccharides were recovered, 2-AB-labeled, and separated on NP HPLC as described. Oligosaccharides from denatured arylphorin by PNGase F were used as control.

**Acknowledgments**

This work was financially supported by the BMC program of Korea Basic Science Institute, and by the National Research Laboratory Program, Republic of Korea. We thank C. Colominas for early experiments and T.D. Butters for helpful discussion.

**Abbreviations**

2-AB, 2-aminobenzamide; ER, endoplasmic reticulum; HPAEC, high-pH anion exchange chromatography; HPLC, high-performance liquid chromatography; NP, normal phase; MALDI MS, matrix-assisted laser desorption/ionization; MS, mass spectrometry; NMR, nuclear magnetic resonance; PNGase F, peptide-N-glycosidase F; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis.

**References**


Endo, T., Hoshi, M., Endo, S., Arata, Y., and Kobata, A. (1987) Structures of the sugar chains of a major glycoprotein present in the...
N-Glycan structure of lepidopteran arylphorin