Glycobiology vol. 14 no. 9 pp. 767–774, 2004
doi:10.1093/glycob/cwh105
Advance Access publication on June 2, 2004

Structural characterization of the *N*-glycan moiety and site of glycosylation in vitellogenin from the decapod crustacean *Cherax quadricarinatus*

Isam Khalaila[1,2,3], Jasna Peter-Katalinic[3], Clarence Tsang[4], Catherine M. Radcliffe[4], Eliahu D. Aflalo[5], David J. Harvey[4], Raymond A. Dwek[4], Pauline M. Rudd[4], and Amir Sagi[5]

[3]Institute for Medical Physics and Biophysics University of Münster, Robert-Koch-Str. 31 D-48149, Münster, Germany; [4]Oxford Glycobiology Institute, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK; [5]Department of Life Sciences and the Institute for Applied Biosciences, Ben Gurion University, P.O. Box 653, Beer Sheva, Israel

Received on February 17, 2004; revised on May 26, 2004; accepted on May 27, 2004

Glycosylation is of importance for the structure and function of proteins. In the case of vitellin (Vt), a ubiquitous protein accumulated into granules as the main yolk protein constituent of oocytes during oogenesis, glycosylation could be of importantance for the folding, processing and transport of the protein to the yolk and also provides a source of carbohydrate during embryogenesis. Vt from the crayfish *Cherax quadricarinatus* is synthesized as a precursor protein, vitellogenin (Vg), in the hepatopancreas, transferred to the hemolymph, and mobilized into the growing oocyte via receptor-mediated endocytosis. The gene sequence of *C. quadricarinatus* shows a 2584-amino-acid protein with 10 putative glycosylation sites. In this study a combined approach of lectin immunoblotting, in-gel deglycosylation, and mass spectrometry was used to identify the glycosylation sites and probe the structure of the glycan moieties using *C. quadricarinatus* Vg as a model system. Three of the consensus sites for *N*-glycosylation—namely, Asn^{152}, Asn^{160} and Asn^{2493}—were glycosylated with the high-mannose glycans, $Man_{5-9}GlcNAc_2$, and the glucose-capped oligosaccharide $Glc_1Man_9GlcNAc_2$.

Key words: *N*-glycan/oligosaccharides/site of glycosylation/vitellogenin/vitellin

Introduction

Vitellin (Vt), is the major yolk protein in crustaceans, as in all oviparous animals. Vt is accumulated in yolk bodies in the oocytes during vitellogenesis, a process that has been well characterized in oviparous vertebrates and insects and lately in crustaceans. Vt later serves the developing embryo as an important source of proteins, lipids, and carbohydrates. As early as 1967, crustacean Vt was shown to be a high-density glycolipoprotein (Wallace *et al*., 1967) containing a high percentage of lipids, invariably conjugated to carotenoid pigments. Other studies have shown that crustacean Vt is a high-density lipoprotein (HDL) complex (Meusy and Payen, 1988).

In the decapod *Cherax quadricarinatus*, vitellogenin (Vg), the precursor of Vt, is produced in the hepatopancreas of vitellogenic females (Abdu *et al*., 2002). Vg is a high-molecular-weight lipoglycocarotenoprotein composed of several subunits with similar biochemical and immunological characteristics to Vt (Chang *et al*., 1994; Derelle *et al*., 1986; Kerr, 1969; Lee and Watson, 1994; Meusy, 1980). The partial amino acid sequencing of Vt subunits has allowed degenerate oligonucleotide primers to be produced and then used for complete cDNA isolation and sequencing of *C. quadricarinatus* Vg. The cDNA sequence indicates that Vg is composed of 2584 amino acids (Abdu *et al*., 2002). Furthermore, northern analysis of the Vg cDNA indicated a single 8000-nucleotide mRNA band present only in the hepatopancreas of vitellogenic females, suggesting that in *C. quadricarinatus*, Vg is produced in the hepatopancreas and transported to the maturing oocyte via the hemolymph (Abdu *et al*., 2002). Several proteins, with molecular weights of 86, 177, and 196 kDa, have been specifically found in the hemolymph of secondary vitellogenic *C. quadricarinatus* females (Yehezkel *et al*., 2000), and it has been suggested that these proteins are recognized and transported to the growing oocyte through receptor-mediated endocytosis (Warrier and Subramoniam, 2002), a process known to be related to the glycan content of the mobilized protein. Glycosylation has been shown to be involved in a number of functional roles, including protection of proteins from proteases (Marinaro *et al*., 2000), assistance in protein folding, targeting of proteins to specific locations, and regulation of protein activity (Helenius and Aebi, 2001). Thus vitellogenesis in *C. quadricarinatus* is a unique model for studying glycans and their roles in protein folding, processing, and transport, because the target protein is produced at an extra-ovarian site, released to the hemolymph, and taken up by the oocyte through receptor-mediated endocytosis.

The protein and lipid components of Vt have been characterized in both invertebrate and vertebrate systems (de Chaffoy de Courcelles and Kondo, 1980; Fyffe and O'Connor, 1974; Ohlendorf *et al*., 1977; Raikhel and Dhadialla, 1992; Tirumalai and Subramoniam, 1992, 2001), but very little information is available on the carbohydrate components of this major yolk protein, despite the structural and functional importance of these units. Many studies have been performed on insect glycoprotein due to the

[1]To whom correspondence should be addressed; e-mail: isam.khalaila@epfl.ch
[2]Present address: Ecole Polytechnique Fédérale de Lausanne, EPFL-BCH-LCOM, CH-1015 Lausanne, Switzerland

wide use of insect cell systems to produce glycosylated recombinant proteins. Although these studies suggested the presence of an *N*-glycosylation pathway in insect cells similar to that seen in mammalian cells, the most frequent structure of insect *N*-glycans is the paucimannosidic glycan, $Man_3GlcNAc_2(\pm Fuc)$, not found in mammalian cells (Jarvis and Finn, 1995). Structural analysis of the glycan moiety of insect Vt revealed high-mannose oligosaccharides (Raikhel and Dhadialla, 1992). Using radiolabeling and chromatography, lipovitellin from the crustacean *Emerita asiatica* was shown to contain five different *O*-linked oligosaccharides and four different *N*-linked oligosaccharides (Tirumalai and Subramoniam, 2001), but no structural details of these glycan moieties are available. In contrast, the sequence of *C. quadricarinatus* Vg shows 10 putative *N*-glycosylation sites with the consensus sequence Asn-X-Ser/Thr (Abdu *et al*., 2002). In this study, the sites of glycosylation and glycan structure have been elucidated in Vg from the hemolymph and Vt from the eggs of *C. quadricarinatus* using a combined approach based on lectin immunoblotting, enzymatic in-gel deglycosylation, exoglycosidase sequencing, and chromatographic and mass spectrometric analysis.

Results

Lectin blotting of Vg and Vt proteins

Five and three major proteins could be found in the hemolymphatic and ovarian HDL fraction respectively, with apparent molecular weights of 86, 96, 177, 196, and 208 kDa in the hemolymph and 75, 86, and 96 kDa in the ovary (Figure 1A). Excluding the 75-kDa protein, all of the proteins were shown to exhibit a strong interaction with *Galanthus nivallis* agglutinin (GNA), which recognizes *N*-glycans and/or *O*-linked mannoses (Figure 1B), but no interactions were observed with the other lectins used, *Sambucus nigra* agglutinin (SNA), *Maackia amurensis* agglutinin (MAA), peanut (*Arachis hypogaea*) agglutinin (PNA), and *Dature stramonium* agglutinin (DSA), which recognize specificly sialic acid linked α(2-6) to galactose, sialic acid linked α(2-3) to galactose, galactose β(1-3) N-acetylgalactosamine and galactose β(1-4)GlcNAc, respectively, (data not shown).

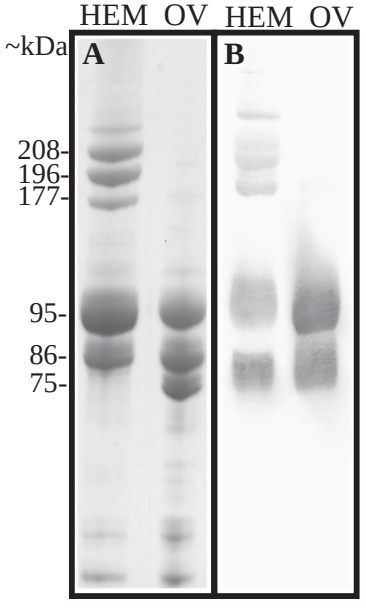

Fig. 1. Ovarian (OV) and hemolymph (HEM) HDL from *C. quadricarinatus* separated on 7% SDS–PAGE. (**A**) Coomasie blue–stained SDS–PAGE; (**B**) blot probed with GNA, specific for high-mannose glycan chains.

Mass spectrometry analysis of Vg proteins and glycosylation site

Before the carbohydrate moiety could be obtained, the identity of the subunits of the HDL fractions 86, 177, and 196 kDa from the hemolymph and 86 kDa from the ovary (Figure 1A) were confirmed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) of the tryptic digest as deduced from the *C. quadricarinatus* Vg cDNA (accession number AF306784). All the proteins were found to correspond to specific regions of the Vg gene. Subsequently, all the corresponding regions included potential consensus *N*-glycosylation sites. The LC-MS/MS analysis of the 86-kDa protein revealed 9.7% coverage of the 2584 amino acids deduced from Vg cDNA. Coverage of 18.3% and 7.6% have been obtained from the LC-MS/MS analyses of 177 kDa and 196 kDa proteins, respectively.

In the LC-MS analysis of the tryptic digest of the electroeluted intact 86-kDa Vg, a series of quadruply charged molecular ions were found. The most intense peak at *m/z* 1692.46^{4+} could be assigned to the glycopeptide 140–169 amino acid residues (calculated mass = 3034.48) glycosylated at Asn^{152} and Asn^{160} with $Hex_{18}(HexNAc_4)$ moieties (calculated = 3729.18) that together give a mass of $[6763.66 + 4H]/4 = [m/z\ 1691.92]^{4+}$ (Figure 2A). The spacing between the other peaks (162/4) corresponded to hexose. Thus the quadruply charged ions at *m/z* 1651.9, 1611.4, 1570.9, 1530.4, and 1489.9 corresponded to $Hex_{17}(HexNAc_4)$, $Hex_{16}(HexNAc_4)$, $Hex_{15}(HexNAc_4)$, $Hex_{14}(HexNAc_4)$, and $Hex_{13}(HexNAc_4)$, respectively. After peptide *N*-glycosidase F (PNGase F) treatment and simultaneous digestion with trypsin and chymotrypsin, a shift of two mass units was observed in the doubly charged parent ion at *m/z* 1253.0 from the peptide containing amino acids 146–169 (calculated *m/z* 1252.1), and a shift of one or two mass units was observed in the peptide fragments bearing one or both originally glycosylated Asn residues, respectively. This shift was caused by conversion of Asn to Asp in the deglycosylation reaction (Figure 2B and Table I). In the tryptic digest of the PNGase F–treated 196-kDa Vg, a doubly charged ion at *m/z* 917.46 was observed by LC-MS corresponding to amino acids 2493–2508 (calculated *m/z* 916.94) (Table I); LC-MS/MS confirmed the sequence and the transition of Asn^{2493} to Asp (Table I).

MS of the N-glycans

Monoisotopic $[M + Na]^+$ ions at *m/z* 1257.4, 1419.5, 1581.5, 1743.5, 1905.5, and 2067.6 were detected by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) MS in the *N*-linked oligosaccharides from the 86-kDa Vg

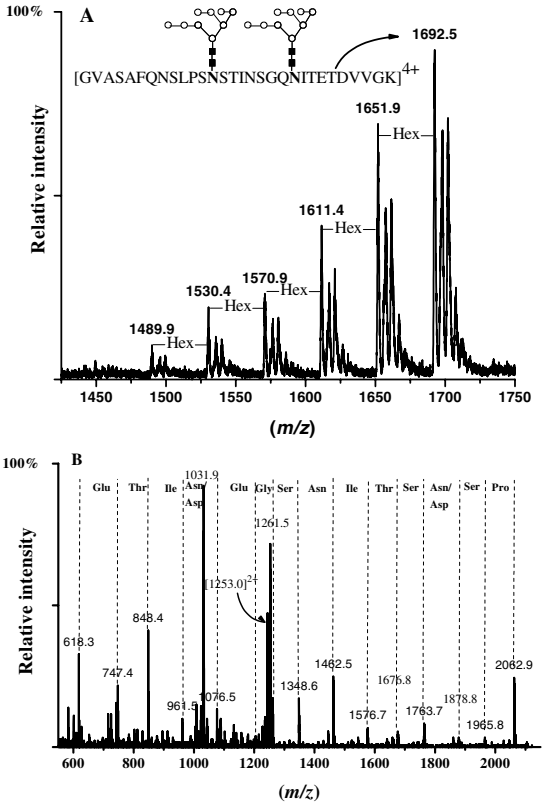

Fig. 2. LC-MS/MS for the glycosylation sites of *C. quadricarinatus* Vg. (**A**) Electrospray mass spectrum of the intact glycopeptide 140–169 amino acid residues, glycosylated at N^{152} and N^{160}, numbers above the peaks represent $[M + 4H]^{4+}$ of different number of hexoses. Hex represent the mass difference of a hexose between the peaks, adjacent peaks at higher mass than the main peaks are the sodiated adducts. (**B**) LC-MS/MS spectrum of the doubly charged ion at *m/z* 1253.0 of the glycopeptide 146–160 amino acid residues after PNGase F treatment; the transition of Asn to Asp acid is shown in the fragment profile. Molecular representations of the sugars are as indicated in the legend to Figure 4.

Table I. Assignment of molecular ions of Vg peptides before and after PNGase F treatment

Peptide	Predicted mass $[M + H]^+$	Detected mass $[M + H]^+$	Sequence
146–169	2503.2	2505.0	QNSLPS**N**^{152} **ST**INSGQ**N**^{160}**IT**ETDVVGK
2493–2508	1832.9	1833.9	**N**^{2493}**AS**SLQELVASWQEDR

Assignment of glycosylation sites of *C. quadricarinatus* Vg referring to peptides deduced from the cDNA containing *N*-glycosylation sites. The assignment was made on the basis of a 1-Da shift from Asn to Asp as the result of PNGase F deglycosylation. The glycosylation sequons are indicated in bold.

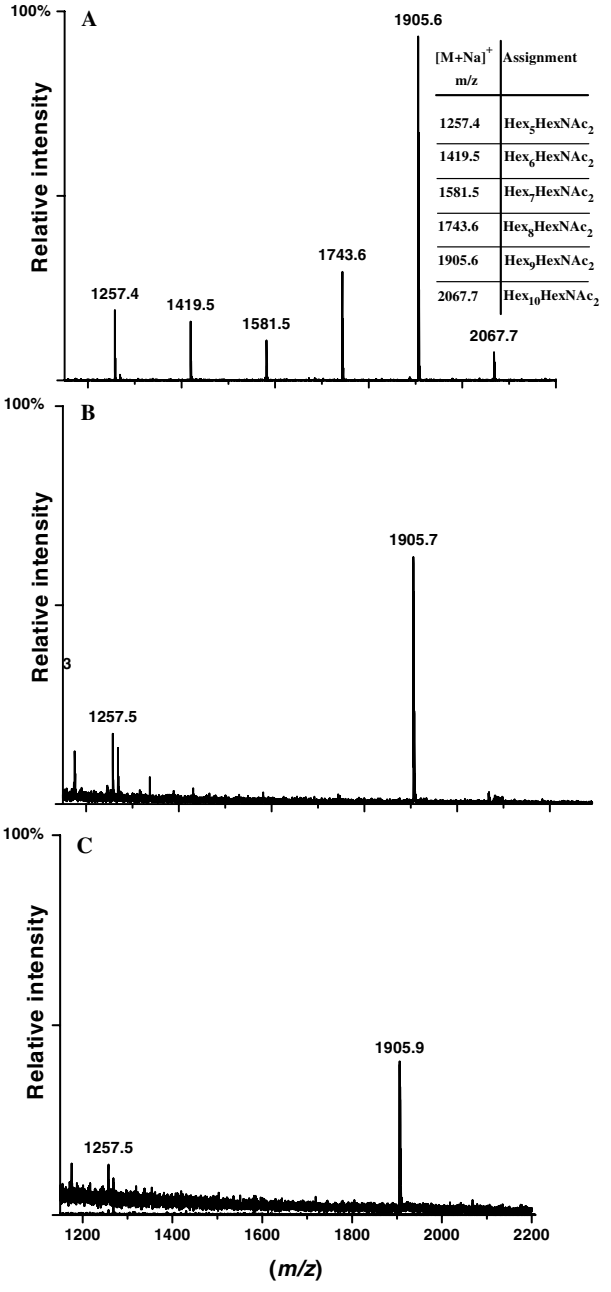

Fig. 3. MALDI mass spectra of oligosaccharides released from *C. quadricarinatus* Vg subunits by PNGase F treatment in gel. Molecular weights correspond to the monoisotopic masses of the $[M + Na]^+$ ions. Derived compositions are shown in the inset table. Hex: Hexose and HexNAc: *N*-acetylhexosamine. (**A**) Spectrum of *N*-oligosaccharides released from the 86-kDa Vg protein. (**B**) Spectrum of *N*-oligosaccharides released from 177-kDa Vg protein. (**C**) Spectrum of *N*-oligosaccharides released from 196-kDa Vg protein.

subunits from the hemolymph and the ovary treated with PNGase F (Figure 3 and Table II). The 162-Da spacing of the ions is indicative of hexose units, and thus different hexose-type structures [$Hex_{5–10}HexNAc_2$] could be deduced from the measured monoisotopic masses (Table II). The *N*-linked oligosaccharide released by PNGase F from both the ovarian and hemolymphatic 86-kDa subunit gave the highest signal intensity compared to the *N*-glycan mixture released from 177- and 196-kDa Vg subunits (Figure 3). The $[M + Na]^+$ ion at *m/z* 1905.6 [$Hex_9HexNAc_2$] gave rise to the most intense peak within the spectrum of the 86-kDa Vg subunit

Table II. Structure details of the glycans $[M + Na]^+$ of *C. quadricarinatus* 86-kDa Vg

Peak[a]	Glycan	Mass[b] Found	Mass[b] Calculated	GU[c] values	% Area[d]
1	$Man_5HexNAc_2$	1257.5	1257.4	6.20	4.7
2	$Man_6HexNAc_2$	1419.5	1419.5	7.07	14.1
3	$Man_7HexNAc_2$	1581.5	1581.5	7.94	13.9
4	$Man_8HexNAc_2$	1743.5	1743.5	8.83	10.8
5	$Man_9HexNAc_2$	1905.5	1905.6	9.51	44.2
6	$Glc_1Man_9HexNAc_2$	2067.6	2067.7	10.17	10.9

[a]Peak numbers in Figure 2.
[b]Monoisotopic mass of the $[M + Na]^+$ ion.
[c]Glucose unit.
[d]Amount of glycan expressed as a percentage of the total glycans measured by NP HPLC.

(Figure 3A and Table II). In the spectra of oligosaccharides released from the hemolymphatic 177- and 196-kDa proteins, the ion at *m/z* 1905.6 was almost completely dominant.

Exoglycosidase sequencing of oligosaccharides

The *N*-glycans released from the protein bands by in-gel digestion with PNGase F were labeled with the fluorophore 2-aminobenzamide (2AB) by reductive amination (Bigge *et al.*, 1995). The chromatogram obtained by normal-phase high-performance liquid chromatography (NP HPLC, Figure 4, upper chromatogram) shows the intact glycan pool of the 86-kDa band with a series of oligomannose sugars, $Man_{5-9}HexNAc_2$, and the monoglucosylated oligomannose, $Glc_1Man_9HexNAc_2$. Glycan structures were assigned from the glucose unit (GU) values of the intact glycan pool combined with data from digestions with the exoglycosidases jack bean α-mannosidase (JBM) and glucosidase II (glcII). JBM cleaves the nonreducing terminal mannose α 1-2, 3, and 6 linkages, and glcII cleaves the nonreducing terminal glucose α 1-3 mannose linkage. Digestion with JBM is shown in the lower chromatogram in Figure 4, in which $Man_{5-9}HexNAc_2$ is digested to $Man_1HexNAc_2$ (GU 2.65) and $Glc_1Man_9HexNAc_2$ is digested to $Glc_1Man_4HexNAc_2$ (GU 5.99). Figure 5 shows the intact glycan pool (upper chromatogram) together with the digestion with glcII (lower chromatogram) in which peak 6, $Glc_1Man_9HexNAc_2$ (GU 10.16) is completely digested to $Man_9HexNAc_2$ (GU 9.52). The peak areas of the glycans in the enzyme digests correlated with those of the undigested glycan pool. The assignments were confirmed by MALDI MS of the complete glycan pool (Table II).

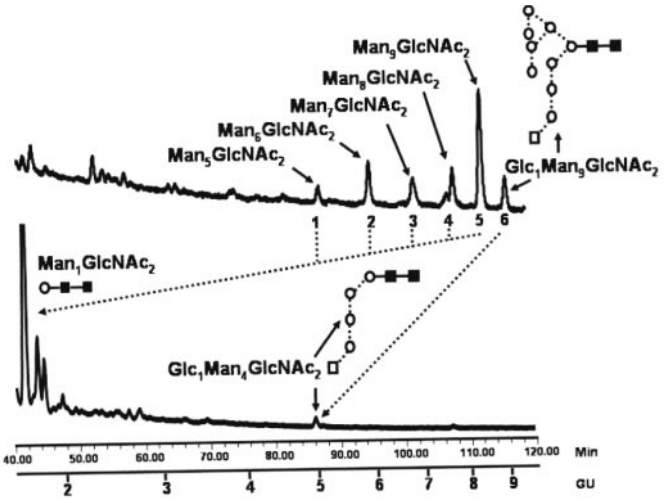

Fig. 4. NP HPLC chromatogram of the glycans released from the 86-kDa band from the SDS–PAGE gel showing the oligomannose series $Man_{5-9}GlcNAc_2$ and $Glc_1Man_9GlcNAc_2$. The lower chromatogram shows the JBM digest of the glycan pool. The products of the digestion are $Man_1GlcNAc_2$ and $Glc_1Man_4GlcNAc_2$. Molecular representations of the sugars are included. The individual monosaccharides are represented as follows: open square, glucose; open circle, mannose; dotted line, α linkage; solid line, β linkage. The linkage positions of the oligosaccharides are represented by the angle of the line linking adjacent monosaccharides. The sugar on the left is linked via C1 to ring carbons C2 (180°), C3 (225°), C4 (270°), or C5 (315°) of the sugar on the right.

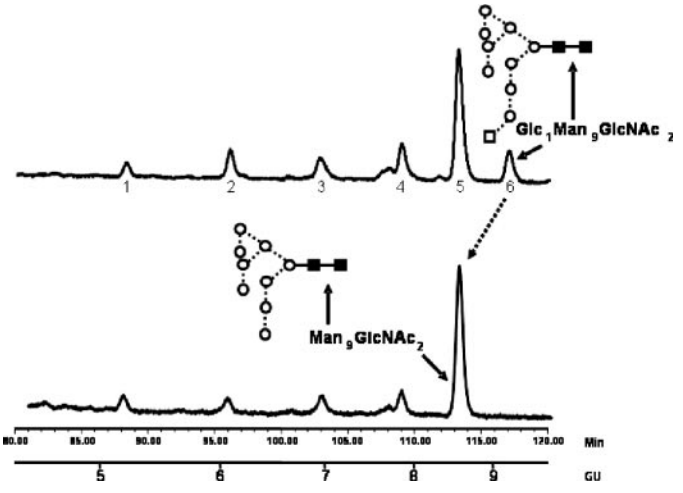

Fig. 5. NP HPLC chromatogram of the glycans released from the 86-kDa band from the SDS–PAGE gel showing the oligomannose series $Man_{5-9}GlcNAc_2$ (peaks 1–5) and $Glc_1Man_9GlcNAc_2$ (peak 6). The lower chromatogram shows the glucosidase II digest of the glycan pool where $Glc_1Man_9GlcNAc_2$ has been digested to $Man_9GlcNAc_2$. Molecular representations of the sugars are as indicated in the legend to Figure 4.

Discussion

Glycosylation is important for protein structure and function. In the case of Vt, the glycan moieties also serve as a carbohydrate source; in fact, this protein is a major source of carbohydrate, protein, and lipid for the developing embryo (Lee *et al.*, 1997). In *C. quadricarinatus* Vt originates from a precursor Vg, which is produced in the hepatopancreas (Abdu *et al.*, 2002) and secreted to the hemolymph. We have confirmed that Vg from the hemolymph is posttranslationally modified by *N*-linked

oligosaccharides. Furthermore, lectin blotting and exoglycosidase sequencing have identified the glycan moieties as those of the high-mannose type. The oligomannose moieties ranged from $Man_5GlcNAc_2$ to $Man_9GlcNAc_2$, which is similar to those of crayfish *Astacus leptodactylus* hemocyanin (Tseneklidou-Stoeter *et al*., 1995) and other glycosylated proteins from invertebrates, such as the deep-sea tube worm *Riftia pachyptila* hemoglobins (Zal *et al*., 1998), the larval serum protein of *Drosophila melanogaster* (Williams *et al*., 1991), and the storage glycoprotein arylphorin of lepidopteran insects (Kim *et al*., 2003; Ryan *et al*., 1985). Thus the *N*-glycan processing pathways seem to be conserved within Arthropoda.

In insects, glucose and mannose trimming of the *N*-linked $Glc_3Man_9GlcNAc_2$ core is possible, and the partially trimmed $Glc_1Man_9GlcNAc_2$ oligosaccharide structure has also been found in the *N*-linked glycan of *C. quadricarinatus* Vg. Furthermore, recent results show that the glucose-capped oligomannose glycans, such as $Glc_1Man_9GlcNAc_2$, are conserved in glycoproteins from many species, including hemoglobins from *R. pachyptila* (Zal *et al*., 1998), the insect storage protein arylphorin (Kim *et al*., 2003), immunoglobulins from egg yolk of hen or Japanese quail (Matsuura *et al*., 1993; Ohta *et al*., 1991), glycoprotein from the egg jelly coat of the starfish *Asterias amurensis* (Endo *et al*., 1987), ovarian vitellogenic substances of *Asterias rubens* (De Waard *et al*., 1987), and membrane protein gp63 of the parasite *Leishmania* (Funk *et al*., 1997). The latter authors have suggested that the presence of monoglucosylated oligosaccharides is due to the low activity or low level of glucosidase II enzyme (Funk *et al*., 1997). Kim *et al*. (2003) have gone further and have given additional examples for storage proteins, such as arylphorin with monoglucosylated *N*-linked oligosaccharides, mainly $Glc_1Man_9GlcNAc_2$, and have suggested a role for this unique structure in storage proteins. All of these mature glycoproteins, including *C. quadricarinatus* Vg, contain glucose-capped oligosaccharides and are large hydrophobic proteins. Hence it is possible that protein size or hydrophobicity could be linked to the unique glucose-capped oligomannose.

The deduced sequence of *C. quadricarinatus* Vg is composed of 2584 amino acid residues, from which a mass of 292 kDa for the unmodified protein is calculated. Within this molecule we confirm that of the 10 putative *N*-glycosylation sites Asn^{152} and Asn^{160} from the 86-kDa and Asn^{2493} from the 177- and 196-kDa subunits are glycosylated (Table I). The exact role of this glycosylation is not yet known. However, it is well known that glycans play a pivotal role in protein folding (Imperiali and O'Connor, 1999; Wormald and Dwek, 1999), oligomerization, quality control, sorting, and transport in the endoplasmic reticulum and Golgi apparatus (Helenius and Aebi, 2001). In protein folding, glycosylation alters the conformational preferences close to the glycosylation site, which leads to more compact conformations (Wormald and Dwek, 1999). More compact conformations may restrict the activity of glucosidase II and mannosidases, thereby producing an abundance of high-mannose and glucose-capped structures.

The lectins calreticulin and calnexin interact with the glycan moieties of substrate glycoproteins that have been trimmed by glucosidases I and II to the monoglucosylated form (Hammond *et al*., 1994; Spiro *et al*., 1996). Such interactions can lead to correctly paired disulfide bonds (Huppa and Ploegh, 1998). Thus $Glc_1Man_9GlcNAc_2$ may have a similar function in Vg. It is possible that glycosylation of Vg has an important role in folding and subunits assembly to achieve the mature protein in the hemolymph and ovary. In our model organism, Vg is secreted to the hemolymph and, because glycosylation is known to increase solubility of proteins (Jaenicke, 1991), *N*-glycans could have a significant role in keeping this large, hydrophopic protein in the hemolymph to improve its transport to the ovary. Vg uptake into the oocytes is known to be mediated by receptors (Raikhel and Dhadialla, 1992; Warrier and Subramoniam, 2002). In this respect, the glycan moiety might play a role in recognition and receptor-mediated endocytosis. On uptake into the oocytes, it might also have a role in packaging and compacting the Vt in yolk bodies. The hypotheses concerning the roles of N-glycan in Vg processing need to be clarified experimentally.

This study of the *N*-glycosylation of Vg has demonstrated the presence of $Glc_1Man_9GlcNAc_2$ for the first time in Crustacea. The present study focused on *N*-glycans of Vg; nevertheless, to complete the sequence and assembly of Vg subunits and their sites of glycosylation, future work on *O*-glycosylation is needed. Our study of Vg oligosaccharides during vitellogenesis provides a first step in understanding the cascade of assembly, recognition, and transport of Vg, its packaging in the egg and utilization during embryogenesis in crustaceans.

Materials and methods

Materials

PNGase F (EC 3.5.1.52), trypsin, and chymotrypsin, sequencing grade, were obtained from Roche Diagnostics GmbH (Mannheim, Germany). JBM (EC 3.2.1.24) was obtained from Glyko (Upper Heyford, Oxfordshire, UK). Glucosidase II was prepared in the Glycobiology Institute (Oxford); its activity, which was measured by the hydrolysis of [^{14}C]glucose-labeled $Glc_2Man_9GlcNAc_2$ (Karlsson *et al*., 1993), was determined as 5.7% min^{-1} of labeled glucose.

HDL purification

The HDL fractions from *C. quadricarinatus* secondary vitellogenic ovaries and hemolymph were isolated as described by Abdu *et al*. (2000) and Yehezkel *et al*. (2000), respectively.

Structural characterization by lectins

HDLs from *C. quadricarinatus* vitellogenic ovaries and hemolymph (5 μg and 15 μg, respectively) were separated on 7% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE), and proteins were electroblotted onto a nitrocellulose membrane before blotting with GNA, SNA, PNA, DSA, and MAA from the digoxigenin (DIG) glycan differentiation kit (Roche Diagnostics). The membrane was then incubated with anti-DIG conjugated with alkaline phosphatase, and 5-bromo-4-chloro-3-indolylphosphate was used to determine the presence of the alkaline phosphatase, with enzyme catalysis resulting in a dark color.

N-linked glycan analysis and exoglycosidase sequencing

N-linked glycans were released from five gel bands of the purified HDLs with apparent molecular weights of 208, 196, 177, 95, and 86 kDa according to the method described by Küster *et al*. (1997) as modified by Radcliffe *et al*. (2002). Briefly, the HDLs were reduced with 45 mM dithiothreitol for 10 min at 70°C, before alkylation with 100 mM iodoacetamide at room temperature. Fourteen micrograms of alkylated HDLs were separated on 6% SDS–PAGE and stained with Coomassie blue R-250. After destaining with 5% (v/v) methanol/7% (v/v) acetic acid, individual protein bands were excised from the gel. The gel pieces were washed twice with 1 ml acetonitrile followed by 1 ml 20 mM $NaHCO_3$, pH 7.0. Subsequently, the gel pieces were dehydrated with acetonitrile and dried by vacuum centrifugation (SpeedVac). *N*-glycans were released with PNGase F, labeled with 2AB by reductive amination, and separated by NP HPLC, using a low-salt buffer system (Guile *et al*., 1996). The system was calibrated using an external standard of hydrolyzed and 2AB-labeled glucose oligomers to produce a dextran ladder from which the retention times of the individual glycans were converted into GU. These GU values were compared with a database of experimental values to obtain preliminary assignments for the glycans. The HPLC profiles of the glycans from all the gel bands showed the same mannose sugars in similar proportions, but because the release from the 86-kDa band produced the most glycans, this sample was chosen for enzyme digestions. The assignments were confirmed by NP HPLC, following digestion of a 2AB-labeled glycan pool with JBM and glucosidase II and by MALDI MS of an unlabeled glycan pool.

Glycopeptide analysis: in-solution or in-gel digestion of Vg

Intact or deglycosylated in-gel Vg proteins (see earlier description) were digested with trypsin 12.5 ng/μl in 5 mM $(NH_4)HCO_3$, pH 7.5, overnight at 37°C. The resultant peptides were extracted from the gel pieces by incubation with 25 mM $(NH_4)HCO_3$; the buffer was transferred to a separate vial and exchanged with 5% formic acid and then with acetonitrile. All incubations were performed in a volume of 200 μl for 30 min using a sonication bath; the supernatants were combined and the solvents were removed using a SpeedVac.

HDLs were separated on 7% SDS–PAGE and detected by imidazole-zinc staining. After development, the gel was rinsed with water for 30 s and incubated in 0.2 M imidazole, 0.1% SDS solution for 15 min. The imidazole-SDS solution was discarded, and the gel was stained with 0.2 M zinc sulfate until the gel background became white. Detected bands corresponding to Vg subunits were excised, and the proteins were electroeluted overnight with the Bio-Rad (Hercules, CA) model 422 electroeluter using 25 mM Tris buffer containing192 mM glycine and 0.1% SDS. The electroeluted protein was then precipitated at 4°C overnight with 20% trichloroacetic acid and centrifuged for 15 min at 14,000 rpm, and the trichloroacetic acid supernatant was discarded. The pellet was washed with ice-cooled acetone, centrifuged, and dried. The protein was resuspended in 30 μl 20 ng/μl trypsin solution in 25 mM $(NH_4)HCO_3$, pH 7.5, containing 10% acetonitrile. The solution was incubated overnight at 37°C with continuous shaking. At the end of the trypsin digestion the samples were dried in a SpeedVac. Pellets from in-gel or in-solution digestion were twice resuspended in pure water and dried to reduce the $(NH_4)HCO_3$ concentration. Before LC-MS analysis, the tryptic digests were dissolved in 10 μl 5% acetonitrile containing 0.05% formic acid.

MALDI-TOF MS

Positive ion reflection MALDI-TOF mass spectra were acquired using Bruker Reflex III instrument (Bruker Daltonik, Bremen, Germany) equipped with delayed extraction. For exoglycosidase sequencing, spectra were recorded with a Micromass TofSpec 2E reflectron-TOF mass spectrometer (Waters-Micromass, Manchester, UK) fitted with delayed extraction and a nitrogen laser (337 nm). The acceleration voltage was 20 kV, the pulse voltage was 3200 V, and the delay for the delayed extraction ion source was 500 ns. Aqueous solution (0.5 μl) of the sample was added to the matrix solution (0.5 μl of a solution of 2,5-dihydroxybenzoic acid in acetonitrile [10 mg/ml]) on a stainless steel target plate and dried at room temperature before recrystallization from ethanol.

Electrospray ionization MS/MS and LC-MS/MS analysis

Conventional and tandem MS were performed on an orthogonal hybrid quadrupole time-of-flight mass spectrometer (Q-TOF, Waters-Micromass) fitted with a Micromass Z-spray ion source. Tandem MS was performed by collision-induced dissociation at low energy, using argon as a collision gas. The collision energy was adjusted to that appropriate to the mass of the ions bring fragmented, typically between 25 and 45 eV. Data acquired by the Q-TOF mass spectrometer were processed with a MassLynx data system. LC-MS and LC-MS/MS were used to analyze the peptides and glycopeptides of Vg subunits obtained after tryptic digests. An LC-Packings Ultimate nano-HPLC system equipped with the LC-Packings, Dionex PepMap C-18 column (75 μm internal diameter, 15 cm length) was interfaced with the Q-TOF mass spectrometer. Chromatographic separations were achieved using a linear gradient elution of 5–42.5% acetonitrile (0.04% formic acid) over 60 min at 200 nl/min. All buffers were prepared from HPLC-grade solutions.

Acknowledgments

We thank Dr. Claire S. Allardyce for editing the manuscript. This study was supported in part by grants from DFG (KE 206/17-1), BSF (2000116), and a Minerva Stiftung Postdoctoral Research Fellowship to I.K. The TofSpec mass spectrometer was purchased with a grant from the Biotechnology and Biological Sciences Research Council.

Abbreviations

2AB, 2-aminobenzamide; DIG, digoxigenin; DSA, *Datura stramonium* agglutinin; GNA, *Galanthus nevallis* agglutinin; GU, glucose unit; HDL, high-density lipoprotein;

HPLC, high-performance liquid chromatography; JBM, jack bean α-mannosidase; LC, liquid chromatography; MAA, *Maackia amurensis* agglutinin; MALDI, matrix-assisted laser desorption/ionization; MS, mass spectrometry; NP, normal phase; PNA, peanut (*Arachis hypogaea*) agglutinin; Q-TOF, quadrupole time-of-flight; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; SNA, *Sambucus nigra* agglutinin; Vg, vitellogenin; Vt, vitellin.

References

Abdu, U., Yehezkel, G., and Sagi, A. (2000) Oocyte development and polypeptide dynamics during ovarian maturation in the red-claw crayfish *Cherax quadricarinatus*. *Invert. Reprod. Develop.*, **37**, 75–83.

Abdu, U., Davis, C., Khalaila, I., and Sagi, A. (2002) The vitellogenin cDNA of *Cherax quadricarinatus* encodes a lipoprotein with calcium binding ability, and its expression is induced following the removal of the androgenic gland in a sexually plastic system. *Gen. Comp. Endocrinol.*, **127**, 263–272.

Bigge, J.C., Patel, T.P., Bruce, J.A., Goulding, P.N., Charles, S.M., and Parekh, R.B. (1995) Nonselective and efficient fluorescent labeling of glycans using 2-amino benzamide and anthranilic acid. *Anal. Biochem.*, **230**, 229–238.

Chang, C.F., Lee, F.Y., Huang, Y.S., and Hong, T.H. (1994) Purification and characterization of the female-specific protein (vitellogenin) in mature female hemolymph of the prawn, *Penaeus-monodon*. *Invertebr. Reprod. Dev.*, **25**, 185–192.

de Chaffoy de Courcelles, D. and Kondo, M. (1980) Lipovitellin from the crustacean, *artemia salina*. Biochemical analysis of lipovitellin complex from the yolk granules. *J. Biol. Chem.*, **255**, 6727–6733.

Derelle, E., Grosclaude, J., Meusy, J.J., Junera, H., and Martin, M. (1986) ELISA titration of vitellogenin and vitellin in the fresh water prawn *Macrobrachium rosenbergii*, with monoclonal antibody. *Comp. Biochem. Physiol.*, **85**, 1–4.

De Waard, P., Kamerling, J.P., Van Halbeek, H., Vliegenthart, J.F., and Broertjes, J.J. (1987) Characterization of *N*-linked gluco-oligomannose type of carbohydrate chains of glycoproteins from the ovary of the starfish *Asterias rubens* (L.). *Eur. J. Biochem.*, **168**, 679–685.

Endo, T., Hoshi, M., Endo, S., Arata, Y., and Kobata, A. (1987) Structures of the sugar chains of a major glycoprotein present in the egg jelly coat of a starfish, *Asterias amurensis*. *Arch. Biochem. Biophys.*, **252**, 105–112.

Funk, V.A., Thomas-Oates, J.E., Kielland, S.L., Bates, P.A., and Olafson, R.W. (1997) A unique, terminally glucosylated oligosaccharide is a common feature on *Leishmania* cell surfaces. *Mol. Biochem. Parasitol.*, **84**, 33–48.

Fyffe, W.E. and O'Connor, J.D. (1974) Characterization and quantification of a crustacean lipovitellin. *Comp. Biochem. Physiol. B*, **47**, 851–867.

Guile, G.R., Rudd, P.M., Wing, D.R., Prime, S.B., and Dwek, R.A. (1996) A rapid high-resolution high-performance liquid chromatographic method for separating glycan mixtures and analyzing oligosaccharide profiles. *Anal. Biochem.*, **240**, 210–226.

Hammond, C., Braakman, I., and Helenius, A. (1994) Role of *N*-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control. *Proc. Natl Acad. Sci. USA*, **91**, 913–917.

Helenius, A. and Aebi, M. (2001) Intracellular functions of *N*-linked glycans. *Science*, **291**, 2364–2369.

Huppa, J.B. and Ploegh, H.L. (1998) The eS-Sence of -SH in the ER. *Cell*, **92**, 145–148.

Imperiali, B. and O'Connor, S.E. (1999) Effect of *N*-linked glycosylation on glycopeptide and glycoprotein structure. *Curr. Opin. Chem. Biol.*, **3**, 643–649.

Jaenicke, R. (1991) Protein folding: local structures, domains, subunits, and assemblies. *Biochemistry*, **30**, 3147–3161.

Jarvis, D.L. and Finn, E.E. (1995) Biochemical analysis of the *N*-glycosylation pathway in baculovirus-infected lepidopteran insect cells. *Virology*, **212**, 500–511.

Karlsson, G.B., Butters, T.D., Dwek, R.A., and Platt, F.M. (1993) Effects of the imino sugar *N*-butyldeoxynojirimycin on the *N*-glycosylation of recombinant gp120. *J. Biol. Chem.*, **268**, 570–576.

Kerr, M.S. (1969) The hemolymph proteins of the blue crab, *Callinectes sapidus*. II. A lipoprotein serologically identical to oocyte lipovitellin. *Dev. Biol.*, **20**, 1–17.

Kim, S., Hwang, S.K., Dwek, R.A., Rudd, P.M., Ahn, Y.H., Kim, E.H., Cheong, C., Kim, S.I., Park, N.S., and Lee, S.M. (2003) Structural determination of the *N*-glycans of a lepidopteran arylphorin reveals the presence of a monoglucosylated oligosaccharide in the storage protein. *Glycobiology*, **13**, 147–157.

Küster, B., Wheeler, S.F., Hunter, A.P., Dwek, R.A., and Harvey, D.J. (1997) Sequencing of *N*-linked oligosaccharides directly from protein gels: in-gel deglycosylation followed by matrix-assisted laser desorption/ionisation mass spectrometry and normal-phase high performance liquid chromatography. *Anal. Biochem.*, **250**, 82–101.

Lee, C.Y. and Watson, R.D. (1994) Development of a quantitative enzyme-linked-immunosorbent-assay for vitellin and vitellogenin of the blue-crab *Callinectes sapidus*. *J. Crustac. Biol.*, **14**, 617–626.

Lee, F.Y., Shih, T.W., and Chang, C.F. (1997) Isolation and characterization of the female-specific protein (vitellogenin) in mature female hemolymph of the freshwater prawn, *Macrobrachium rosenbergii:* comparison with ovarian vitellin. *Gen. Comp. Endocrinol.*, **108**, 406–415.

Marinaro, J.A., Casley, D.J., and Bach, L.A. (2000) O-glycosylation delays the clearance of human IGF-binding protein-6 from the circulation. *Eur. J. Endocrinol.*, **142**, 512–516.

Matsuura, F., Ohta, M., Murakami, K., and Matsuki, Y. (1993) Structures of asparagine linked oligosaccharides of immunoglobulins (IgY) isolated from egg-yolk of Japanese quail. *Glycoconj. J.*, **10**, 202–213.

Meusy, J.J. (1980) Vitellogenin, the extraovarian precursor of the yolk protein in Crustacea: a review. *Reprod. Nutr. Dev.*, **20**, 1–21.

Meusy, J.J. and Payen, G.G. (1988) Female reproduction in malacostracan crustacea. *Zool. Sci.*, **5**, 217–265.

Ohlendorf, D.H., Barbarash, G.R., Trout, A., Kent, C., and Banaszak, L.J. (1977) Lipid and polypeptide components of the crystalline yolk system from *Xenopus laevis*. *J. Biol. Chem.*, **252**, 7922–8001.

Ohta, M., Hamako, J., Yamamoto, S., Hatta, H., Kim, M., Yamamoto, T., Oka, S., Mizuochi, T., and Matsuura, F. (1991) Structures of asparagine-linked oligosaccharides from hen egg-yolk antibody (IgY). Occurrence of unusual glucosylated oligo-mannose type oligosaccharides in a mature glycoprotein. *Glycoconj. J.*, **8**, 400–413.

Radcliffe, C.M., Diedrich, G., Harvey, D.J., Dwek, R.A., Cresswell, P., and Rudd, P.M. (2002) Identification of specific glycoforms of major histocompatibility complex class I heavy chains suggests that class I peptide loading is an adaptation of the quality control pathway involving calreticulin and ERp57. *J. Biol. Chem.*, **277**, 46415–46423.

Raikhel, A.S. and Dhadialla, T.S. (1992) Accumulation of yolk proteins in insect oocytes. *Annu. Rev. Entomol.*, **37**, 217–251.

Ryan, R.O., Anderson, D.R., Grimes, W.J., and Law, J.H. (1985) Arylphorin from *Manduca sexta*: carbohydrate structure and immunological studies. *Arch. Biochem. Biophys.*, **243**, 115–124.

Spiro, R.G., Zhu, Q., Bhoyroo, V., and Soling, H.D. (1996) Definition of the lectin-like properties of the molecular chaperone, calreticulin, and demonstration of its copurification with endomannosidase from rat liver Golgi. *J. Biol. Chem.*, **271**, 11588–11594.

Tirumalai, R. and Subramoniam, T. (1992) Purification and characterization of vitellogenin and lipovitellins of the sand crab *Emerita asiatica*: molecular aspects of crab yolk proteins. *Mol. Reprod. Dev.*, **33**, 16–26.

Tirumalai, R. and Subramoniam, T. (2001) Carbohydrate components of lipovitellin of the sand crab *Emerita asiatica*. *Mol. Reprod. Dev.*, **58**, 54–62.

Tseneklidou-Stoeter, D., Gerwig, G.J., Kamerling, J.P., and Spindler, K.D. (1995) Characterization of *N*-linked carbohydrate chains of the crayfish, *Astacus leptodactylus* hemocyanin. *Biol. Chem. Hoppe-Seyler*, **376**, 531–537.

Wallace, R.A., Walker, S.L., and Hauschka, P.V. (1967) Crustacean lipovitellin. Isolation and characterization of the major

high-density lipoprotein from the eggs of decapods. *Biochemistry*, **6**, 1582–1590.

Warrier, S. and Subramoniam, T. (2002) Receptor mediated yolk protein uptake in the crab *Scylla serrata*: crustacean vitellogenin receptor recognizes related mammalian serum lipoproteins. *Mol. Reprod. Dev.*, **61**, 536–548.

Williams, P.J., Wormald, M.R., Dwek, R.A., Rademacher, T.W., Parker, G.F., and Roberts, D.R. (1991) Characterisation of oligosaccharides from *Drosophila melanogaster* glycoproteins. *Biochim. Biophys. Acta*, **1075**, 146–153.

Wormald, M.R. and Dwek, R.A. (1999) Glycoproteins: glycan presentation and protein-fold stability. *Structure Fold. Des.*, **7**, R155–R160.

Yehezkel, G., Chayoth, R., Abdu, U., Khalaila, I., and Sagi, A. (2000) High-density lipoprotein associated with secondary vitellogenesis in the hemolymph of the crayfish *Cherax quadricarinatus*. *Comp. Biochem. Physiol. B*, **127**, 411–421.

Zal, F., Küster, B., Green, B.N., Harvey, D.J., and Lallier, F.H. (1998) Partially glucose-capped oligosaccharides are found on the hemoglobins of the deep-sea tube worm *Riftia pachyptila*. *Glycobiology*, **8**, 663–673.