Glycoconj J (2011) 28:113–123
DOI 10.1007/s10719-011-9328-3

N-glycosylation of ovomucin from hen egg white

Marina Offengenden · Messele A. Fentabil · Jianping Wu

Received: 18 February 2011 / Revised: 26 March 2011 / Accepted: 28 March 2011 / Published online: 12 April 2011
© Springer Science+Business Media, LLC 2011

Abstract Ovomucin is a bioactive egg white glycoprotein responsible for the gel properties of fresh egg white and is believed to be involved in egg white thinning, a natural process that occurs during storage. Ovomucin is composed of two subunits: a carbohydrate-rich β-ovomucin with molecular weight of 400–610 KDa and a carbohydrate-poor α-ovomucin with molecular mass of 254 KDa. In addition to limited information on O-linked glycans of ovomucin, there is no study on either the N-glycan structures or the N-glycosylation sites. The purpose of the present study was to characterize the N-glycosylation of ovomucin from fresh eggs using nano LC ESI-MS, MS/MS and MALDI MS. Our results showed the presence of N-linked glycans on both glycoproteins. We found 18 potential N-glycosylation sites in α-ovomucin. 15 sites were glycosylated, one site was found in both glycosylated and non-glycosylated forms and two potential glycosylation sites were found unoccupied. The N-glycans of α-ovomucin found on the glycosylation sites are complex-type structures with bisecting *N*-acetylglucosamine. MALDI MS of the N-glycans released from α-ovomucin by PNGase F revealed that the most abundant glycan structure is a bisected type of composition $GlcNAc_6Man_3$. Two N-glycosylated sites were found in β-ovomucin.

Keywords Ovomucin · Glycosylation · Glycans · Egg white · Mucus · Glycoprotein

M. Offengenden · M. A. Fentabil · J. Wu (✉)
Department of Agricultural, Food and Nutritional Science,
University of Alberta,
4–10 Ag/For Centre,
Edmonton, Alberta T6G 2P5, Canada
e-mail: jwu3@ualberta.ca

Abbreviations

GalNAc	*N*-acetylgalactosamine
GlcNAc	*N*-acetylglucosamine
PTS domain	mucin domain
MALDI	matrix-assisted laser desorption/ionization
TOF	time of flight
MS	mass spectrometry
LC	liquid chromatography
ESI	electro-spray ionization
VWF	von-Willebrand factor
PNGase F	peptide N-glycosidase
Hex	hexose
HexNAc	*N*-acetylhexosamine
CK	cystine knot

Introduction

Egg white is a gelatinous fluid surrounding the yolk in avian eggs. The hen egg white proteins, constituting approximately 11% of albumen's weight, provide the developing embryo with nutrients and protect it from physical damage and bacterial infection [1, 2]. Ovomucin, a glycoprotein that is responsible for the gel properties of the egg white [3], comprises approximately 3.5% of the egg white proteins [4]. Its high viscosity provides mechanical protection from penetration of pathogens into egg yolk [1]. Ovomucin also possesses anti-bacterial, anti-viral and anti-tumor activities [4–6]. It is composed of two components: a carbohydrate-rich β-ovomucin with approximately 60% carbohydrates and a carbohydrate-poor α-ovomucin with approximately 15% carbohydrates that form a linear high molecular weight polymeric structure via disulfide bonds [7, 8]. From rheological [8], structural [9] and genetic [10] aspects, ovomucin is similar to other gel forming mucins

[11]. Mucins are high molecular weight glycoproteins responsible for the viscoelastic and gel-like properties of the mucus that covers epithelial cells thus protecting them from dehydration and microbial infection [12]. Characterized by a high amount of O-linked glycans linked via GalNAc to serine or threonine residues, mucin genes belong to the MUC gene family [13]. In addition to a central O-glycosylated domain (referred also as PTS domain), which contributes to the ability of mucins to interact with various microorganisms and their properties, non-PTS domains such as von-Willebrand-factor (VWF) domains, C-terminal domain and cysteine rich domains are present. These domains are conserved in mucins, contain N-linked carbohydrates and are important for their dimerization and polymerization [11, 14, 15].

Gel forming mucins in the chicken are coded from a locus that is similar to a corresponding locus of human mucins of MUC2, MUC5AC, MUC5AB and MUC6 located on chromosome 11 [10, 13]. β-ovomucin is an orthologue of human MUC6; while α-ovomucin, a protein similar to MUC2 with a missing PTS domain, is coded by an additional gene that is inserted between MUC2 and MUC5AC [10]. Similar to non-PTS domains in other mucins, α-ovomucin contains von-Willerbrand factors D, C and cystine knot domains [9]. The reason for an additional mucin-type protein with only the non-PTS N-glycosylated domains in the egg white is unknown; however, it is possible that these domains along with the N-linked carbohydrates have a particular role in the egg white, such as hatching [9].

Several O-linked carbohydrates released from ovomucin by reductive β-elimination were characterized previously. These partially sulfated short oligosaccharides composed of 3–5 units of *N*-acetylgalactosamine, galactose, *N*-acetylglucosamine, and sialic acid [16–19], might contribute to the gelling properties of ovomucin in a similar manner to other mucins [8, 20], and to its ability to interact with various microorganisms. In contrast, information about N-linked glycans of ovomucin is not available. N-glycans are covalently linked to asparagines on the protein that are located within a consensus sequence of Asn-Xaa-Ser/Thr, where Xaa can be any amino acid with the exception of proline. They share a common pentasaccharide core and can be classified as high-mannose-type, complex-type, and hybrid-type [21]. Many egg white proteins are N-glycosylated. Ovalbumin, the major egg white protein [2] has high mannose and hybrid type N-glycans [22–24]. Another major egg white protein is ovomucoid, an inhibitor of proteolytic enzymes [2]. It possesses complex type N-glycans composed of the pentasaccharide core and a bisecting GlcNAc, some of which have pentaantennary structure and galactose in a terminal position [25, 26].

α-ovomucin contains 24 potential N-glycosylation sites in its amino acid sequence [9]. The 827 amino acid fragment of β-ovomucin submitted to EMBL/GenBank/DDBJ databases by Watanabe *et al.* (http://www.uniprot.org/uniprot/Q6L608) contains 9 potential N-glycosylation sites; however, neither the N-glycan structures nor the N-glycosylation sites of ovomucin were investigated. In the present study, we utilized mass spectrometric techniques (LC-ESI MS, MS/MS and MALDI-TOF MS) to determine the major N-glycan structures, site occupancy and site-specific N-glycosylation of ovomucin.

Materials and methods

Materials

Fresh eggs from White Leghorn laid within 24 h were obtained from the Poultry Research Centre of the University of Alberta (Edmonton, Canada). Sodium chloride was purchased from Acros Organics (Morris Plains, NJ, USA). Laemmli sample buffer, 20% SDS solution, molecular weight protein marker for SDS-PAGE and iodoacetamide were purchased from BIO-RAD (Bio-Rad Laboratories, Inc., Hercules, CA, USA). Hydrochloric acid and sodium hydroxide were bought from Fisher Scientific Inc. (Fisher Scientific, Ottawa, ON, Canada). Proteomics grade PNGase F from *Elizabethkingia meningosepticum*, ammonium bicarbonate, 2-mercaptoethanol, dithiothreitol and dimethyl sulfoxide were purchased from Sigma-Aldrich (St. Louis, MO, USA). Sequencing grade modified trypsin was purchased from Promega (Madison, WI, USA). Acetonitrile (HPLC grade) was from ACROS (Fair Lawn, NJ, USA). Water used for the experiments was obtained from the Milli-Q water supply system (Millipore Corporation, Billerica, MA, USA).

SDS-PAGE

Egg white from fresh egg was manually separated from the yolk at the day of collection (within 24 h of laying), vortexed and diluted to give an aqueous solution of 20% v/v egg white with 5% SDS and 1% β-mercaptoethanol. This solution was incubated overnight at 25°C, diluted in a ratio of 1:1 with Laemmli sample buffer containing 5% β-mercaptoethanol and loaded on 4-20% Mini-PROTEAN TGX precast gel (Bio-Rad Laboratories, Inc.). Loading volume was 20 µl and a molecular weight standard (Bio-Rad) was loaded to a separate well. SDS-PAGE was performed [27] in a Mini-PROTEAN tetra cell with a PowerPac Basic electrophoresis apparatus (Bio-Rad Laboratories, Inc., Hercules, CA, USA) at constant voltage of 200 V. The gel was stained with Commassie Brilliant Blue and scanned in Alpha Innotech gel scanner

(Alpha Innotech Corp., San Leandro, CA, USA) with FluorChem SP software.

In-gel PNGase F digestion

PNGase F digestion was performed according to Sigma-Aldrich technical bulletin for PNGase F with slight modifications. To briefly describe, excised gel bands of ovomucin subunits were transferred into low-retention microcentrifuge tubes (Fisher Scientific, Ottawa, ON, Canada) and destained by double incubation with 200 µl of 200 mM ammonium bicarbonate with 40% acetonitrile for 30 min. After drying in a Speed Vac (Savant Automatic Environmental SpeedVac, AES2000, Savant Instruments, Inc., Farmingdale, NY), 15 µl of prepared PNGase F solution (prepared by adding 100 µl of water to a vial containing 50 units of enzyme) was added, centrifuged briefly and incubated overnight at 37°C. After overnight incubation, released carbohydrates were extracted with 200 µl of water four times, collected and lyophilized. Deglycosylated gel pieces were dried in a SpeedVac and subjected to digestion with trypsin.

In-gel trypsin digestion

Excised gel pieces containing the protein of interest with and without PNGase F treatments (destaining was carried out as for the PNGase F treated sample), were digested following a modified protocol [28] based on the method of Shevchenko *et al*. [29]. Briefly, the proteins in the gel pieces were reduced with 10 mM DTT, alkylated with 50 mM iodoacetamide, washed with 100 mM ammonium bicarbonate and acetonitrile, and dried in a SpeedVac. Dried gel bands were subjected to overnight proteolytic digestion at 37°C with trypsin. Digestion was performed with 0.8 µg enzyme (solution of 20 ng/µl in 50 mM ammonium bicarbonate). After incubation, peptides were extracted into new microcentrifuge tubes with 30 µl of 100 mM ammonium bicarbonate followed by extraction with 30 µl solution containing 5% formic acid and 50% acetonitrile in water twice and dried to ~15 µl in a SpeedVac.

Nano LC-ESI MS and MS/MS analysis

The tryptic digest was analyzed by a hybrid quadrupole orthogonal acceleration time-of-flight mass spectrometer, QToF Premier (Waters, Milford, MA), online connected to Waters nanoACQUITY ultra high performance liquid chromatography (UPLC) system. 5 µl of the tryptic digest containing (glyco-) peptides was loaded onto a nanoAcquity UPLC system with peptide trap (180 µm× 20 mm, Symmetry® C18 nanoAcquity™ column, Waters, Milford, MA) and a nano analytical column (75 µm×100 mm, Atlantis™ dC_{18} nanoAcquity™ column, Waters, Milford, MA). Desalting of the trapped sample was achieved by flushing the trap column with a solution of 1% acetonitrile and 0.1% formic acid in water (Solvent A) at a flow rate of 10 µl/min for 3 min. Peptides were separated with a gradient of 1-65% solvent B (acetonitrile, 0.1% formic acid) over 35 min at a flow rate of 300 nL/min. The column was connected to a QToF premier (Waters Corporation) for ESI-MS and MS/MS analysis of the effluent.

Permethylation of N-glycans and MALDI-TOF MS and MS/MS

Permethylation of released N-glycans was performed according to Dell *et al*. [30] using the sodium hydroxide/dimethyl sulfoxide slurry method. Lyophilized permethylated glycans were dissolved in 20 µl of methanol. 0.8 µl of sample solution was spotted to an 800 µm Bruker's anchorChip™ MALDI target (Bruker Daltonics, Billerica, MA, USA). 0.8 µl of matrix solution was then spotted on top of the sample and left to dry. 2, 5-dihydroxybenzoic acid (DHB) was used as the matrix compound. Matrix solution was prepared in 50% acetonitrile at a final concentration of 3.5 mg/ml. MALDI MS and MS/MS were performed on an ultraflexXtreme™ MALDI-TOF/TOF (Bruker Daltonics, Billerica, MA, USA) mass spectrometer in positive MS or lift mode.

Bioinformatic methods

Database search was performed by Mascot search engine (Matrix Science). The mass spectra were submitted to Mascot search engine as a pkl file. The parameters used for database search were as follows: type of search defined was MS/MS Ion Search, Carbamidomethylation (C) was defined as a fixed modification, while Deamidation (NQ) (for PNGase F treated samples) and Oxidation (M) were defined as variable modifications. Sequence alignment was performed by ClustalW in Uniprot database accessible from http://www.uniprot.org, motifs search was performed by ScanProsite [31, 32], accessed from http://ca.expasy.org/prosite.

Results

Ovomucin subunits separation and identification by SDS-PAGE

Reduced fresh egg white was subjected to SDS-PAGE analysis. Three bands potentially containing ovomucin at approximately 350–400, 250 and 150 kDa were labeled as 1, 2 and 3 respectively, as shown in Fig. 1. These bands

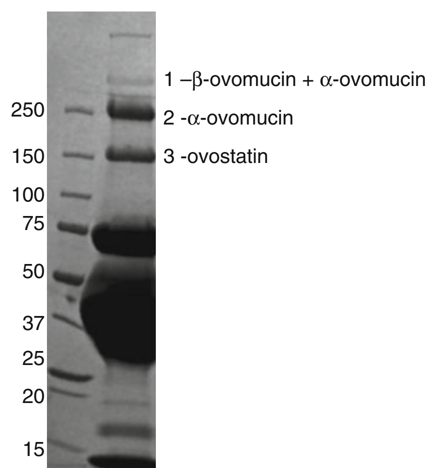

Fig. 1 SDS-PAGE separation of ovomucin subunits. Molecular weight markers (in kDa) appear on the left, analyzed bands containing ovomucin, labeled 1, 2 and 3, on the right

were excised and treated consecutively with PNGase F and trypsin. The peptides obtained were analyzed by nano LC-ESI MS and MS/MS. The mass spectra were submitted for database search using Mascot search engine, as explained in Materials and methods section. α-ovomucin was detected in bands 1 and 2, while β-ovomucin was detected in band 1 only. No ovomucin was detected in band 3, where the primary protein was characterized as ovostatin.

Identification of N-glycosylation sites of α-ovomucin

In order to characterize the potential N-glycosylation sites, PNGase F treated and trypsinized samples were analyzed. The enzyme PNGase F removes N-linked carbohydrates from proteins while deamidating the originally glycosylated asparagine into aspartic acid [33] that results in an increase in molecular mass by 0.9840 Da. This increase in mass was exploited to identify glycosylation sites of α and β-ovomucin.

Tryptic peptides obtained from digestion of deglycosylated ovomucins were subjected to nano LC-ESI MS and MS/MS. The data was analyzed by Mascot database search engine to identify tryptic peptides containing deglycosylated asparagines, supplemented by manual analysis of the data. By using this approach, 18 of the potential glycosylation sites of α-ovomucin (see Table 1 for details) and two sites on β-ovomucin (Table 2) were

Table 1 Glycosylation sites in α-ovomucin, obtained from LC-MS/MS data of tryptic (glyco)peptides by MASCOT database and manual spectrum analysis

Site	Peptide position	Mass[a]	*m/z* Obs	Sequence	Glycosylation
N^{69}	54-78	3019.3[c]	1007.1^{3+}	FTFPGTC*TYVFASHC*NDSYQDFNIK	-
N^{381}	368-385	2069.8	1035.9^{2+}	VYSSGGTYSTPC*QNC*TC*K	+
N^{528}	510-529	2374.2	594.5^{4+} 792.4^{3+}	VQMKPVMQLSITVDHSYQNR	+
N^{599}	589-618	3396.6	850.2^{4+}	FAQHWC*ALLSNTSSTFAAC*HSVVDPSVYIK	+
N^{673}	657-677	2506.1[c]	1253.5^{2+}	QGIC*DPSEEC*PETMVYNYSVK	-
N^{680}	678-684	987.4	494.7^{2+}	YC*NQSC*R	+
N^{772}	763-787	2767.1	1384.6^{2+} 923.4^{3+}	DC*PAPMYYFNC*SSAGPGAIGSEC*QK	+
N^{855}	853-862	1322.5	662.3^{2+}	QWNC*TDNPC*K	+
N^{1036}	1028-1047	2316.0	1159.0^{2+}	ITSTC*SNINMTDLC*ADQPFK	+
N^{1219}	1215-1230	1844.9	923.4^{2+}	TYPLNETIYSQTEGTK	+
N^{1371}	1370-1392	2746.2	1374.1^{2+}	FNESWDFGNC*QIATCLGEENNIK	+[(b)]
N^{1452}	1451-1467	2009.0	1005.4^{2+}	ENC*TYVLVELIQPSSEK	+
N^{1567}	1565-1580	1924.8	963.4^{2+}	YYNNTMGLC*GTC*TNQK	+
N^{1639}	1636-1645	1340.7[c] 1341.6	671.4^{2+} 671.8^{2+}	IIWNLTEC*HR	−/+
N^{1792}	1775-1802	3576.5	895.1^{4+}	EAWEHDC*QYC*TC*DEETLNISC*FPRPC*AK	+
N^{1807}	1803-1810	917.5	$916.5^{1+(b)}$	SPPINC*TK	+ [(b)]
N^{1841}	1837-1842	749.4	749.4^{1+}	TC*IINK	+
N^{1964}	1960-1984	2815.2	$1408.1^{2+(b)}$	APYDNC*TQYTC*TESGGQFSLTSTVK	+ [(b)]

[a] Calculated mass (deamidated Asn) with modified Cys.

[b] Peptides detected manually (not detected by Mascot database search) in PNGase F untreated sample

[c] Calculated mass (Asn not deamidated) with modified Cys.

* Carbamidomethylated cysteines

Table 2 Glycosylation sites of β-ovomucin, obtained from LC-MS/MS data of tryptic peptides by MASCOT database

Site	Peptide position	Mass[(a)]	*m/z* Obs	Sequence	Glycosylation
N^{238}	238-249	1477.6	739.8^{2+}	NC*TC*STLSEYSR	+
N^{945}	944-952	1182.6	592.3^{3+}	YNMTLIWNK	+

[a] Calculated mass (deamidated Asn) with modified Cys.

* Carbamidomethylated cysteines

characterized. Two sites on α-ovomucin were not glycosylated (N^{69}, N^{673}), one site – N^{1639} existed in both glycosylated and not glycosylated forms, while all other identified sites on α-ovomucin and β-ovomucin were glycosylated. A representative MS and MS/MS data of one of the glycosylation sites containing a deamidated asparagine, Asn^{772}, is presented in Fig. 2; data of a not-glycosylated site, Asn^{637}, is represented in Fig. 3.

Determination of site specific N-glycan structures by LC-MS/MS

Glycosylation site occupancy and glycan heterogeneity on a particular site were determined from glycopeptide analysis. Tryptic glycopeptides obtained from in-gel digested α-ovomucin without PNGase F treatment were subjected to nano LC-ESI MS and MS/MS analysis. Glycopeptides

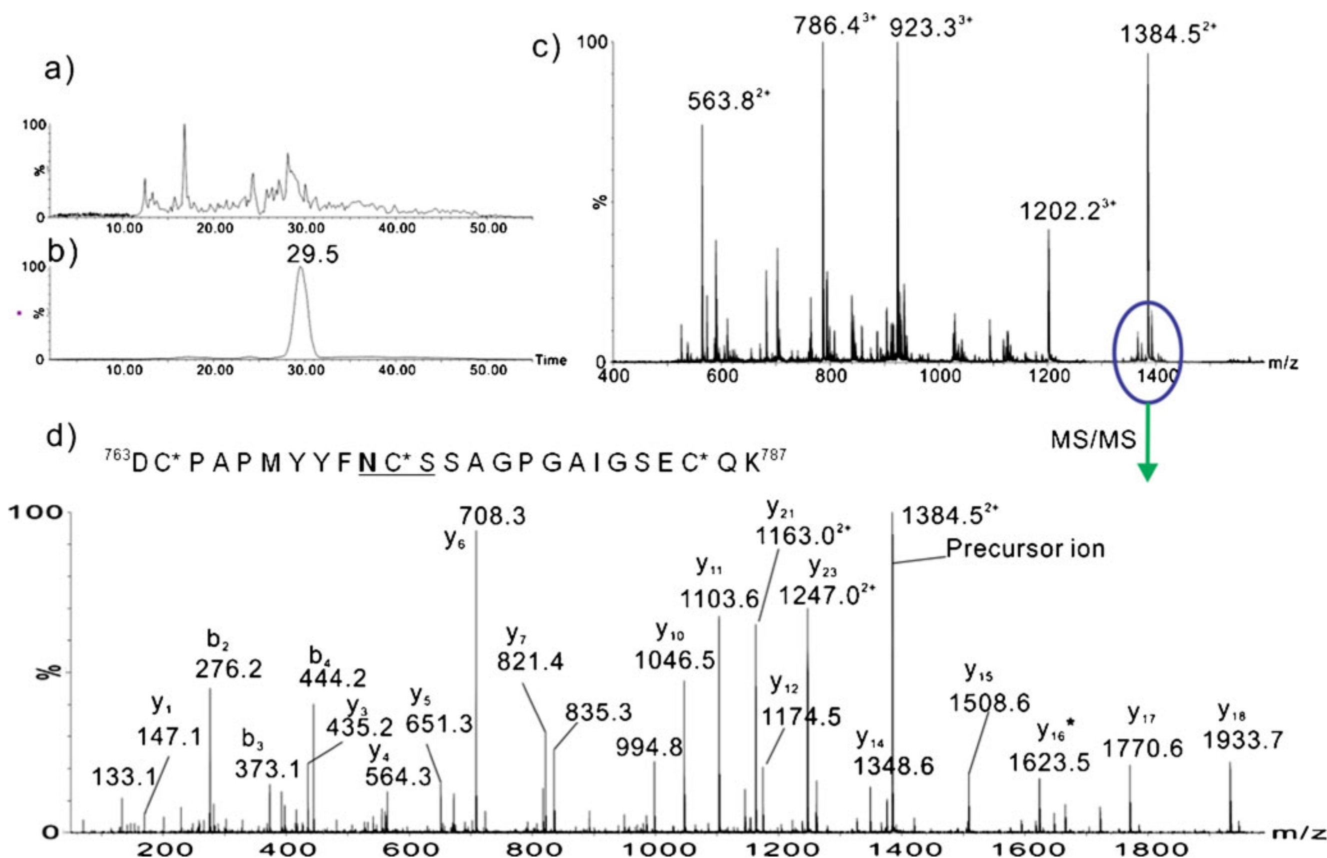

Fig. 2 LC-MS and MS/MS ion chromatogram of a peptide ^{763}DC*PAPMYYFNC*SSAGPGAIGSEC*QK^{787} with deglycosylated asparagine - N^{772}. **a** Ion chromatogram of a sample untreated with PNGase F. The peptide was not detected in this sample, due to glycosylation. **b** Sample treated with PNGase F. The peptide eluted at 29.5 min. **c** MS extracted from PNGase F treated sample. The doubly (m/z 1384.5^{2+}) and the triply (m/z 923.4^{3+}) charged forms of the peptide were detected. **d** MS/MS of the doubly charged ion (m/z 1384.5^{2+}) shows b and y ions as labeled. The mass difference between y16 and y15, which corresponds to Asn, is 115 Da instead of 114 Da as a result of deglycosylation with PNGase F. Cysteines marked with an asterisk are carbamidomethylated

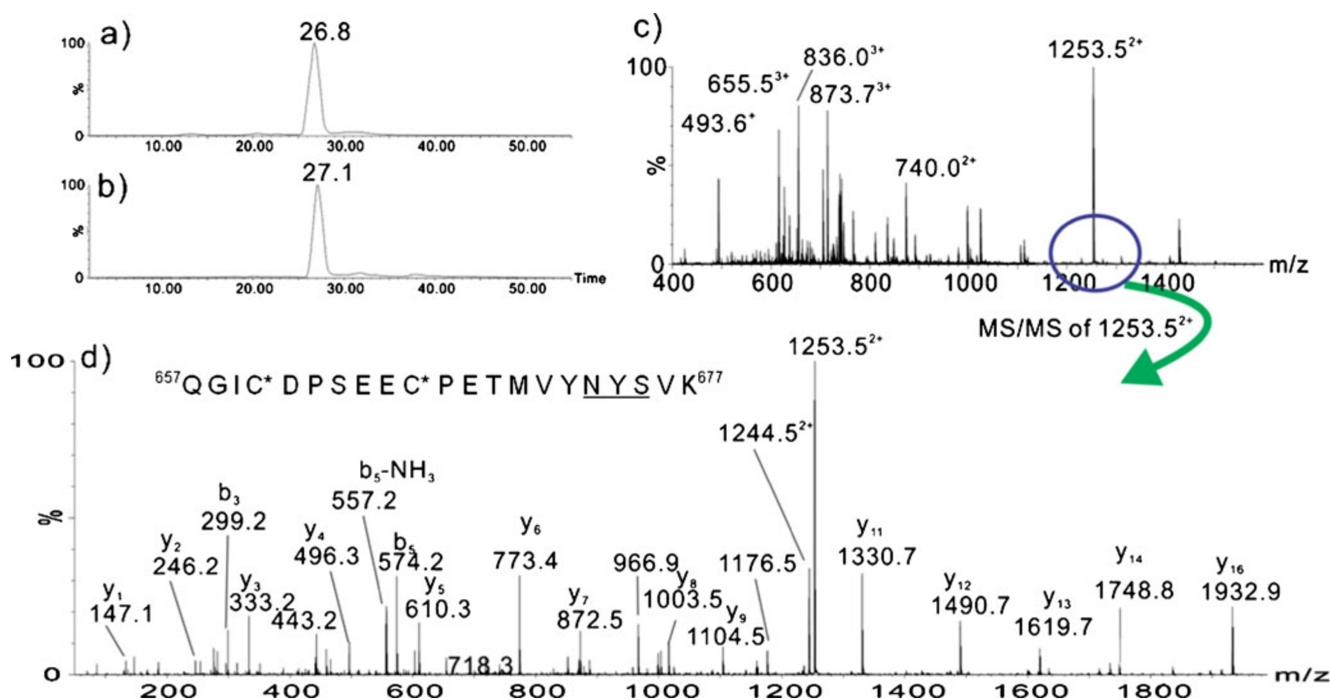

Fig. 3 LC-MS and MS/MS ion chromatogram of a peptide [657]QGIC*DPSEEC*PETMVYNYSVK[677], which does not contain a glycans on its potential glycosylation site – N[673]. Ion chromatogram of **a** untreated and **b** treated samples with PNGase F. The peptide eluted around 27 min in both samples, which implies that the peptide is not glycosyalted. **c** MS extracted from PNGase F treated sample: The doubly (m/z 1253.5^{2+}) and the triply (m/z 836.0^{3+}) charged forms of the peptide were detected. **d** MS/MS of the doubly charged ion (m/z 1253.5^{2+}) shows b and y ions as labeled. The mass difference between y4 and y5, which corresponds to Asn, is 114 Da which is not deamidated by PNGase F. Cysteines marked with an asterisk are carbamidomethylated

were detected by searching for diagnostic glycan ions such as HexNAc (m/z 204) and Hex-HexNAc (m/z 366) resulting from fragmentation [33]. Representative MS/MS data of glycopeptides is presented in Fig. 4.

MS/MS data of a doubly charged glycopeptide ion at m/z 1227.5 is presented in Fig. 4a. In the MS/MS data, the peak at m/z 748.4 corresponds to a peptide predicted by ExPAsy proteomics server (www.expasy.org), TC*IINK, which contains the potential N-glycosylation site, Asn[1841]. The peaks at m/z 951.5 and 1154.6 correspond to the peptide with one and two *N*-acetylglucosamines, respectively. The peak at m/z 1316.6 corresponds to $GlcNAc_2Man$ attached to the peptide. The peak at m/z 1519.7 has the same structure as the previous peak with additional HexNAc that indicates the presence of a bisecting GlcNAc. The peak at 1843.8 corresponds to the pentasaccharide core of $GlcNAc_2Man_3$ with a bisecting GlcNAc. From the mass of the glycopeptide (2453.0 Da) calculated based on the detected doubly charged ion at m/z 1227.5 and the peptide mass (747.4 Da), we determined that the glycan part has a mass of 1705.5 Da and is composed of $GlcNAc_2Man_3$+$HexNAc_4$.

Figure 4b and c show MS/MS of two triply charged glycopeptides of m/z 1016.4 and 1084.5 with similar MS/MS spectra. The glycopeptides have a common peptide at m/z 671.4 (doubly charged ion of IIWNLTEC*HR). MS/MS of the 1016.4 ion shows that the glycopetide has the same glycan structure ($GlcNAc_2Man_3$+$HexNAc_4$) as the glycopeptides in Fig. 4a, while the ion at 1084.5 has one more HexNAc.

Following the same strategy, five different glycosylation sites carrying different glycans were identified (Table 3). Although sulfated glycans were reported previously for O-glycans of ovomucin [16, 17, 19], there was no evidence of sulfation in N-glycans of α-ovomucin.

Determination of N-Glycan profile of α-ovomucin by MALDI-TOF MS

Glycans were released from α-ovomucin (band 2 in Fig. 1) by treatment with PNGase F, permethylated and analyzed by MALDI-TOF MS and MS/MS. MALDI-TOF MS spectrum of permethylated glycans is presented in Fig. 5. The peaks at m/z 1906.5, 2151.6, 2396.7, 2600.7 and 2641.8 correspond to the complex type N-glycans obtained from LC-ESI MS and MS/MS analysis of glycopeptides presented in Table 3. MALDI-TOF MS results revealed two major glycan structures at m/z of 1906.5 and 2151.6 that

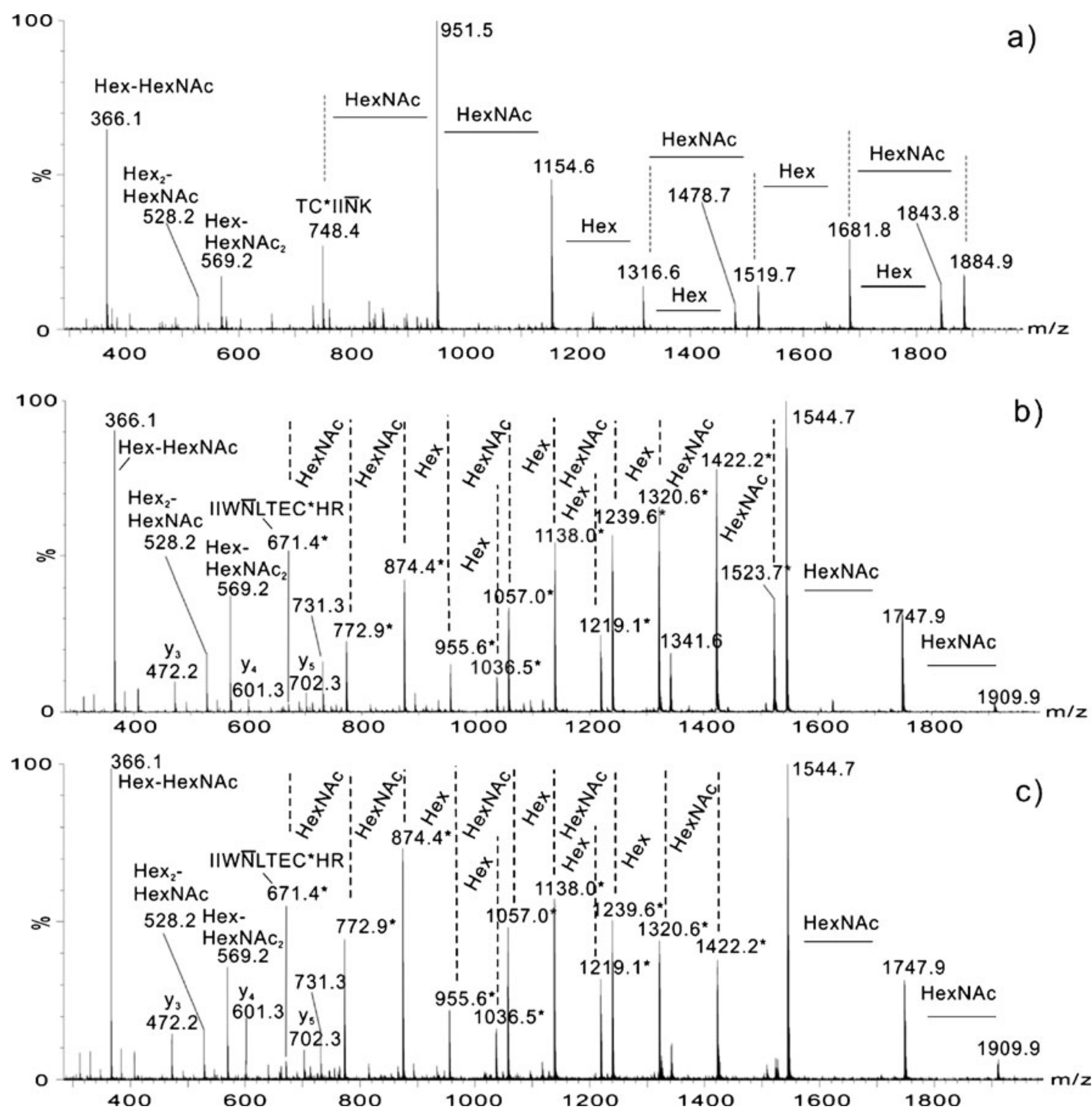

Fig. 4 LC-MS/MS ion chromatogram of representative glycopeptides **a** LC-MS/MS of a doubly charged ion at m/z 1227.5 corresponding to a peptide TC*II**N**K and a glycan $GlcNAc_2Man_3$+$HexNAc_4$. **b** LC-MS/MS of a triply charged glycopeptides ion at m/z 1016.4 corresponding to a peptide IIW**N**LTEC*HR carrying a glycan $GlcNAc_2Man_3$+$HexNAc_4$. **c** MS/MS of a triply charged glycopeptides ion at m/z 1084.5 corresponding to a peptide IIW**N**LTEC*HR carrying a glycan $GlcNAc_2Man_3$+$HexNAc_5$. Cysteines marked with an asterisk are carbamidomethylated

are composed of a pentasaccharide core+$HexNAc_3$ and the core+$HexNAc_4$, respectively. The peaks at m/z 2109.6 and 2355.7 had the same composition as the peaks at m/z of 1906.5 and 2151.6, respectively, with an additional hexose on each. The glycan structures (see Fig. 5) for these and other peaks are based on our findings that ovomucin had a bisecting GlcNAc from LC-ESI MS/MS results, known N-glycan structures of other egg white proteins [23–26, 34] and carbohydrate composition of α-ovomucin that contained mainly *N*-acetylglucosamine (3.1-8.5%), mannose (2.9-4.6%), small amount of galactose (1.1-4.3%) and sialic acid (0.3-1.3%) [7, 35–37]. A glycan with composition of Core+Hex $HexNAc_6$ with m/z of 1511.6 on site N^{1841} (Table 3) was not detected by MALDI-TOF/MS, possibly due to its low abundance, which resulted in a weak signal.

Motifs search in α-ovomucin sequence and alignment of its N-glycosylation sites with human MUC2

α-ovomucin sequence was searched by ScanProsite against its database to identify motifs. ScanProsite identified four VWF D domains, two VWF C domains and a cystine knot. VWF D domains contained 6 glycosylation sites, one

Table 3 Tryptic glycopeptides from α-ovomucin and site specific N-glycosylation

Site	Peptide	m/z obs	Peptide mass	Glycan mass	Perm. mass [(a)]	Glycan structure
N[1639]	IIWNLTEC*HR	1016.4^{3+}	1341.7	1705.5	2151.6	Core+$HexNAc_4$
		1084.5^{3+}	1341.7	1908.6	2396.7	Core+$HexNAc_5$
N[1219]	TYPLNETIYSQTEGTK	1183.8^{3+}	1844.9	1705.5	2151.6	Core+$HexNAc_4$
N[1841]	TC*IINK	1227.5^{2+}	748.4	1705.5	2151.6	Core+$HexNAc_4$
		1328.5^{2+}	748.4	1908.6	2396.7	Core+$HexNAc_5$
		1430.6^{2+}	748.4	2111.8	2641.8	Core+$HexNAc_6$
		1409.6^{2+}	748.4	2070.7	2600.7	Core+Hex $HexNAc_5$
		1511.6^{2+}	748.4	2273.8	-	Core+Hex $HexNAc_6$
N[1807]	SPPINC*TK	1210.0^{2+}	916.5	1502.4	1906.5	Core+$HexNAc_3$
		1311.0^{2+}	916.5	1705.5	2151.6	Core+$HexNAc_4$
		1412.6^{2+}	916.5	1908.6	2396.7	Core+$HexNAc_5$
		1010.0^{3+}	916.5	2111.8	2641.8	Core+$HexNAc_6$
N[599]	FAQHWC*AL-LSNTSSTFAAC*HSVVDPS-VYIK	1276.3^{4+}	3396.6	1705.5	2151.6	Core+$HexNAc_4$
		1225.5^{4+}	3396.6	1502.4	1906.5	Core+$HexNAc_3$

* Carbamidomethylated cysteines

[a] Perm. Mass refers to mass of permethylated glycans as $[M+Na]^+$ molecular ions detected by MALDI-TOF MS (see Fig. 5).

glycosylation site was found in VWF C2 domain and no glycosylation sites were found in the cystine knot, as illustrated in Fig. 6. To determine N-glycosylation sites conservation, alignment of α-ovomucin sequence with the sequence of human MUC2 was performed by ClustalW. Four glycosylation sites were conserved among the two proteins: N[1371], N[1452], N[1567] and N[1964] as presented in Fig. 6.

Discussion

Ovomucin, a glycoprotein responsible for the gel properties of fresh egg white, is composed of two components: β-ovomucin, with molecular weight of 400–610 kDa [35, 38, 39] and α-ovomucin - a polypeptide of 2087 amino acids with an estimated molecular mass of 254 kDa [9, 37]. Two components of α-ovomucin -

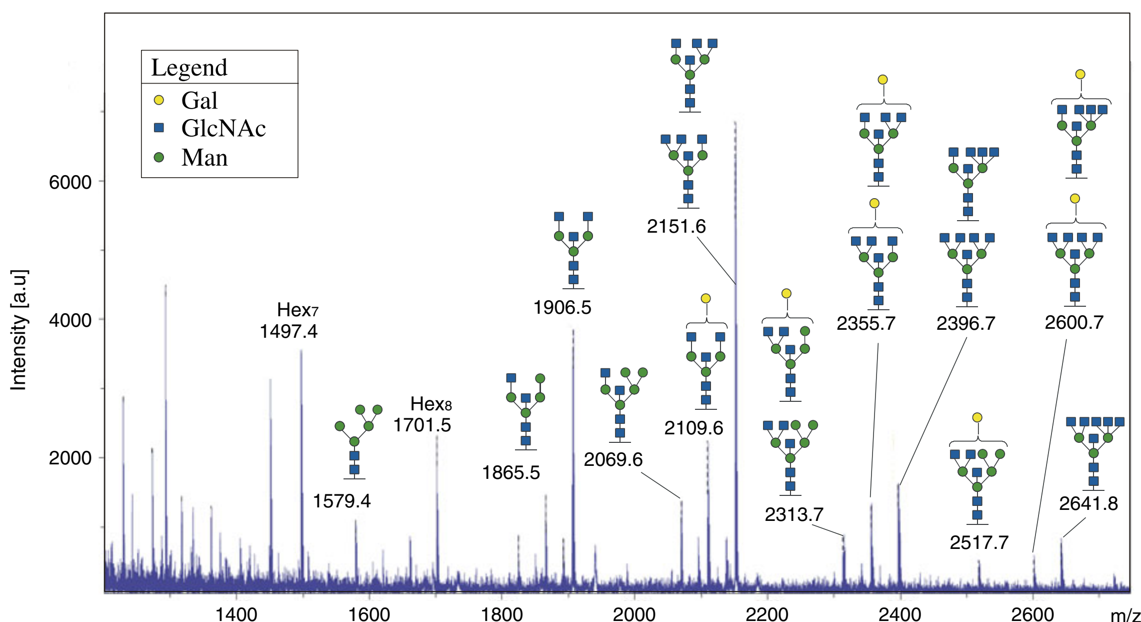

Fig. 5 MALDI-TOF ion chromatogram of permethylated glycans released by in-gel digestion with PNGase F from SDS-PAGE band 2 that contained α-ovomucin. Possible glycan structures are presented

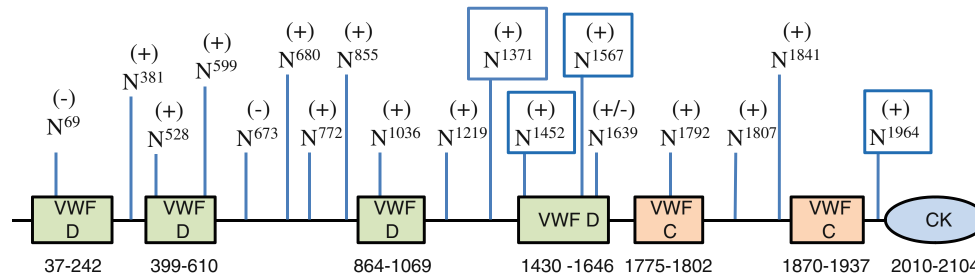

Fig. 6 Location of N-glycosylation sites on α-ovomucin. Domains were assigned by ScanProsite. Glycosylation status as was determined in Results section (Table 1) is presented as well as conserved glycosylation sites are surrounded by a square (aligned with human MUC2 by ClustalW)

α1 and α2-ovomucins with estimated molecular weights of 150 and 220 kDa respectively were reported [35] and another α-ovomucin component with molecular weight of 350 kDa was reported as well [39]. Our results confirmed the presence of β-ovomucin in the 350–400 kDa gel region, in agreement with previous reports [35, 38]. α-ovomucin was present in gel bands that corresponded to molecular weights of ~350 kDa and 250 kDa; however, it was not detected in 150 kDa region where the major protein was ovostatin. The absence of α1-ovomucin in the 160 kDa fraction obtained by gel filtration was reported previously [39]. However, LC-MS/MS analysis of egg white revealed α-ovomucin on SDS-PAGE gel in the area of 100–250 kDa [40]. In the present experiment, fresh eggs were analyzed and no further ovomucin purification was performed to minimize the impact of sample preparation on the structure of ovomucin. Depolymerization of ovomucin complex through reduction of disulfide bonds was suggested as a cause for egg white thinning, occurring during storage [4]. Degradation of α-ovomucin could possibly occur during egg storage or ovomucin preparation, leading to α-ovomucin migration to lower molecular weight regions on SDS-PAGE. Therefore, previous observations of a smaller α-ovomucin in ovomucin extracts might be due to degradation of α-ovomucin occurred during egg sample storage or ovomucin preparation.

N-glycans play important roles in folding, oligomerization, sorting and transport of proteins [41]. In order to understand how glycosylation influences the biological function of a protein, it is indispensable not only to elucidate the composition and structures of the glycans but also to identify the sites on the protein that are glycosylated. Protein glycosylation depends on various factors such as its 3D structure, location of the potential site, its accessibility by glycosylating enzymes and glycosylation of other sites on the protein, leading to 10-30% of the potential sites to be not occupied [42]. We determined glycosylation status of 18 glycosylation sites, out of the potential 24 [9] for the first time. 15 sites were occupied by glycans, two sites (N^{69} and N^{673}) were not glycosylated and one site (N^{1639}) existed in both glycosylated and non-glycosylated forms. Only 2 of the potential N-glycosylation sites in β-ovomucin were identified, possibly due to a high amount of densely O-glycosylated regions, as characteristic to mucus glycoproteins [10, 43] that prevented cleavage. Therefore, other approaches should be applied for further investigation of β-ovomucin.

Glycopeptide analysis of 5 glycosylation sites revealed glycan heterogeneity with up to five glycans on each site (see Table 3), except for N^{1219} on which only one glycan type was determined, possibly due to the large mass of the glycopeptides that prevented it from being detected. This is not surprising, as glycoforms are very common due to the fact that glycosylation results from a series of not always complete enzymatic reactions [44].

Identified glycans from α-ovomucin glycopeptides had a pentasaccaride core of $GlcNAc_2Man_3$, which is characteristic to N-linked glycans, and a bisecting GlcNAc. Bisected type N-glycans are common in other egg white proteins such as ovotransferrin, ovomucoid and chicken riboflavin binding protein [24–26, 34]. As glycosylation is species and tissue specific [42], we assumed that N-glycans from α-ovomucin may have similar structures to the above egg white proteins. Taking that into account, we proposed possible glycans structures for the peaks obtained by MALDI-TOF MS (Fig. 5) and determined that the most abundant glycan is composed of $GlcNAc_2Man_3$ core, a bisecting GlcNAc and another 3 GlcNAc antennae located on the mannoses of the core. The fact that this glycan was present on all glycopeptides identified by LC-MS/MS supports this observation.

N-glycans from α-ovomucin were released by PNGase F from SDS-PAGE band corresponding to a molecular weight of 250 kDa. However, this band contained ovalbumin as a contaminant, although with a much lower score than that of α-ovomucin, as can be seen in Table 4. Therefore, it is likely that our N-glycan profile

Table 4 Proteins identified in band 2 by Mascot [(a)]

Protein	NCBI accession number	Mascot score
Ovomucin α-subunit	gi 12583679	4972
PNGase F	gi 157833480	1122
Ovalbumin	gi 129293	693
Ovotransferrin	gi 1351295	238
Ovotransferrin - Chain A, N-terminal lobe, iron loaded open form	gi 6729806	238
Trypsin	gi 136429	137

[a] Only proteins with scores higher than 80 were included.

contained small amounts of glycans from ovalbumin. We identified several relatively weak peaks at m/z 1579.4, 2069.6, 2313.7 and 2517.7 that can possibly correspond to abundant glycans from ovalbumin [22–24]. In addition to the glycan structures that were detected by both LC-MS/MS and MALDI-TOF MS (m/z 1906.5, 2151.6, 2396.7, 2600.7 and 2641.7), three additional peaks at m/z 1865.5, 2109.6 and 2355.7 were detected. While the latter two could result from galactose addition to the abundant glycans (m/z 1906.5 and 2151.6, respectively) it is unknown whether the peak at m/z of 1865.5 is a glycan from ovomucin or other source.

α-ovomucin was found to possess sequence similarity of 33% with the N-terminal D-domain, 27% with the C-terminal D4 domain, 34% with C1 domain and 29% with the cystine knot of human pre-pro-VWF domains [9]. It was shown that specific N-glycosylation sites on VWF in D1, D', B1 and CK domains influence its synthesis and secretion [45]. We used ScanProsite in order to assign specific N-glycosylation sites into domains. Seven sites were found within four VWF D domains and one site within VWF C2 domain as shown in Fig. 6. When aligned with human MUC2 (*results not shown*), four conserved glycosylation sites (N^{1371}, N^{1452}, N^{1567}, N^{1964}) were found, one site more than in previously reported alignment [9]. This conservation might suggest that these specific sites play an important role in the properties of α-ovomucin, possibly in the same way as specific N-linked glycans are involved in dimerization of rat MUC2 [15]. However, studies of tyrosinase related protein family show that N-glycosylation site conservation does not guarantee conserved roles in glycoprotein family among different species [46]. Therefore, the roles of these conserved glycosylation sites in α-ovomucin, the exact structures of its N-glycans along with the role of α-ovomucin in the egg white and the factors contributing to its degradation should be further investigated.

Acknowledgements This work was supported by grants from Natural Sciences and Engineering Research Council of Canada (NSERC) and Alberta Livestock Meat Agency Inc. (ALMA) to J. Wu.

References

1. Stevens, L.: Egg proteins: what are their functions? Sci. Prog. **79** (Pt 1), 65–87 (1996)
2. Burley, R.W., Vadehra, D.V.: The albumen: chemistry. In: Burley, R.W., Vadehra, D.V. (eds.) The Avian Egg: Chemistry and Biology, pp. 65–128. Wiley, New York (1989)
3. Brooks, J., Hale, H.P.: The mechanical properties of the thick white of the hen's egg. Biochim. Biophys. Acta **32**, 237–250 (1959)
4. Omana, D.A., Wang, J., Wu, J.: Ovomucin - a glycoprotein with promising potential. Trends Food Sci. Technol. **21**, 455–463 (2010)
5. Hiidenhovi, J.O.: In: Huopalahti, R., Lopez-Fandino, R., Anton, M., Schade, R. (eds.) Bioactive Egg Compounds, pp. 61–68. Springer, Berlin; New York (2007)
6. Mine, Y., D'Silva, I.: Bioactive components in egg white. In: Mine, Y. (ed.) Egg Bioscience and Biotechnology, pp. 141–184. Wiley-Interscience, Hoboken (2008)
7. Robinson, D.S., Monsey, J.B.: Studies on the composition of egg-white ovomucin. Biochem. J. **121**, 537–547 (1971)
8. Rabouille, C., Aon, M.A., Thomas, D.: Interactions involved in ovomucin gel-forming properties: a rheological-biochemical approach. Arch. Biochem. Biophys. **270**, 495–503 (1989)
9. Watanabe, K., Shimoyamada, M., Onizuka, T., Akiyama, H., Niwa, M., Ido, T., Tsuge, Y.: Amino acid sequence of α-subunit in hen egg white ovomucin deduced from cloned cDNA. DNA Seq. **15**, 251–261 (2004)
10. Lang, T., Hansson, G.C., Samuelsson, T.: An inventory of mucin genes in the chicken genome shows that the mucin domain of Muc13 is encoded by multiple exons and that ovomucin is part of a locus of related gel-forming mucins. BMC Genomics. **7** (2006)
11. Perez-Vilar, J.: Gastrointestinal mucus gel barrier. In: Bernkop-Schnürch, A. (ed.) Oral Delivery of Macromolecular Drugs: Barriers, Strategies and Future Trends, pp. 21–48. Springer, Dordrecht [The Netherlands], New York (2009)
12. Bansil, R., Turner, B.S.: Mucin structure, aggregation, physiological functions and biomedical applications. Curr. Opin. Colloid. Interface Sci. **11**, 164–170 (2006)
13. Dekker, J., Rossen, J.W.A., Büller, H.A., Einerhand, A.W.C.: The MUC family: an obituary. Trends Biochem. Sci. **27**, 126 (2002)
14. Strous, G.J., Dekker, J.: Mucin-type glycoproteins. Crit. Rev. Biochem. Mol. Biol. **27**, 57–92 (1992)
15. Bell, S.L., Xu, G., Khatri, I.A., Wang, R., Rahman, S., Forstner, J. F.: N-linked oligosaccharides play a role in disulphide-dependent dimerization of intestinal mucin Muc2. Biochem. J. **373**, 893–900 (2003)
16. Strecker, G., Wieruszeski, J., Martel, C., Montreuil, J.: Determination of the structure of sulfated tetra- and pentasaccharides obtained by alkaline borohydride degradation of hen ovomucin. A fast atom bombardment-mass spectrometric and 1H-NMR spectroscopic study. Glycocnj. J **4**, 329–337 (1987)
17. Strecker, G., Wieruszeski, J.-M., Martel, C., Montreuil, J.: Complete 1H- and 13C-N.M.R. assignments for two sulphated oligosaccharide alditols of hen ovomucin. Carbohydr. Res. **185**, 1–13 (1989)
18. Strecker, G., Wieruszeski, J.-M., Cuvillier, O., Michalski, J.C., Montreuil, J.: 1H and 13C-NMR assignments for sialylated oligosaccharide-alditols related to mucins. Study of thirteen

components from hen ovomucin and swallow nest mucin. Biochimie **74**, 39–52 (1992)
19. Kato, A., Hirata, S., Kobayashi, K.: Structure of the sulfated oligosaccaride chain of ovomucin. Agr. Biol. Chem. **42**, 1025–1029 (1978)
20. Sellers, L.A., Allen, A., Morris, E.R., Ross-Murphy, S.: Mucus glycoprotein gels. Role of glycoprotein polymeric structure and carbohydrate side-chains in gel-formation. Carbohydr. Res. **178**, 93–110 (1988)
21. Varki, A. (ed.): Essentials of glycobiology. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (1999)
22. Tai, T., Yamashita, K., Ogata-Arakawa, M., Koide, N., Muramatsu, T., Iwashita, S., Inoue, Y., Kobata, A.: Structural studies of two ovalbumin glycopeptides in relation to the endo-β-N-acetyl-glucosaminidase specificity. J. Biol. Chem. **250**, 8569–8575 (1975)
23. Yamashita, K., Tachibana, Y., Kobata, A.: The structures of the galactose-containing sugar chains of ovalbumin. J. Biol. Chem. **253**, 3862–3869 (1978)
24. Harvey, D.J., Wing, D.R., Kuster, B., Wilson, I.B.H.: Composition of N-linked carbohydrates from ovalbumin and co-purified glycoproteins. J. Am. Soc. Mass Spectrom. **11**, 564–571 (2000)
25. Yamashita, K., Kamerling, J.P., Kobata, A.: Structural study of the carbohydrate moiety of hen ovomucoid. Occurrence of a series of pentaantennary complex-type asparagine-linked sugar chains. J. Biol. Chem. **257**, 12809–12814 (1982)
26. Yamashita, K., Kamerling, J.P., Kobata, A.: Structural studies of the sugar chains of hen ovomucoid. Evidence indicating that they are formed mainly by the alternate biosynthetic pathway of asparagine-linked sugar chains. J. Biol. Chem. **258**, 3099–3106 (1983)
27. Laemmli, U.K.: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**, 680–685 (1970)
28. Towndrow, K.M., Jia, Z., Lo, H.H., Person, M.D., Monks, T.J., Lau, S.S.: 11-Deoxy,16,16-dimethyl prostaglandin E2 induces specific proteins in association with its ability to protect against oxidative stress. Chem. Res. Toxicol. **16**, 312–319 (2003)
29. Shevchenko, A., Wilm, M., Vorm, O., Mann, M.: Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. Anal. Chem. **68**, 850–858 (1996)
30. Dell, A., Khoo, K.-H., Panico, M., McDowell, R.A., Etienne, A. T., Reason, A.J., Morris, H.R.: FAB-MS and ES-MS of glycoproteins. In: Fukuda, M., Kobata, A. (eds.) Glycobiology: A Practical Approach, pp. 187–222. IRL Press at Oxford University Press, Oxford, New York (1993)
31. Sigrist, C.J.A., Cerutti, L., de Castro, E., Langendijk-Genevaux, P., Bulliard, V., Bairoch, A., Hulo, N.: PROSITE, a protein domain database for functional characterization and annotation. Nucleic Acids Res. **38**, D161–D166 (2010)
32. de Castro, E., Sigrist, C.J.A., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P., Gasteiger, E., Bairoch, A., Hulo, N.: ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res. **34**, W362–W365 (2006)
33. Medzihradszky, K.F.: Characterization of protein N-glycosylation. Methods Enzymol. **405**, 116–138 (2006)
34. Spik, G., Coddeville, B., Montreuil, J.: Comparative study of the primary structures of sero-, lacto- and ovotransferrin glycans from different species. Biochimie **70**, 1459–1469 (1988)
35. Itoh, T., Miyazaki, J., Sugawara, H., Adachi, S.: Studies on the characterization of ovomucin and chalaza of the hen's egg. J. Food Sci. **52**, 1518–1521 (1987)
36. Hayakawa, S., Sato, Y.: Physicochemical identity of alpha ovomucins or beta ovomucins obtained from the sonicated insoluble and soluble ovomucins. Agric. Biol. Chem. 1185–1191 (1977)
37. Tsuge, Y., Shimoyamada, M., Watanabe, K.: Structural features of newcastle disease virus- and anti-ovomucin antibody-binding glycopeptides from pronase-treated ovomucin. J. Agric. Food Chem. **45**, 2393–2398 (1997)
38. Hayakawa, S., Sato, Y.: Subunit structures of sonicated α and β-ovomucin and their molecular weights estimated by sedimentation equilibrium. Agric. Biol. Chem. **42**, 957–961 (1978)
39. Hiidenhovi, J., Aro, H.S., Kankare, V.: Separation of ovomucin subunits by gel filtration: Enhanced resolution of subunits by using a dual-column system. J. Agric. Food Chem. **47**, 1004–1008 (1999)
40. Mann, K.: The chicken egg white proteome. Proteomics **7**, 3558–3568 (2007)
41. Helenius, A., Aebi, M.: Intracellular functions of N-linked glycans. Science **291**, 2364 (2001)
42. Dwek, R.A.: Glycobiology: toward understanding the function of sugars. Chem. Rev. **96**, 683–720 (1996)
43. Perez-Vilar, J., Hill, R.L.: The structure and assembly of secreted mucins. J. Biol. Chem. **274**, 31751–31754 (1999)
44. Rudd, R.M., Dwek, R.A.: Glycosylation: heterogeneity and the 3D structure of proteins. Crit. Rev. Biochem. Mol. Biol. **32**, 1–100 (1997)
45. McKinnon, T.A.J., Goode, E.C., Birdsey, G.M., Nowak, A.A., Chan, A.C.K., Lane, D.A., Laffan, M.A.: Specific N-linked glycosylation sites modulate synthesis and secretion of von Willebrand factor. Blood **116**, 640–648 (2010)
46. Gupta, G., Surolia, A., Mitra, N., Sinha, S.: Probing into the role of conserved N-glycosylation sites in the Tyrosinase glycoprotein family. Glycoconj. J. **26**, 691–695 (2009)