

Article

Mass spectrometry and two-dimensional electrophoresis to characterize the glycosylation of hen egg white ovomacroglobulin

Fang Geng, Xi Huang, Kaustav Majumder, Zihui Zhu, Zhaoxia Cai, and Meihu Ma

J. Agric. Food Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.jafc.5b02618 • Publication Date (Web): 31 Aug 2015

Downloaded from <http://pubs.acs.org> on September 5, 2015

Just Accepted

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.



ACS Publications

**Mass spectrometry and two-dimensional electrophoresis to characterize the
glycosylation of hen egg white ovomacroglobulin**

Fang Geng[†], Xi Huang[†], **Kaustav Majumder**[†], Zhihui Zhu[†], Zhaoxia Cai[†], Meihu Ma^{*†}

[†]National R&D Center for Egg Processing, College of Food Science and Technology,
Huazhong Agricultural University, Wuhan, 430070, China

^{*}Department of Food Science, University of Guelph, Guelph, Ontario N1G 2W1,
Canada

Corresponding Author:

*Meihu Ma. Phone: +86-27-87283177; Fax: +86-27-87283177; E mail:
mameihuhn@163.com.

ABSTRACT

Glycosylation of proteins plays an important role in their biological functions, such as allergenicity. Ovomacroglobulin (OVMG) is a glycoprotein from hen egg white, but few studies have done so far to delineate the glycosylated sites of OVMG. The present study characterized the glycosylation of OVMG using mass spectrometry and two-dimensional electrophoresis. MALDI-TOF-MS showed that OVMG subunit $[M+H]^+$ ion has a peak at 183,297 m/z, therefore the carbohydrate moieties is calculated as 11.5 % of the whole OVMG molecule. HPLC-ESI-MS/MS confirmed that out of 13 potential N-glycosylation sites of OVMG, 11 sites were glycosylated, 1 site (N¹²²¹) was found in both glycosylated and non-glycosylated forms. On the two-dimensional electrophoresis gel, a series of OVMG spots horizontally distributed at 170 kDa with an isoelectric point range of 5.03-6.03, indicating the heterogeneity of glycosylation of OVMG. These results provided important information for understanding of structure, function, and potential allergenic sites of OVMG.

KEYWORDS

Egg white; Ovomacroglobulin; Glycosylation; Mass spectrometry; Two-dimensional electrophoresis

INTRODUCTION

Egg is a rich source of dietary proteins and well known for their nutritional value. Despite the high nutritional value egg white proteins are associated with food allergy. The carbohydrate moieties of egg white glycoproteins play important roles in egg white gel properties; egg induced allergies, and biological activities. It was considered that the highly glycosylated ovomucin is mainly contributed to the gel properties of egg white, and liberation carbohydrate units of ovomucin could result in the thinning of egg white during egg storage^{1, 2}. Ovomucin-lysozyme aggregation, which combined with the negative charges of the terminal sialic acid in ovomucin and the positive charges of lysyl ϵ -amino groups in lysozyme, also play an important role in the keeping of gelation of thick egg white^{3, 4}. Glycans of egg white proteins are also intimately involved in the allergic reactions. The main egg allergens ovomucoid, ovalbumin, and ovomucin are all glycoproteins⁵. And it had been reported that deglycosylation of ovomucoid domain III decreased its binding ability with human IgE⁶. Glycation modification of egg white protein could affect the sensitization of egg-induced allergies, as it has demonstrated that the mannosylated egg white protein and ovalbumin showed an attenuation of orally induced egg allergy in mice^{7, 8}. Additionally, there is evidence that glycosylation was crucial to the biological activities of egg proteins. For instance, Mg^{2+} ions can interact with the carbohydrate moiety of ovomucin and result in an increase of antivirus activity⁹.

Therefore, researchers focus on the characterization of glycosylation sites and glycan structure of egg white proteins. β -ovomucin, the most heavily glycosylated component in egg white, contains approximately 60% carbohydrates that are mostly O-linked glycans¹⁰. While the carbohydrates of α -ovomucin are mainly N-glycans, and 16 N-glycosylation sites have been identified¹¹. Ovomuroid, a highly

glycosylated protein containing 20-25% of carbohydrates moieties, possesses complex type N-glycans that distributed in 5 potential N-glycosylation sites¹²⁻¹⁴. Ovalbumin also possesses 1 identified N-glycosylation site (N²⁹³) with high mannose and hybrid type N-glycans^{15, 16}. These investigations proved the important information to reveal the role of glycosylation of egg white proteins in allergenicity and biological activities, as well as to elucidate the underlying mechanism of egg quality reduction during storage.

However, there are some minor egg white proteins were identified as glycoproteins but lacked information about its glycosylation, such as ovomacroglobulin (OVMG)^{17, 18}. OVMG, also known as “ovostatin”, is a protease inhibitor and possesses diverse biological activity. As a member of the ancient macroglobulin family, OVMG shows a considerable homology with other members of the amino acid sequence and shares typical characteristics with other macroglobulins. But there are still some differences between OVMG and other macroglobulin family members. Recent studies have showed that OVMG and its human homolog, alpha-2-macroglobulin (A2MG), could binds to the surfactant protein D via lectin-carbohydrate interactions¹⁹, indicating that the glycosylation of OVMG might play an important role in the immune function. Hence, the characterization of the glycosylation of OVMG would be helpful to elucidate the mechanism of its functions, and also to understand the difference between OVMG and other macroglobulins.

According to sequence analysis, OVMG contains 13 potential N-glycosylation sites^{17, 20}, but it has not been confirmed experimentally yet. The present study explored the glycosylation of OVMG using mass spectrometry and two-dimensional electrophoresis (2-DE). OVMG was purified from fresh egg white, and the molecular weight of OVMG subunit was measured by MALDI-TOF MS, thus the mass of

carbohydrate moieties was calculated. Then, OVMG was deglycosylated by PNGase F and subsequently digested by proteases, and the N-glycosylation sites were identified using HPLC-ESI-MS/MS. Finally, 2-DE analysis of OVMG was performed to display the heterogeneity of glycosylation of OVMG.

MATERIALS AND METHODS

Preparation of OVMG: Fresh hen eggs laid within 24 h from White Leghorns were used for the purification of OVMG. The operation performed as previously described²¹. Briefly, fresh hen egg white (100 mL) was diluted with an equal volume of distilled water, followed by two-step PEG (Polyethylene glycol 8000) precipitation (5-9 %, w/v) to obtain OVMG-rich precipitate (P₅₋₉). And then the precipitate was dissolved in PBS and further purified by gel filtration chromatography (Sephacryl S-200 HR, GE Healthcare Bio-sciences AB, Sweden) at a flow rate of 0.75 mL/min with an automatic liquid chromatography system (JiaPeng Technology, Shanghai, China). The first peak of eluent was collected, and desalted with distilled water using an Amicon Ultra-15 Centrifugal Filter Devices (100 kDa NMWL, Millipore, USA), then stored at 4 °C. The purity of OVMG was tested by HPLC using a TSK gel G2000SWXL column (7.8 × 300 mm; TOSOH, Tokyo, Japan) with a Waters 2695 Separations Module (Waters, USA) as previous description²¹, and the result was shown in Figure S1.

Measure the molecular weight of OVMG subunit by MALDI-TOF MS: Molecule weight of OVMG subunit was measured by a Bruker Reflex™ III MALDI-TOF mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany) working in positive ion mode. A saturated solution of α -cyano-4-hydroxycinnamic acid in acetonitrile and 0.1% TFA was used for the matrix^{22, 23}. Purified OVMG (about 2 μ M) was reduced with

109 10 mmol/L DTT, and then 1 μL of sample solution was mixed with 1 μL of matrix
110 solution. Deposit the mixed solution on the stainless steel target plate and drying at
111 room temperature. The ions were ionized under a laser intensity of 80 %, and
112 accelerated with an acceleration voltage of 20 kV, then measured in a linear mode and
113 m/z range of 40000-220000²².

114 **PNGase F deglycosylation and protease digestion:** PNGase F digestion was
115 performed following the instructions of the manufacturer. OVMG was dissolved in
116 distilled water to 1.5 mg/mL, and adjusted pH to 8.0 by 100 mM NH_4HCO_3 .
117 Denatured OVMG was prepared by adding 30 μL of 0.15% SDS with 80 mM
118 2-mercaptoethanol to 255 μL of OVMG solution, and incubating for 20 min at 95°C.
119 After that, the solution was kept at room temperature to cool down, 15 μL of PNGase
120 F enzyme solution (500 U/ml, Sigma-Aldrich) was then added, and incubate at 37 °C
121 overnight to release asparagine-linked oligosaccharides¹¹.

122 Deglycosylated OVMG solution (300 μL) was alkylated with 50 mM
123 iodoacetamide for 30 min at room temperature. Then the solution was equally divided
124 into 3 parts and separately digested with trypsin, chymotrypsin, and endoproteinase
125 Glu-C (Sigma-Aldrich) at an enzyme-to-substrate ratio of 1:50 w/w overnight at
126 37 °C, and then stop the reaction by heating to 100 °C for 5 min. The samples were
127 then lyophilized for the subsequent LC-MS/MS analysis.

128 **Identify the N-glycosylation sites of OVMG using HPLC-ESI-MS/MS:** The
129 digests (by trypsin, chymotrypsin, and Glu-C) of deglycosylated OVMG were
130 analyzed separately to identify the N-glycosylation sites. Mass spectrometric
131 measurements were performed by HPLC-ESI-MS/MS using a LTQ mass spectrometer
132 (Thermo-Fisher Science, Bremen, Germany) equipped with a nanospray source and
133 Eksigent RP-HPLC (Eksigent Technologies, Dublin, USA). 10 μL of the digests were

loaded onto PepMap C18 analytical column (75 $\mu\text{m} \times 15 \text{ cm}$, Dionex), and separation was carried out at a flow rate of 300 nL/min using a gradient constructed from solution A (2% ACN, 0.1% formic acid) and solution B (80% ACN, 0.1% formic acid): 2-40% B for 90 min; 40-100% B for 15 min; 100% B for 15 min. The spray voltage was set as 2.2 kV, and temperature of the ion transfer capillary was 200 °C. The mass ranges for the MS scan was set to 200–2000 m/z. For MS/MS analysis, peptides were subjected to fragmentation by collision-induced decomposition (CID), and normalized collision energy was 35%.

For the identification of N-glycosylation sites, matching search was performed against a local database containing the OVMG FASTA file (P20740-OVOS_CHICK) using the MASCOT search engine (Matrix Science). Types of the search were MS/MS Ion search, with either trypsin, chymotrypsin, or Glu-C as the enzyme allowing up to two missed cleavages. Carbamidomethylation (C) was set as a fixed modification; Deamidation (NQ) and Oxidation (M) were specified as variable modifications. The data was searched with a peptide ion mass tolerance and a fragment ion mass tolerance of $\pm 0.25 \text{ Da}$. Only peptides with Mascot ion score higher than 23 ($p < 0.005$) and containing the PNGase consensus sequence (N-X-S/T, $X \neq P$) were considered for further analysis²⁴.

Two-dimensional electrophoresis and protein identification: 2-DE analysis was carried out with the Ettan IPGphor 3 System for the first-dimension isoelectric focusing (IEF) and the Ettan DALTSix System (GE Healthcare) for the second-dimension SDS-PAGE. Sample (100 μL , 80 μg of protein) was loaded onto DryStrip IPG strips (24 cm; pH 4–7) and isoelectrically focused for a total of 64000 Vh per gel using the same conditions previously described (20 °C; 300 V for 0.5 h, 700 V for 0.5 h, 1500 V for 1.5 h, 9000 V for 3 h, and 9000 V for 5 h)²⁵. After

159 equilibrating to resolubilize proteins and reducing disulfide bonds, the
160 second-dimension electrophoresis was performed using a 10 % SDS-PAGE. The gels
161 were run at 2 W per gel for 40 min and followed by 17 W per gel for approximately 4
162 h. Protein spots on 2-DE gels were visualized by Coomassie Brilliant Blue staining.
163 Gel evaluation and data analysis were carried out using the Image Master V 7.0
164 program (GE Healthcare)²⁵. For identification, the candidate protein spots were
165 excised manually from the stained gels and digested with trypsin (Promega, Madison,
166 WI, USA). The digested peptides were extracted and then identified using a
167 MALDI-TOF MS/MS (Bruker, Karlsruhe, Germany). Data were searched against the
168 protein database (NCBIInr, *Gallus gallus*) via the MASCOT program ([http://www.](http://www.matrixscience.com)
169 [matrixscience.com](http://www.matrixscience.com))²⁶.

170 **Bioinformatics Analysis.** For comparing the N-glycosylation sites of chicken OVMG
171 (UniProtKB accession: P20740) with its homologous proteins, another 9 egg
172 macroglobulins (OVMGs) from *Aves* were selected: common turkey (G1NMI3,
173 *Meleagris gallopavo*), duck (R0J9X6, *Anas platyrhynchos*), emperor penguin
174 (A0A087R4G8, *Aptenodytes forsteri*), adelic penguin (A0A093NQZ9, *Pygoscelis*
175 *adeliae*), dalmatian pelican (A0A091SNW1, *Pelecanus crispus*), flamingo
176 (A0A091V5J9, *Phoenicopterus ruber*), red-legged seriema (A0A091M0T0, *Cariama*
177 *cristata*), white-tailed sea-eagle (A0A091P9X7, *Haliaeetus albicilla*), and hoatzin
178 (A0A091VDK9, *Opisthocomus hoazin*). The sequences of OVMGs were obtained
179 from UniProt database, and N-glycosylation sites were predicted by NetNGlyc 1.0
180 Server (<http://www.cbs.dtu.dk/services/NetNGlyc/>). Alignment of the ten sequences
181 was performed online (<http://www.uniprot.org/align>). Finally, the comparison of three
182 representative OVMGs (chicken, P20740; turkey, G1NMI3; duck, R0J9X6) were
183 visualized and presented by Illustrator for Biological Sequences (IBS, Version 1.0)²⁷.

RESULTS AND DISCUSSION

Molecular mass of glycan moiety in OVMG: OVMG includes a 36-residue signal peptide and a 1437-residue main chain¹⁷. Molecular weight (MW) of the main chain is 162.2 kDa²⁰, but the measured MW of the subunit was reported from 165 kDa to 195 kDa by different methods²⁸⁻³⁰. In the present study, MW of OVMG subunit was measured by MALDI-TOF MS. As shown in Figure 1, four ion signals peaks were observed at m/z 44791, 62154, 92203, and 183297, which respectively correspond to $([M+4H]^{4+})$, $([M+3H]^{3+})$, $([M+2H]^{2+})$, and $([M+H]^+)$. Thereby the MW of OVMG subunit is 183.3 kDa, and the MW of glycan moiety was calculated as 21.1 kDa, accounting for 11.5% of the whole OVMG molecular.

This result provided a more accurate MW of OVMG, and was consistent with those of sequence speculation¹⁷ and experimental analysis^{31, 32}. Additionally, the Full Width at Half Maximum (FWHM) of molecular ion peak of OVMG subunit $([M+H]^+)$ was approximately 6000 (m/z), indicating the glycosylation of OVMG is not homogeneous.

Comparing with other egg white glycoproteins, the proportion of OVMG glycan is higher than ovalbumin, ovotransferrin, and lysozyme, but lower than ovomucin and ovomucoid^{10, 33}. Since glycans play an important role in the recognition of allergens by the immune system³⁴, such a high proportion of glycans in OVMG strongly suggest that OVMG should be considered as a suspected allergen, and its potential allergenicity need be confirmed in future studies.

N-glycosylation sites of OVMG: In order to identify the N-glycosylation sites, OVMG was deglycosylated by PNGase F, and subsequently digested by proteinase, then the digests were analyzed by HPLC-ESI MS/MS. PNGase F cleaves N-linked oligosaccharides from glycoproteins, and results in a transfer of the originally

glycosylated asparagine to aspartic acid with a 0.98 Da increase in molecular weight. The mass change of peptide fragments in which the glycosylated asparagine is converted to aspartic acid can be exploited to identify N-glycosylation sites.

For getting higher sequence coverage, de-glycosylated OVMG was digested by three different proteases (trypsin, chymotrypsin, and Glu-C), and then digests were subjected to HPLC-ESI MS/MS. The MS/MS data were searched against a local database containing the OVMG FASTA file by Mascot engine to identify peptides containing de-glycosylated asparagines. A total of 439 peptides (Table S1) were found out by the following conditions: Mascot ion score higher than 23 ($p < 0.005$), and including PNGase consensus sequence. These peptides contain 12 N-glycosylation sites, among them, only one site (N^{1221}) existed in both glycosylated and not glycosylated forms, while all of the other 11 sites were glycosylated (Table 1 and Figure 2A). Another potential glycosylation site (N^{342}) was not covered in the MS/MS analysis. Two representative MS/MS spectrums of the N-glycosylation sites, N^{67} and N^{1221} were presented in Figure 2, and other spectrums of N-glycosylation sites were showed in Figure S2.

Alignment of OVMG N-glycosylation sites with its homologous proteins:

Glycosylation has an important impact on the folding, structure, and function of protein, and results in the diversity of glycoproteins. For well understanding the role of glycosylation in the function of a protein, glycosylated sites are indispensable information. Macroglobulin family is ancient and conservative, BLAST the protein sequence of OVMG against UniProtKB database (<http://www.uniprot.org/blast/>), and results showed that the Sequence Identity between OVMGs from different species of *Aves* are higher than 80 %. Two representative birds OVMGs (G1NMI3, from common turkey; and R0J9X6, from duck) were selected, and the sequence, structural

234 domain, and N-glycosylation sites of them were compared with chicken OVMG
235 (P20740). The results were visualized and showed in Figure 3. It was observed that
236 the domains of these three proteins are highly conserved (Figure 3A), suggesting that
237 they possess structural similarities. The number of predicted N-glycosylation sites of
238 these three proteins are 13 (P20740), 17 (G1NMI3) and 17 (R0J9X6), respectively,
239 and the alignment results revealed that total of 9 predicted N-glycosylation sites were
240 located at identical position among all the three proteins.

241 Interestingly, G1NMI3 has a higher Sequence Identity with P20740, however,
242 shares more predicted N-glycosylation sites with R0J9X6 (Figure 3B). For further
243 questioning this finding, another 7 sequences of OVMGs from birds (*Aves*) were selected,
244 and the predicted N-glycosylation sites of them were compared with the above three
245 OVMGs. Results showed that total numbers of predicted N-glycosylation sites of the
246 ten OVMGs are between 13 and 17, and chicken OVMG possesses the fewest sites
247 (Table S2). For further analysis, we define a “conserved site” to be the
248 N-glycosylation site which exist in more than six sequences. The alignment results
249 revealed that the total number of conserved N-glycosylation sites of the ten OVMGs
250 is 15. Chicken OVMG (P20740) has 10, common turkey OVMG (G1NMI3) and
251 mallard OVMG (R0J9X6) have 13, and other 7 OVMGs have 14-15 conserved
252 N-glycosylation sites.

253 The above results showed that chicken OVMG has some differences with other
254 OVMGs in the aspect of N-glycosylation sites. The fewest total and conserved
255 N-glycosylation sites indicated that chicken OVMG loses some of them, but the
256 reason are not clear yet. Since the sequence identity of the ten OVMGs is higher than
257 83 %, there should be some other reasons besides the genetic factors. Among these ten
258 species, only chicken, turkey and duck are domesticated. The domestication of

chicken begins approximately 5400 BC³⁵, much earlier than duck (500 BC) and turkey (300 BC–AD 100)^{36,37}. Therefore, the domestication processes might play a certain role in the losing of OVMGs' N-glycosylation sites. Considering most of food allergens are glycoproteins, losing of glycosylation sites may result in reduce food allergenicity. Recently transcriptome analyses of carrot also suggested the human selection to reduce allergy: comparing with wild carrot roots, putative allergen-related protein genes were silenced in cultivated carrot roots³⁸. A systematic bioinformatics analysis is needed to confirm whether the losing of N-glycosylation sites is widespread in other domesticated animals and plants. The study of this phenomena will provide important information for the research of food allergenicity.

The sequence identity between OVMG and mammalian macroglobulin is 39-44%¹⁷. Comparison of OVMG and human alpha-2-macroglobulin (A2MG) showed that the cysteine residues are highly conserved between these two proteins. 24 of the total 27-cysteine residues of OVMG are place at identical positions in the sequence comparing with A2MG, suggesting these two proteins share the similar folding pattern and tertiary structure. In contrast, the alignment results revealed that OVMG and A2MG only share 3 N-glycosylation sites at positions N⁶⁷, N⁸², and N⁴⁰³. Few conserved sites implied that the differences between OVMG and A2MG, as reported in the previous researches^{28, 29, 39}, might be partly due to their difference in glycosylation sites.

Heterogeneity of OVMG glycosylation: Glycoproteins usually existed in multiple glycoforms, and the heterogeneity is considered to affect the physical and chemical properties of the protein, immunogenicity, and biology functions. Results of MW of glycan moiety and N-glycosylation sites had revealed the heterogeneity of OVMG

glycosylation. Moreover, an intuitive demonstration of the heterogeneity of OVMG was also achieved by 2-DE assay.

After expanding in 2-DE gel, OVMG showed as two series of spots around the molecular weight of 170 kDa (Figure 4). Spots 1, 2, and 3 were relatively separate and with pI values ranging from 5.18 to 5.30, spots 4 and 5 were undivided spots with a train of vertical streaks and pI values ranging from 5.90 to 6.05. All of the 5 spots were identified as OVMG (gi|45382565, ovostatin [Gallus gallus]; Table 2). Spot 7, which has a molecular weight about 100 kDa and with a pI value range of 5.30, also was identified as OVMG, corresponding to the degradation fragment of OVMG^{28,39}. The wide range of pI and vertical streaks could be considered as evidence of the heterogeneity of OVMG glycosylation.

The heterogeneity of glycoprotein usually occurs at glycosylation sites (macroheterogeneity) with variable occupancy of glycosylation sites, and also at glycan profile on each site with occupied by more than one kinds of glycan (microheterogeneity). Combined with the N-glycosylation sites results, it was initially speculated that OVMG has a low heterogeneity in N-glycan sites, but a high heterogeneity in the glycan profiles.

Interestingly, spot 6, which has a MW of 123.2 kDa and a pI of 5.07, was identified as ovalbumin. The MW is approximately three times than normal ovalbumin, indicating that ovalbumin might self-crosslinked by non-disulfide covalent bonds. This kind of ovalbumin covalent aggregates also appeared in the previous comparative proteomic research of egg white proteins²⁶, but how they are formed has not been studied yet. A hypothesis is that ovalbumin are oxidized during the purification processes or storage, and results in the covalent crosslinking of ovalbumin.

308 This study experimentally confirmed the glycosylation of chicken OVMG, and
309 reported the N-glycosylation sites for the first time. The data also demonstrated the
310 heterogeneity of OVMG glycosylation. These results will provide valuable
311 information to understand the bioactivity function and identify potential allergenic
312 sites of OVMG. However, for fully characterize the glycosylation of OVMG, the
313 glycan profile and glycopeptide structure should be investigated in the future.

314

315

316 ASSOCIATED CONTENT

317 Supporting Information

318 The purity of OVMG, and MS/MS spectrum of peptides containing de-glycosylated
319 asparagine are shown in *Supporting Figures* (PDF). Mascot results of peptides
320 containing PNGase consensus sequence are shown in *Table S1. Comparison of*
321 *OVMGs' N-glycosylation sites are included in Table S2.* These information are
322 available free of charge via the Internet at <http://pubs.acs.org>.

323

324

325 AUTHOR INFORMATION

326 Corresponding Author

327 *Meihu Ma: Phone, +86-27-87283177; Fax, +86-27-87283177; E mail,
328 mameihuhn@163.com.

329 Funding

330 Research was supported by Special Fund for Agro-scientific Research in the Public
331 Interest (201303084), Earmarked Fund for Modern Agro-industry Technology

332 Research System (CARS-41-K23), and Fundamental Research Funds for the Central
333 Universities (2014PY050).

334 **Notes**

335 The authors declare no competing financial interest

336

337

338 **ABBREVIATIONS USED**

339 2-DE, two-dimensional electrophoresis; A2MG, α 2-macroglobulin; ACN, acetonitrile;
340 CID, collision induced decomposition; FWHM, full width at half maximum; IEF,
341 isoelectric focusing ; MW, molecular weight; NMWL, nominal molecular weight limit;
342 PEG, polyethylene glycol ; OVMG, ovomacroglobulin.

REFERENCE

- (1) Robinson, D.; Monsey, J. Changes in the composition of ovomucin during liquefaction of thick egg white. *J. Sci. Food Agric.* **1972**, *23*, 29-38.
- (2) Kato, A.; Ogino, K.; Kuramoto, Y.; Kobayashi, K. Degradation of the O-glycosidically linked carbohydrate units of ovomucin during egg white thinning. *J. Food Sci.* **1979**, *44*, 1341-1344.
- (3) Kato, A.; Sato, Y. The release of carbohydrate rich component from ovomucin gel during storage. *Agric. Biol. Chem.* **1972**, *36*, 831-836.
- (4) Kato, A.; Imoto, T.; Yagishita, K. The binding groups in ovomucin-lysozyme interaction. *Agric. Biol. Chem.* **1975**, *39*, 541-544.
- (5) Mine, Y.; Yang, M. Recent advances in the understanding of egg allergens: basic, industrial, and clinical perspectives. *J. Agric. Food Chem.* **2008**, *56*, 4874-4900.
- (6) Matsuda, T.; Nakamura, R.; Nakashima, I.; Hasegawa, Y.; Shimokata, K. Human IgE antibody to the carbohydrate-containing third domain of chicken ovomucoid. *Biochem. Biophys. Res. Commun.* **1985**, *129*, 505-510.
- (7) Rupa, P.; Nakamura, S.; Katayama, S.; Mine, Y. Effects of ovalbumin glycoconjugates on alleviation of orally induced egg allergy in mice via dendritic - cell maturation and T - cell activation. *Mol. Nutr. Food Res.* **2014**, *58*, 405-417.
- (8) Rupa, P.; Nakamura, S.; Katayama, S.; Mine, Y. Attenuation of Allergic Immune Response Phenotype by Mannosylated Egg White in Orally Induced Allergy in Balb/c Mice. *J. Agric. Food Chem.* **2014**, *62*, 9479-9487.
- (9) Shan, Y.; Xu, Q.; Ma, M. Mg ²⁺ binding affects the structure and activity of ovomucin. *Int. J. Biol. Macromol.* **2014**, *70*, 230-235.
- (10) Robinson, D.; Monsey, J. Studies on the composition of egg-white ovomucin. *Biochem. J* **1971**, *121*, 537-547.
- (11) Offengenden, M.; Fentabil, M. A.; Wu, J. N-glycosylation of ovomucin from hen egg white. *Glycoconj. J.* **2011**, *28*, 113-123.
- (12) Yamashita, K.; Kamerling, J.; 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.* **1982**, *257*, 12809-12814.
- (13) Kato, I.; Schrode, J.; Kohr, W. J.; Laskowski Jr, M. Chicken ovomucoid: determination of its amino acid sequence, determination of the trypsin reactive site, and preparation of all three of its domains. *Biochemistry* **1987**, *26*, 193-201.
- (14) <http://www.uniprot.org/uniprot/P01005>.
- (15) Stein, P. E.; Leslie, A. G.; Finch, J. T.; Carrell, R. W. Crystal structure of uncleaved ovalbumin at 1·95 Å resolution. *J. Mol. Biol.* **1991**, *221*, 941-959.
- (16) Harvey, D.; Wing, D.; Küster, B.; Wilson, I. Composition of N-linked carbohydrates from ovalbumin and co-purified glycoproteins. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 564-571.
- (17) Nielsen, K. L.; Sottrup-Jensen, L.; Nagase, H.; Thøgersen, H. C.; Etzerodt, M. Amino acid sequence of hen ovomacroglobulin (ovostatin) deduced from cloned cDNA. *Mitochondrial DNA* **1994**, *5*, 111-119.
- (18) Mann, K.; Gautron, J.; Nys, Y.; McKee, M. D.; Bajari, T.; Schneider, W. J.; Hincke, M. T. Disulfide-linked heterodimeric clusterin is a component of the chicken eggshell matrix and egg white. *Matrix Biol.* **2003**, *22*, 397-407.

- (19) Craig-Barnes, H. A.; Doumouras, B. S.; Palaniyar, N. Surfactant protein D interacts with α 2-macroglobulin and increases its innate immune potential. *J. Biol. Chem.* **2010**, *285*, 13461-13470.
- (20) <http://www.uniprot.org/uniprot/P20740>.
- (21) Geng, F.; Huang, X.; Yan, N.; Jia, L.; Ma, M. Purification of hen egg white ovomacroglobulin using one-step chromatography. *J. Sep. Sci.* **2013**, *36*, 3717-3722.
- (22) Signor, L.; Erba, E. B. Matrix-assisted Laser Desorption/Ionization Time of Flight (MALDI-TOF) Mass Spectrometric Analysis of Intact Proteins Larger than 100 kDa. *Journal of visualized experiments: JoVE* **2013**, *79*, 50635, doi:10.3791/50635.
- (23) Jiang, K.; Wang, C. J.; Sun, Y. J.; Liu, Y.; Zhang, Y.; Huang, L. J.; Wang, Z. F. Comparison of Chicken and Pheasant Ovotransferrin N-Glycoforms via Electrospray Ionization Mass Spectrometry and Liquid Chromatography Coupled with Mass Spectrometry. *J. Agric. Food Chem.* **2014**, *62*, 7245-7254.
- (24) Melo-Braga, M. N.; Verano-Braga, T.; León, I. R.; Antonacci, D.; Nogueira, F. C.; Thelen, J. J.; Larsen, M. R.; Palmisano, G. Modulation of protein phosphorylation, N-glycosylation and Lys-acetylation in grape (*Vitis vinifera*) mesocarp and exocarp owing to *Lobesia botrana* infection. *Mol. Cell. Proteomics* **2012**, *11*, 945-956.
- (25) Qiu, N.; Ma, M.; Cai, Z.; Jin, Y.; Huang, X.; Huang, Q.; Sun, S. Proteomic analysis of egg white proteins during the early phase of embryonic development. *J. Proteomics* **2012**, *75*, 1895-1905.
- (26) Qiu, N.; Ma, M.; Zhao, L.; Liu, W.; Li, Y.; Mine, Y. Comparative Proteomic Analysis of Egg White Proteins under Various Storage Temperatures. *J. Agric. Food Chem.* **2012**, *60*, 7746-7753.
- (27) Liu, W.; Xie, Y.; Ma, J.; Luo, X.; Nie, P.; Zuo, Z.; Lahrmann, U.; Zhao, Q.; Zheng, Y.; Zhao, Y. IBS: an illustrator for the presentation and visualization of biological sequences. *Bioinformatics* **2015**, *btv362*, doi: 10.1093/bioinformatics/btv362.
- (28) Kitamoto, T.; Nakashima, M.; Atsushi, I. Hen egg white ovomacroglobulin has a protease inhibitory activity. *J. Biochem.* **1982**, *92*, 1679-1682.
- (29) Nagase, H.; Harris, E.; Woessner, J.; Brew, K. Ovostatin: a novel proteinase inhibitor from chicken egg white. I. Purification, physicochemical properties, and tissue distribution of ovostatin. *J. Biol. Chem.* **1983**, *258*, 7481-7489.
- (30) Saxena, I.; Tayyab, S. Protein proteinase inhibitors from avian egg whites. *Cellular and Molecular Life Sciences CMLS* **1997**, *53*, 13-23.
- (31) Kitamoto, T.; Nakashima, M.; Ikai, A. HEN EGG-WHITE OVOMACROGLOBULIN HAS A PROTEASE INHIBITORY ACTIVITY. *Journal of Biochemistry* **1982**, *92*, 1679-1682.
- (32) Nagase, H.; Harris, E. D.; Woessner, J. F.; Brew, K. OVOSTATIN - A NOVEL PROTEINASE-INHIBITOR FROM CHICKEN EGG-WHITE .1. PURIFICATION, PHYSICOCHEMICAL PROPERTIES, AND TISSUE DISTRIBUTION OF OVOSTATIN. *Journal of Biological Chemistry* **1983**, *258*, 7481-7489.
- (33) Mine, Y.; D'Silva, I. Bioactive components in egg white. In *Egg bioscience and biotechnology*, Mine, Yoshinori ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2008; pp 141-184.
- (34) Al-Ghoul, A.; Johal, R.; Sharquie, I. K.; Emara, M.; Harrington, H.; Shakib, F.; Ghaemmaghami, A. M. The glycosylation pattern of common allergens: the recognition and uptake of Der p 1 by epithelial and dendritic cells is carbohydrate dependent. *PLoS One* **2012**, *7*, e33929.
- (35) Underhill, A. P. Current issues in Chinese Neolithic archaeology. *Journal of World Prehistory* **1997**, *11*, 103-160.

- 429 (36) Larson, G.; Fuller, D. Q. The evolution of animal domestication. *Annual Review of Ecology,*
430 *Evolution, and Systematics* **2014**, *45*, 115-136.
- 431 (37) Thornton, E. K.; Emery, K. F.; Steadman, D. W.; Speller, C.; Matheny, R.; Yang, D. Earliest
432 Mexican Turkeys (*Meleagris gallopavo*) in the Maya region: implications for pre-hispanic animal
433 trade and the timing of Turkey domestication. *PLoS One* **2012**, *7*, e42630.
- 434 (38) Rong, J.; Lammers, Y.; Strasburg, J. L.; Schidlo, N. S.; Ariyurek, Y.; de Jong, T. J.; Klinkhamer, P.
435 G.; Smulders, M. J.; Vrieling, K. New insights into domestication of carrot from root transcriptome
436 analyses. *BMC Genomics* **2014**, *15*, 895.
- 437 (39) Nagase, H.; Harris, E. Ovostatin: a novel proteinase inhibitor from chicken egg white. II.
438 Mechanism of inhibition studied with collagenase and thermolysin. *J. Biol. Chem.* **1983**, *258*,
439 7490-7498.
- 440
- 441

442

Figure Captions

443

444

445 **Figure 1** MALDI-TOF-MS analysis of OVMG subunit. OVMG was reduced by DTT
446 and mixed with matrix solution, then ionized under a laser and measured by TOF MS
447 in a linear mode. Four ion signals peaks was observed at m/z 44791, 62154, 92203,
448 and 183297, which respectively correspond to $([M+4H]^{4+})$, $([M+3H]^{3+})$, $([M+2H]^{2+})$,
449 and $([M+H]^+)$ of OVMG, thereby the MW of OVMG subunit is 183.3 kDa.

450

451 **Figure 2** Location of N-glycosylation sites on OVMG and two representative MS/MS
452 spectrums of peptides which contain the PNGase consensus sequence. A, Location of
453 identified N-glycosylation sites on the sequence of OVMG, domains of OVMG were
454 assigned by Pfam (<http://pfam.xfam.org/>); N¹²²¹ was partly glycosylated, N³⁴² was not
455 covered in the analysis. B, MS/MS spectrum of peptide ⁶⁵NLNQITISVRVVL⁷⁶, in
456 which the glycosylated asparagine (Asn, N⁶⁷) is converted to aspartic acid (Asp), the
457 mass difference between y10 and y9 (or b3 and b2) is 115 Da. C, MS/MS spectrum of
458 peptide ¹²¹⁹KPNRSQEDLTAKSAIVQW¹²³⁶, which does not glycosylated on its
459 potential site (N¹²²¹), the mass difference between y15 and y16 is 114 Da.

460

461 **Figure 3** Comparison of OVMG N-glycosylation sites with its homologous proteins.
462 A, Location of N-glycosylation sites on the sequence of OVMGs from chicken
463 (P20740, *Gallus gallus*), turkey (G1NMI3, *Meleagris gallopavo*) and duck (R0J9X6,
464 *Anas platyrhynchos*). B, the Venn diagram illustrating the number of conserved
465 N-glycosylation sites in the three proteins.

466

467 **Figure 4** Visual illustration of the heterogeneity of OVMG glycosylation using 2-DE.
468 OVMG was separated by 2-DE using 24 cm (pH 4–7) IEF strips and 10 % SDS–
469 PAGE gels. Spots were identified using MALDI-TOF MS/MS. Spots 1-5, and 7 were
470 identified as OVMG, spot 6 was identified as ovalbumin.
471

472 **Table 1** N-Glycosylation sites of OVMG identified by HPLC-ESI MS/MS

Sites	Sequence	Peptide position	Mr (calc)	m/z Obs	Glycosylation
N ⁶⁷	NLNQ <u>T</u> ISVRVVL (CT ¹)	65-76	1355.78	678.90 ²⁺	+
N ⁸²	VVLEYDTINTTIFEK (T ²)	74-88	1784.91	893.46 ²⁺	+
	EYDTINTTIF (CT)	77-86	1216.55	609.28 ²⁺	+
	YDTINTTIFE (G-C ³)	78-87	1216.55	609.28 ²⁺	+
N ⁸⁹	EKN <u>T</u> TTSNGLQCL (CT)	87-99	1465.67	733.85 ²⁺	+
N ¹⁹¹	MYPLIAVQDPQNNRIFQWQ <u>N</u> VTSE (G-C)	172-195	2891.39	964.81 ³⁺	+
	IFQWQ <u>N</u> VTSEINIVQIEFPLTEEPILGNYK (T)	186-215	3562.82	1188.63 ³⁺	+
N ⁴⁰³	VNNKNTHN <u>F</u> TTDINGIAPF (CT)	395-414	2117.02	1059.52 ²⁺	+
	NTHN <u>F</u> TTDINGIAPFSIDTSK (T)	400-420	2293.09	1147.56 ²⁺	+
N ⁵²⁷	TGEIKVNIQADQ <u>N</u> GTF (CT)	515-530	1734.84	868.43 ²⁺	+
N ⁵⁸⁸	QMLTTSN <u>V</u> SLVIEAAANSFCAVR (T)	582-604	2482.22	828.42 ³⁺	+
N ⁷⁵⁷	EFFPETWIWDIILIN <u>S</u> TGK (T)	743-761	2309.16	1155.59 ²⁺	+
	IWDIILIN <u>S</u> TGKASVSY (CT)	750-766	1879.99	941.01 ²⁺	+
	DIILIN <u>S</u> TGKASVSY (CT)	752-766	1580.83	791.43 ²⁺	+
	IILIN <u>S</u> TGKASVSYTIPDTITE (G-C)	753-774	2336.24	1169.13 ²⁺	+
	IN <u>S</u> TGKASVSY(CT)	756-766	1126.55	564.28 ²⁺	+

N ¹¹⁴¹	CLETASEK <u>NIT</u> DIY (CT)	1133-1146	1656.76	829.39 ²⁺	+
N ¹²²¹	ITSYVLLALLYK <u>PNRS</u> QE (G-C)	1208-1225	2107.17	703.40 ³⁺	/
	K <u>PNRS</u> QEDLTKASAIVQW (CT)	1219-1236	2070.09	1036.06 ²⁺	/
	K <u>PNRS</u> QEDLTKASAIVQW (CT)	1219-1236	2071.07	1036.54 ²⁺	+
N ¹³¹⁵	YSLTV <u>NGT</u> GCVLIQTALR (T)	1310-1327	1966.02	984.02 ²⁺	+
	SLTV <u>NGT</u> GCVL (CT)	1311-1321	1120.54	561.26 ²⁺	+
N ¹³⁴⁷	YNIHLPEGAFGFSLSVQTS <u>NASC</u> PR (T)	1328-1352	2752.29	918.44 ³⁺	+
	GAFGFSLSVQTSNASCPRDQPGKFD (G-C)	1335-1359	2673.21	892.08 ³⁺	+
	SLSVQTS <u>NASC</u> PRDQPGKF (CT)	1340-1358	2078.97	1040.50 ²⁺	+
	SVQTS <u>NASC</u> PRDQPGKF (CT)	1342-1358	1878.86	940.44 ²⁺	+

473 1, peptides identified from chymotrypsin digestion of OVMG;

474 2, peptides identified from trypsin digestion of OVMG;

475 3, peptides identified from Glu-C digestion of OVMG.

476

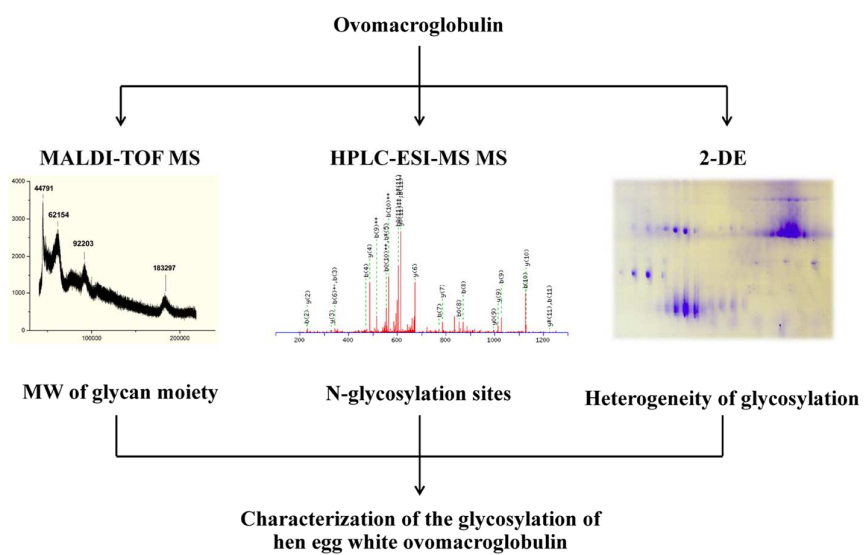
477 **Table 2** Identification of protein spots in 2-DE by MALDI-TOF MS/MS.

Spot	gi No.	Protein name	Score	Matched peptides / Sequence coverage	Exptl pI / Mr (kDa)	Theor pI / Mr (kDa)
1	gi 45382565	ovostatin [Gallus gallus]*	367	10/6%	5.18/178.7	5.64/165.5
2	gi 45382565	ovostatin [Gallus gallus]	342	15/12%	5.24/176.4	5.64/165.5
3	gi 45382565	ovostatin [Gallus gallus]	346	13/10%	5.30/174.5	5.64/165.5
4	gi 45382565	ovostatin [Gallus gallus]	358	9/7%	5.95/175.3	5.64/165.5
5	gi 45382565	ovostatin [Gallus gallus]	364	16/11%	6.03/170.0	5.64/165.5
6	gi 223299	ovalbumin [Gallus gallus]	462	11/42%	5.07/123.2	5.20/43.1
7	gi 45382565	ovostatin [Gallus gallus]	342	18/14%	5.30/100.0	5.64/165.5

478 *Ovomacroglobulin also named “ovostatin”

479

TABLE OF CONTENTS



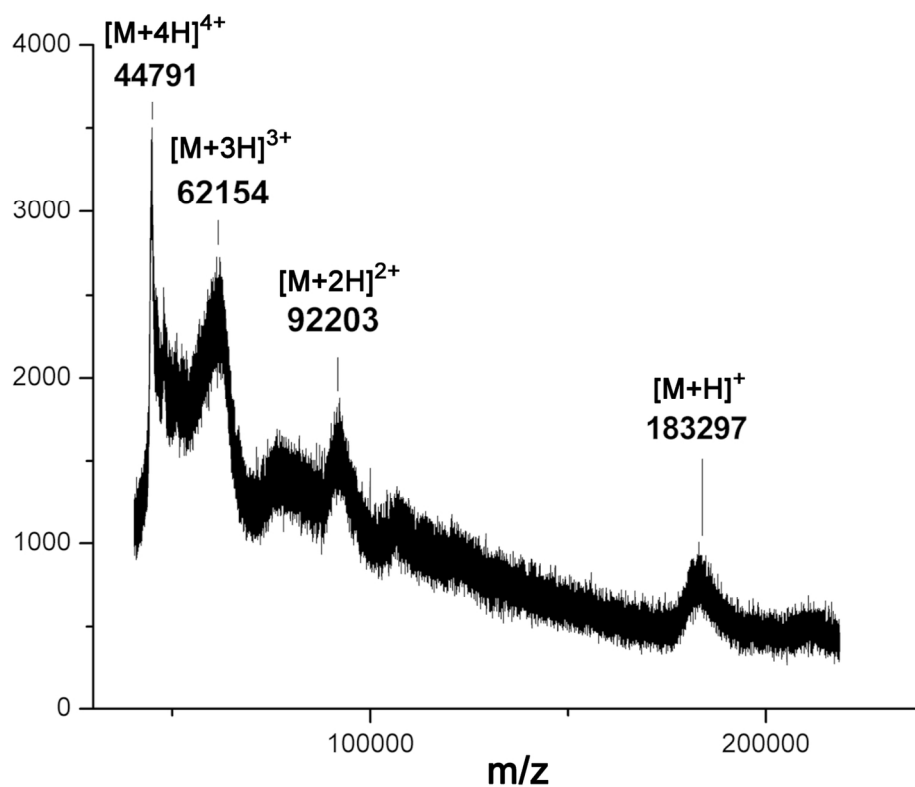


Figure 1 MALDI-TOF-MS analysis of OVMG subunit. OVMG was reduced by DTT and mixed with matrix solution, then ionized under a laser and measured by TOF MS in a linear mode. Four ion signals peaks was observed at m/z 44791, 62154, 92203, and 183297, which respectively correspond to ([M+4H]⁴⁺), ([M+3H]³⁺), ([M+2H]²⁺), and ([M+H]⁺) of OVMG, thereby the MW of OVMG subunit is 183.3 kDa.
69x55mm (600 x 600 DPI)

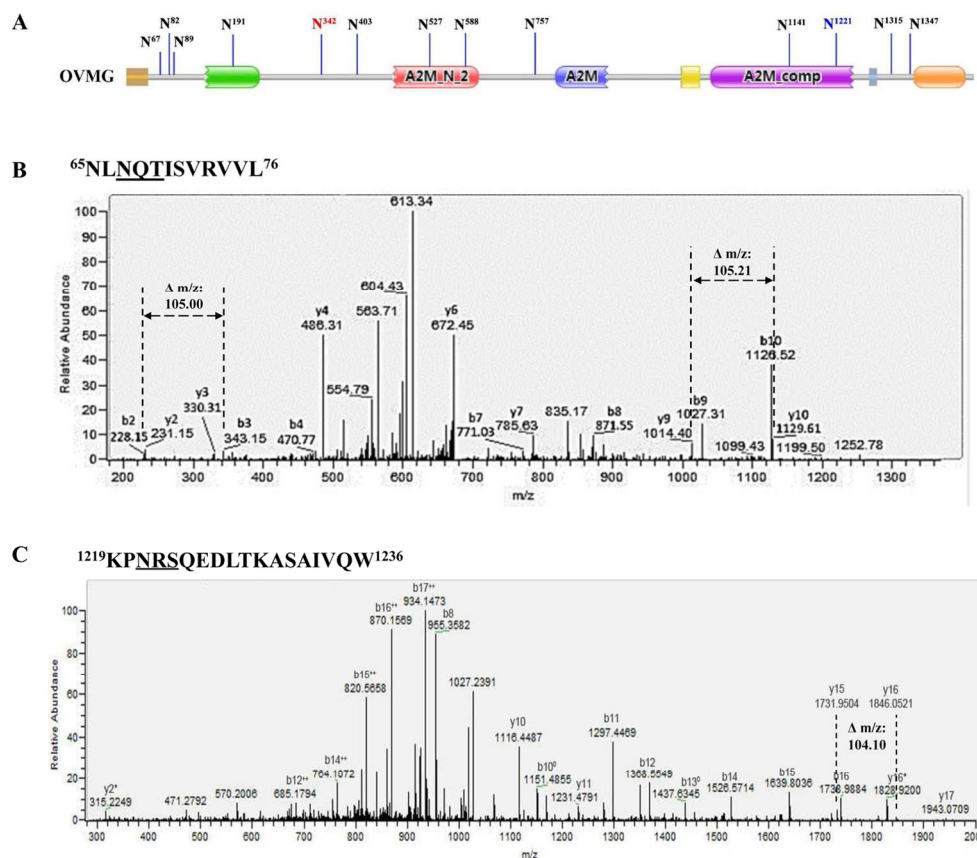


Figure 2 Location of N-glycosylation sites on OVMG and two representative MS/MS spectrums of peptides which contain the PNGase consensus sequence. A, Location of identified N-glycosylation sites on the sequence of OVMG, domains of OVMG were assigned by Pfam (<http://pfam.xfam.org/>); N1221 was partly glycosylated, N342 was not covered in the analysis. B, MS/MS spectrum of peptide 65NLNQTISVRVVL76, in which the glycosylated asparagine (Asn, N67) is converted to aspartic acid (Asp), the mass difference between y10 and y9 (or b3 and b2) is 115 Da. C, MS/MS spectrum of peptide 1219KPNRSQEDLTAKSAIVQW1236, which does not glycosylated on its potential site (N1221), the mass difference between y15 and y16 is 114 Da.

129x111mm (300 x 300 DPI)

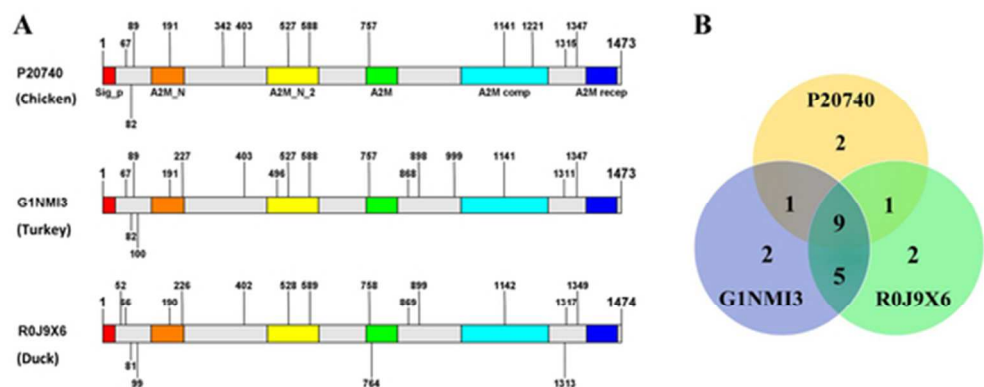


Figure 3 Comparison of OVMG N-glycosylation sites with its homologous proteins. A, Location of N-glycosylation sites on the sequence of OVMGs from chicken (P20740, *Gallus gallus*), turkey (G1NMI3, *Meleagris gallopavo*) and duck (R0J9X6, *Anas platyrhynchos*). B, the Venn diagram illustrating the number of conserved N-glycosylation sites in the three proteins.

49x20mm (300 x 300 DPI)

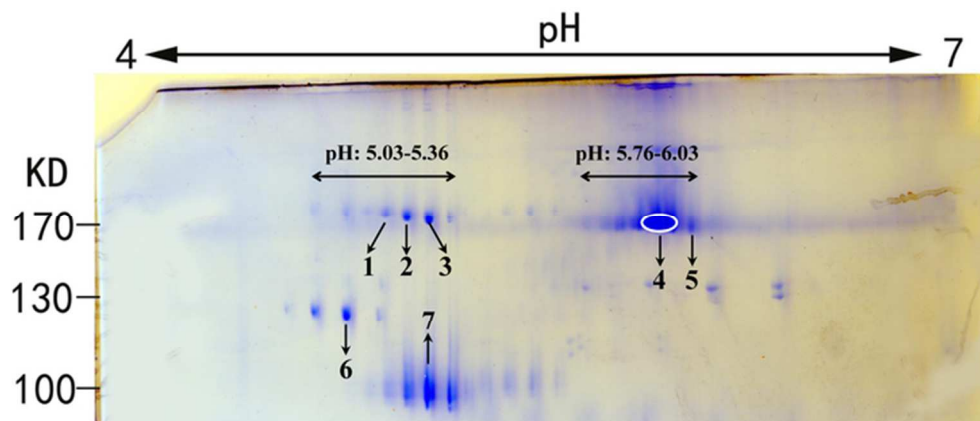


Figure 4 Visual illustration of the heterogeneity of OVMG glycosylation using 2-DE. OVMG was separated by 2-DE using 24 cm (pH 4–7) IEF strips and 10 % SDS–PAGE gels. Spots were identified using MALDI–TOF MS/MS. Spots 1–5, and 7 were identified as OVMG, spot 6 was identified as ovalbumin.
62x26mm (300 x 300 DPI)

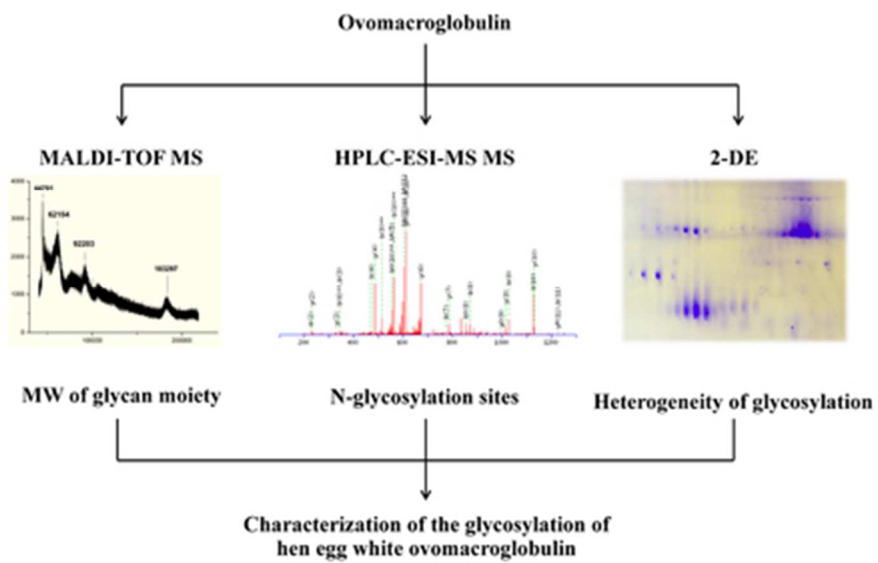


TABLE OF CONTENTS
39x24mm (300 x 300 DPI)