Anal Bioanal Chem (2010) 397:205–214
DOI 10.1007/s00216-009-3326-3

ORIGINAL PAPER

Determination of insoluble avian eggshell matrix proteins

Ivan Mikšík · Pavla Sedláková · Katerina Lacinová · Statis Pataridis · Adam Eckhardt

Received: 17 September 2009 / Revised: 3 November 2009 / Accepted: 16 November 2009 / Published online: 9 December 2009
© Springer-Verlag 2009

Abstract The organic components of bones and other mineralized tissues have a high impact on the organization and deposition of calcium, and consequently influence the mechanical properties of those tissues. The extractable proteins of avian eggshells have been studied extensively and many of them have been identified; insoluble (non-extractable) proteins have been sparsely studied, however. In the work discussed in this paper we studied EDTA-insoluble proteins by gradual decalcification of eggshell with EDTA. The insoluble proteinaceous films were chemically treated with cyanogen bromide and the mixtures of large fragments obtained were gradually precipitated with salt. The separated fractions were digested with trypsin and analyzed by HPLC–MS–MS (ion trap mass spectrometer). Analysis of the entire eggshell matrix (without precipitation steps) only enabled 6 proteins to be determined (ovocalyxins 32 and 36, ovocleidin 17 and 116, clusterin, and ovalbumin). Pretreatment of the individual eggshell layers and gradual precipitation with salt markedly increased the number of proteins identified – 28 proteins were determined. We identified for the first time collagens I (two chains) and III in the eggshell matrix, and Kunitz-like protease inhibitor as a major shell matrix protein. Besides the above mentioned proteins we can also mention EDIL3, fibronectin, sulfhydryl oxidase, tubulin alpha 1, lysozyme, Dickkopf-related protein 3, keratins, and ovotransferrin. The relative abundances of proteins in all eggshell layers were determined using the exponentially modified protein abundance index (emPAI). In the cuticle layer seven proteins were identified, whereas 16 proteins were described in the palisade layer and 23 in the mammillary layer.

Keywords Biological samples · Eggshell · Insoluble proteins · Matrix proteins · Chicken proteins

I. Mikšík · P. Sedláková · K. Lacinová · S. Pataridis · A. Eckhardt
Institute of Physiology, v.v.i,
Academy of Sciences of the Czech Republic,
Videnska 1083,
14220 Prague 4, Czech Republic

I. Mikšík (✉) · A. Eckhardt
Centre for Cardiovascular Research,
Videnska 1083,
14200 Prague 4, Czech Republic
e-mail: miksik@biomed.cas.cz

Introduction

The organic components of bones and other mineralized tissues have a high impact on the organization and deposition of calcium, and, consequently, influence the mechanical properties of those tissues. Avian eggshells have a relatively simple structure: the outermost layer is a relatively thin cuticle, followed by a thick calcified layer composed of calcite, which forms crystalline structures, termed “palisades”, that are terminated by rounded cones, named “mammilllae”. The tips of the mammillary cones serve as anchor points for the fibres of the shell membranes that envelop the albumen [1, 2].

The egg is composed of a central yolk surrounded by the albumen, eggshell membranes, and calcified shell and is progressively formed in specialized regions of the oviduct. After ovulation, the yolk/ovum complex travels through the longest part of the oviduct (magnum) and progressively acquires the water, ions, and proteins that compose the albumen (egg white) during a two to three-hour period [3]. Eggshells are formed during movement along the oviduct by producing a multilayered mineral–organic composite [4]. Shell membranes are formed in the isthmus region of

the oviduct between the 3:00 and 5:00 h postoviposition (p.o.). The mammillary layer is formed in the red isthmus (5:15 and 5:30 h p.o.) and the palisade layer forms in the shell gland (6:00 and 18:00 p.o.). Finally, 1.5 h before oviposition, (egg expulsion) shell pigments and eggshell cuticle are deposited on the surface.

The extractable proteins of the eggshell matrix have been studied extensively and many of them have been identified. Mann et al. [5] described 520 proteins from an acid-soluble organic matrix of the calcified chicken eggshell layer. Recently Mann et al. [6] listed 119 proteins in egg yolk, 78 proteins in egg white, and 528 proteins in the decalcified eggshell organic matrix, and Farinazzo et al. [7] extended the number of yolk proteins to 255. Eggshell-specific matrix proteins were recently reviewed [3]. These proteins are designated "ovocleidins" and "ovocalyxins" and their main functions are regulation of eggshell mineralization and antimicrobial defence. Matrix proteins can modify the growth of calcium carbonate crystals in biomaterials and numerous papers describe this function in the formation of eggshell [8]; a review is available elsewhere [3]. The antimicrobial protective properties of a protein extract from eggshell has been demonstrated in a few papers [9–11]. However, the insoluble (non-extractable) proteins have been only sparsely studied [12] and only six have been described: ovocalyxin-32 was found mainly in the outer layer (the cuticle); ovocleidin-116 and 17 and ovocalyxin-36 were found throughout the whole eggshell; clusterin was also present throughout the whole eggshell apart from the cuticle; and ovalbumin was only found in the inner layer, the mammillary.

It must be stressed that the vast majority of work has been done on chicken (*Gallus gallus*) eggs, however similar proteins can be found in other bird species, for example as ansocalcin in goose eggs with a high homology with ovocleidin-17 [13–15] or in emu eggshell matrix proteins [16].

The objective of the work discussed in this paper is to determine a more complete set of insoluble (EDTA-insoluble) proteins in individual eggshell layers.

Materials and methods

Instrumentation

The HPLC–MS apparatus used was an Agilent 1100 LC/MSD system (Agilent, Palo Alto, CA, USA) consisting of a degasser, a binary pump, an autosampler, a thermostatted column compartment, and a diode-array detector. The instrument was controlled, and the data collected and manipulated by the software ChemStation A.06.03. It was coupled to an ion-trap mass spectrometer (Agilent LC-MSD Trap XCT-Ultra); details of the instrument conditions are given in the section *Conditions for HPLC–MS*.

Analysis of MS–MS data (peptide/protein identification) was carried out using the software SpectrumMill (v. 3.02, Agilent). Initial searches were performed in the IPI-chicken database (http://www.ebi.ac.uk/IPI/IPIhelp.html) [17] followed by searches in the SwissProtdatabase. Searches for collagens were performed on the data extracted (for collagen) from SwissProt, where the variable modification was set for the hydroxylation of proline and lysine.

Chemicals

Trypsin (type IX-S from porcine pancreas, E.C. 3.4.21.4, 15,450 units per mg) and pepsin (Pepsin A, E.C. 3.4.23.1, activity 3,460 units per mg protein), cyanogen bromide (CNBr) ammonium bicarbonate, dithiothreitol (DTT), and iodoacetic acid were obtained from Sigma (St Louis, MO, USA), Tris [Tris(hydroxymethyl)aminomethane] and NaCl were purchased from Lachema (Brno, Czech Republic). Ethylenediaminetetraacetic acid disodium salt (EDTA, Titriplex III), 2-mercaptoethanol, and HPLC gradient-grade acetonitrile were from Merck (Darmstadt, Germany). All solutions were prepared in MilliQ water (Millipore, Bedford, MA, USA).

The eggs used (5) in the experiments were commercially available hens' eggs (Ceska vejce CZ, Praha, Czech Republic) of the same age. However, no influence of nutrition was studied.

Econo-Pac 10 DG columns were obtained from Bio-rad (Bio-rad Laboratories, Hercules, CA, USA)

Sample preparation

Preparation of eggshell fractions

Preparation of the various insoluble layers utilised a modified version of a previously published method [12]. Whole eggs were washed with water and methanol and three types of sample were prepared.

Cuticle layer Eggs were treated with 0.13 mol L^{-1} EDTA (pH 7.6) containing 10 mmol L^{-1} 2-mercaptoethanol (three times the egg volume) for 60 min at room temperature. The resulting insoluble organic layer left on the egg surface after this partial decalcification was scraped off, collected by washing with water, and then centrifuged (1,000*g*, 15 min, room temperature). The resulting pellet was resuspended in water and centrifuged under the above conditions (repeated three times) and then lyophilized. With this treatment, 10% of the shell weight was removed (determined by weighing).

Palisade layer In the next step, the egg that had undergone step A was treated with 0.6 mol L^{-1} EDTA (pH 7.6) containing 10 mmol L^{-1} 2-mercaptoethanol (three times the egg volume) for 150 min at laboratory temperature. The insoluble material (layer) on the eggs was scraped off and the material was subjected to the same procedure as described for the cuticle layer. With this treatment, an additional 80% of the original shell weight was removed (determined by weighing).

Mammillary layer The remaining egg material was again treated with 0.6 mol L^{-1} EDTA (pH 7.6) containing 10 mmol L^{-1} 2-mercaptoethanol (three times the egg volume), but for a prolonged period of time (overnight) at laboratory temperature. After this procedure, only the inner egg content (egg white and yolk covered by soft membranes) remained intact. In this treatment, the rest of the insoluble proteins from the palisade layer and cones (mammillary knob layer) were obtained. With this treatment, the last 10% of the original shell weight was removed (determined by weighing). Shell membranes were not analyzed.

Cleavage of proteins

The first cleavage of proteins was done with CNBr. Samples of individual layers (5 mg mL^{-1}) were incubated in 0.2 mol L^{-1} ammonium bicarbonate, pH 7.0, containing 25% (*v*/*v*) 2-mercaptoethanol to reduce the oxidized methionyl residues and after lyophilisation the samples were cleaved with CNBr in 70% (*v*/*v*) formic acid under nitrogen. The samples were then repeatedly dried (with a controlled warm air flow) and reconstituted in water (three times). This leads to relatively large fragments because CNBr splits proteins specifically at methionine. The mixtures obtained from these peptides were gradually precipitated with salt (1, 2, 3, and 4 mol L^{-1} NaCl) (Scheme 1). These fractions were reduced and alkylated by use of DTT and iodoacetic acid: samples were suspended in 0.5 mL 6.0 mol L^{-1} guanidine HCl, 1.2 mol L^{-1} Tris-HCl, 2.5 mmol L^{-1} Na_2EDTA (pH 8.4) buffer, then their disulfide bridges were reduced by adding 25 μL 1.0 mol L^{-1} DTT and maintaining the temperature at 65 °C for 30 min. *S*-Carboxymethylation (alkylation) was carried out by adding 60 μL 1.0 mol L^{-1} iodoacetic acid and placing in the dark at room temperature for 40 min. The alkylation was stopped by adding 15 μL 1.0 mol L^{-1} DTT (dithiothreitol) and the samples were then desalted and transferred to the digestion buffer by applying the reaction mixture to Econo-Pac 10 DG columns equilibrated with pH 7.8, 20 mmol L^{-1} NH_4HCO_3 digestion buffer. The carboxymethylated protein was eluted and collected with the digestion buffer. The treated proteins were digested by adding trypsin (1:50 ratio enzyme:substrate) and incubated at 37 °C for 3 h. Blank samples were prepared by incubating the enzyme solution alone under identical conditions. The samples were then centrifuged for 5 min at 2000*g* and the supernatants removed and stored at −18 °C.

The insoluble pellet after trypsin cleavage of the 1 mol L^{-1} NaCl precipitate was finally digested with pepsin (1:50 ratio enzyme:substrate) at 37 °C for 3 h in 0.5 mol L^{-1} acetic acid, pH 2) [18].

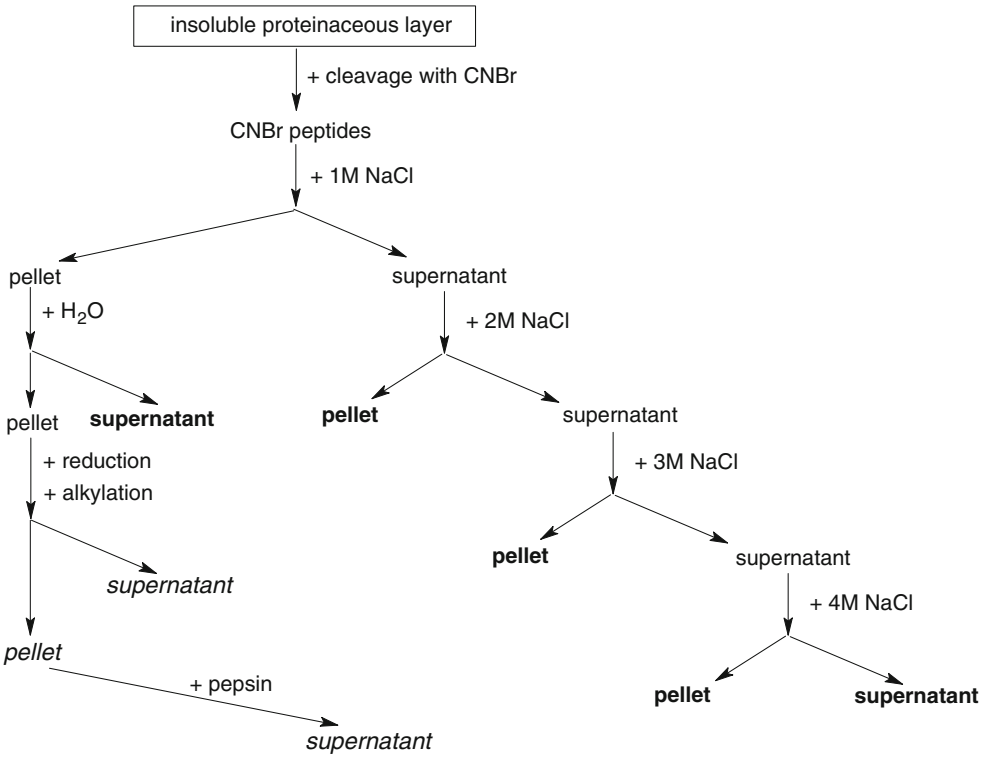
Conditions for HPLC–MS

Chromatographic separation was carried out in a Jupiter 4 μm Proteo 90A column (250×2 mm I.D., Phenomenex, Torrence, CA, USA). A 20-μL sample was injected. Elution was achieved using a linear gradient (A = water with 0.1% formic acid, and B = acetonitrile with 0.085% formic acid). Separation was initiated by running the system isocratically for two minutes with 2% mobile phase B, followed by a gradient elution to 35% B at 40 min. The column was finally eluted with 100% B for 10 min. Equilibration before the next run was achieved by washing with buffer A for 10 min. The flow-rate was 0.25 mL min^{-1}, the column temperature was held at 25 °C and UV absorbance detection was done at 214 nm.

Atmospheric pressure ionization–electrospray ionization (API–ESI) positive mode ion-trap mass spectrometry was used. Operating conditions: drying gas (N_2), 10 $Lmin^{-1}$; drying gas temperature, 350 °C; nebulizer pressure, 172.37 kPa (25 psi); ions were observed over the mass range *m*/*z* 100–2200 (MS – standard mode, MS–MS – enhanced mode). Analysis was done in auto MS–MS mode (10 precursor ions, excluded after two spectra for 0.5 min).

Analysis of MS–MS data

Analysis of MS–MS data (peptide/protein identification and searching for possible post-translational modifications) was carried out using SpectrumMill software (v.3.3; Agilent). The Spectrum-Mill autovalidation of spectra was performed using the default settings, but all spectra were then evaluated manually. The first search was made in the database IPI-chicken (v.3.45, downloaded November 5, 2008) [17] when the fixed modification was carboxymethylation, variable modification was oxidation of methionine, and maximum number of missed cleavages was 2. Another search was made against an in-house set of data for collagen extracted from the SwissProt database where the fixed modification was carboxymethylation and variable modifications were hydroxylations of proline and lysine [19]. The last search was made in the whole of the SwissProt database where the fixed modification was carboxymethylation and there was no variable modification.



Scheme 1 Schematic diagram of sample preparation of proteins in eggshell layers. Analyzed products in *bold* were finally reduced, alkylated, and cleaved with trypsin. Analyzed products in *italics* were trypsinized only (reduction and alkylation were performed earlier)

All searches were performed with CNBr/trypsin-specific cleavage (digestion).

The relative abundances of proteins were determined using the exponentially modified protein abundance index (emPAI) in all eggshell layers.

The relatively abundances of identified proteins were estimated quantitatively by calculating their exponentially modified protein abundance index (emPAI) which is 10^{PAI}–1 [5, 20], where PAI is the number of observed peptides divided by the number of theoretically observable peptides (when this index was calculated for peptides with an MH^+ of 200–4,000).

Mann et al. [5, 21] demonstrated that emPAI could be used for comparison of protein concentrations (e.g. egg white protein) within one experimental procedure and for classification of proteins as major or minor components. Of course this index is not absolute, all proteins should be digested with the same efficiency, but it can give an overview of the relative proportion of proteins studied. However, many egg shell proteins can be modified (cross-linked, glycosylated, phosphorylated) and these modifications are not covered by the present analysis.

It is always complicated to exactly determine the keratins that commonly contaminate a sample. To precisely determine the origin of keratins in the sample (eggshell) we performed sets of blank experiments—we ground a powder of calcium carbonate ($CaCO_3$) and lyophilized protein (collagen) between our fingers. We assumed that any keratins found were impurities and they were then subtracted from the sample. In principle they were human cytoskeletal keratins and dermcidin.

Results and discussion

Analyzing a complex mixture of proteins is always difficult. Our initial approaches to determining the insoluble eggshell proteins in the individual layers of chicken eggshell by HPLC–MS [12] and CE–MS [22] led to the determination of only six proteins. For this reason we tried to use another approach to determine a larger number of protein species – limited digestion (by CNBr) and gradual precipitation with salt. CNBr cleaves proteins at methionine. Because methionine is an amino acid with a relatively low abundance, the resulting peptides are longer and can be precipitated with salt. The results of this method are presented in Table 1. It is clear that this method was successful – the number of proteins identified increased markedly. Table 2 shows the proteins that could only be determined using one peptide (and for this reason their determination confidence is not significant, so this table is informative only).

According to a previously published classification [3, 5], the proteins we determined can be divided into three characteristic groups:

1. Proteins and glycoproteins which are also present in egg white. These are ovalbumin, ovotransferrin, and lysozyme

Table 1 Insoluble eggshell matrix proteins with at least two peptides. The proteins appear in descending order of their distinct summed MS–MS search score. Protein identification was accepted when the data contained at least two peptides confirmed by autovalidation and manual spectra validation and containing at least seven amino acids. New proteins not previously described in chicken eggshell are in *bold*. The first search was in the IPI-chicken database (protein name starts with IPI); the second search was in the SwissProt database (protein name starts with only one letter)

Database accession #	Protein name	Cuticle layer				Palisade layer				Mammillary layer				Summary			
		emPAI	# of peptides observed	Sequence coverage (%)	Distinct summed MS–MS search score	emPAI	# of peptides observed	Sequence coverage (%)	Distinct summed MS–MS search score	emPAI	# of peptides observed	Sequence coverage (%)	Distinct summed MS–MS search score	emPAI	# of peptides observed	Sequence coverage (%)	Distinct summed MS–MS search score
IPI00581368	LOC395256 Ovocleidin-116	0.943	15	27	259.16	1.649	22	35	349.87	1.534	21	34	336.24	1.769	23	35	376.41
IPI00590535	FN1 similar to fibronectin 1 isoform 1 preproprotein					0.357	20	10	269.99	0.219	13	6	153.56	0.464	25	12	322.05
IPI00819356	CLU 49 kDa protein (clusterin)	0.369	6	13	68.32	0.778	11	32	158.86	0.874	12	36	185.91	1.081	14	39	211.47
IPI00573506	LOC419289 Ovocalyxin-36	0.655	7	22	127.77	0.911	9	27	178.73	0.655	7	21	107.09	1.207	11	33	210.79
IPI00591158 IPI00588727	EDIL3 similar to EGF-like repeats and discoidin I-like domains 3					0.448	9	13	124.47	0.572	11	17	161.45	0.778	14	22	205.15
IPI00583974	LOC396058 Ovalbumin						1	2	6.80	0.884	11	37	169.69	0.884	11	37	169.69
IPI00681344	**LOC771972 similar to Kunitz-like protease inhibitor**	**0.840**	**9**	**29**	**154.17**	**0.501**	**6**	**24**	**107.02**	**0.501**	**6**	**19**	**86.16**	**0.840**	**9**	**29**	**156.01**
P02465	**Collagen alpha-2(I) chain OS = Bos taurus GN = COL1A2 PE=1 SV=2**		**1**	**1**	**11.50**	**0.164**	**7**	**8**	**116.10**					**0.164**	**7**	**8**	**116.10**
IPI00572756	Ovocleidin-17	1.581	7	30	71.26	0.719	4	23	60.08	1.254	6	44	72.90	2.875	10	54	109.54
IPI00571581	QSOX1 Sulfhydryl oxidase 1					0.150	4	8	58.25	0.277	7	15	105.30	0.277	7	15	105.30
P02453	**Collagen alpha-1(I) chain OS = Bos taurus GN = COL1A1 PE=1 SV=3**					**0.131**	**6**	**7**	**98.83**					**0.131**	**6**	**7**	**98.83**
IPI00578622	RARRES1 Ovocalyxin-32	0.874	6	21	76.05		1	3	10.56	0.233	2	6	28.79	0.874	6	21	78.38
IPI00575989	TUBA1C Tubulin alpha-1 chain (Fragment)					0.136	2	5	22.32	0.212	3	8	51.15	0.377	5	14	73.47
IPI00588573	MFGE8 similar to Milk fat globule-EGF factor 8 protein						1	2	18.49	0.299	5	10	67.10	0.299	5	10	67.10
IPI00600859	LYZ Lysozyme C									0.931	4	42	65.86	0.931	4	42	65.86
IPI00585935	SEMA3G similar to FLJ00014 protein					0.063	2	2	26.53	0.131	4	7	55.70	0.131	4	7	60.44
IPI00585604	189 kDa protein (von Willebrand factor; Protease inhibitor I8)									0.106	5	4	57.35	0.106	5	4	57.35
P04258	**Collagen alpha-1(III) chain OS = Bos taurus GN = COL3A1 PE=1 SV=1**					**0.089**	**3**	**4**	**56.17**	**0.089**	**3**	**4**	**52.72**	**0.089**	**3**	**4**	**57.00**
IPI00572548	**LOC395532 Collagen alpha-1(I) chain**	**0.042**	**2**	**1**	**23.54**	**0.064**	**3**	**2**	**46.32**					**0.087**	**4**	**3**	**56.56**

Table 1 (continued)

Database accession #	Protein name	Cuticle layer				Palisade layer				Mammillary layer				Summary			
		emPAI	# of peptides observed	Sequence coverage (%)	Distinct summed MS–MS search score	emPAI	# of peptides observed	Sequence coverage (%)	Distinct summed MS–MS search score	emPAI	# of peptides observed	Sequence coverage (%)	Distinct summed MS–MS search score	emPAI	# of peptides observed	Sequence coverage (%)	Distinct summed MS–MS search score
IPI00572786	NPNT similar to nephronectin short					0.194	4	8	41.52	0.142	3	6	32.30	0.248	5	10	52.16
IPI00574331	GKN2 hypothetical protein						1	13	25.86	0.274	2	12	22.98	0.438	3	26	48.84
IPI00822585	KRT4 hypothetical protein					0.151	3	2	22.09	0.151	3	3	31.30	0.207	4	4	42.34
IPI00813745	**COL1A2 Collagen alpha-2(I) chain (Fragment)**		**1**	**1**	**15.18**		**1**	**0**	**22.62**					**0.048**	**2**	**2**	**37.80**
IPI00574055	PLOD1 Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1									0.099	3	4	37.39	0.099	3	4	37.39
IPI00578016	DKK3 Dickkopf-related protein 3		1	4	15.77					0.259	3	11	36.45	0.259	3	11	36.45
IPI00573727	NUCB2 Putative uncharacterized protein		1	1	15.99	0.096	2	4	35.98		1	2	18.78	0.096	2	4	36.14
IPI00683271 IPI00578012	Ovotransferrin; LTF 78 kDa protein									0.099	3	3	30.94	0.099	3	3	30.94
IPI00682643	**TSKU Isoform 1 of Tsukushin**						**1**	**2**	**12.74**		**1**	**3**	**13.13**	**0.150**	**2**	**6**	**25.87**
IPI00685019 IPI00577021	IGLL1 24 kDa protein; IGGL1 Ig lambda chain C region									0.334	2	9	24.05	0.334	2	9	24.05
IPI00585021	SERPINBP1 similar to Ovalbumin-related protein Y									0.184	3	9	21.49	0.184	3	9	21.49
IPI00592753	KRT14 Keratin, type I cytoskeletal 14					0.105	2	3	19.13					0.105	2	3	19.13

Table 2 Insoluble eggshell matrix proteins with only one peptide. The proteins appear in descending order of their distinct summed MS–MS search score. Peptide identification was accepted when the data were confirmed by autovalidation and manual spectra validation and contained at least seven amino acids. New proteins not previously described in chicken eggshell are in *bold*

Database accession #	Protein name	Cuticle layer			Palisade layer			Mammillary layer			Summary		
		# of peptides observed	Sequence coverage (%)	Distinct summed MS–MS search score	# of peptides observed	Sequence coverage (%)	Distinct summed MS–MS search score	# of peptides observed	Sequence coverage (%)	Distinct summed MS–MS search score	# of peptides observed	Sequence coverage (%)	Distinct summed MS–MS search score
IPI00680258	**SERPINF2 similar to alpha-2-plasmin inhibitor**	**1**	**2**	**17.40**							**1**	**2**	**17.40**
IPI00586930	LAMB1 similar to laminin, beta 1							1	0	15.79	1	0	15.79
IPI00589269	SDF4 Putative uncharacterized protein	1	3	14.68							1	3	14.68
IPI00582792	**SCUBE1 similar to OTTHUMP00000028801**				**1**	**1**	**14.47**				**1**	**1**	**14.47**
IPI00599894	34 kDa protein							1	5	14.45	1	5	14.45
IPI00683494 IPI00590350 IPI00590846	HBG1 17 kDa protein; HBG1 Haemoglobin subunit beta; HBE BETA-H globin				1	6	11.49				1	6	11.49
IPI00589782	**CLSTN2 similar to calsyntenin-2**	**1**	**1**	**11.34**				**1**	**1**	**10.74**	**1**	**1**	**11.34**
IPI00581532	**NUP54 similar to nucleoporin 54 kDa**	**1**	**1**	**11.27**							**1**	**1**	**11.27**
IPI00600561	CHRD Chordin				1	1	9.77				1	1	9.77
IPI00594326	DDEF2 similar to KIAA0400							1	0	9.27	1	0	9.27

2. Proteins which are found in many tissues. These are clusterin, serum albumin, keratins, collagens, and haemoglobin
3. Eggshell-specific matrix proteins or proteoglycans. These are the C-type lectin-like protein ovocleidin-17, ovocleidin-116, ovocalyxin-32, and ovocalyxin-36.

Many of these proteins have protective (antimicrobial) functions. It is proposed that lysozyme and all eggshell-specific proteins have antimicrobial functions [3]. Another protein with peptidase inhibitor activity (serine-type endopeptidase inhibitor activity) is the Kunitz-like protease inhibitor, which is present in all layers. We have to stress that this protein (IPI00681344) is definitely not the protein previously described in Ref. [5] as "Similar to serine peptidase inhibitor, kunitz type 1" (IPI00584633) among acid-soluble eggshell proteins. This protein is identified as a major shell matrix protein for the first time. Possibly it corresponds to ovocalyxin-25, as briefly described by Gautron and Nys [23].

There were significant results concerning structural fibrillar proteins – collagens type I and III. Collagens are the main proteins of the animal extracellular matrix (both soft and "hard" tissues) [24]. To date, some 29 types of collagen have been described. Mann et al. [5] determined collagens type II, VI, VII, X, and XVIII in the acid-soluble organic matrix of chicken eggshell. It was surprising that they did not find type I collagens when this collagen type is the most abundant, being present in most tissues, but mainly in bone, tendon, and skin. It comprises approximately 95% of the entire collagen content of bone and about 80% of the total proteins present in bone. Other types of collagen, such as types III and V, are present at low levels in bone and appear to modulate the fibril diameter [25]. It has been stated that bone strength depends not only on the quantity of bone tissue but also on its quality. This quality is characterized by the geometry and shape of bones, the microarchitecture of the trabecular bones, the turnover, the mineral, and the collagen [25]. However collagen type III is often found in association with type I collagen in all soft tissues, but not in mineralized bone. Non-mineralized shell membranes contain type X collagen, which possibly inhibits mineralization [26–28]. Collagens type I and V have been immunohistologically determined in eggshell membranes [29].

We mainly found collagen type I in the palisade layer (with some content in the cuticle layer) and collagen type III in the palisade and mammillary layers. Although confidence in their determination is high, their relative abundance is lower (at least a factor of ten lower than for ovocleidin-116). From these results we can conclude that both collagen types are present in the palisade layer while type III is also present in the mammillary layer. It seems

that the occurrence of these two types are not connected to each other. However the presence of collagen—a connective tissue protein—in the calcified tissue is not surprising (see bone) and can be predicted. We can speculate that these are long-range extensions of the eggshell membrane fibres into the shell. We previously tried to determine it in the overall digest of the layers [12], but in this complex and rich mixture it was impossible. We can speculate that Mann et al. [5] did not find these types of collagen because they analysed the acid-soluble organic matrix, and fibrillar, crosslinked proteins (with a triple helix structure) are resistant to extraction and to trypsin digestion. The better approach for analysing these proteins is using a CNBr cleavage to disturb the protein structure (where a limited number of peptides arises after this cleavage) and then trypsin can subsequently cleave the protein's structure. Another complication in the determination of fibrillar collagens (namely types I and III) is that they contain a high proportion of hydroxylated proline (hydroxyproline) and also hydroxylysine. In the observed peptides we found only three without hydroxylation. For this reason appropriate variable modification in the data processing/searching procedure needs to be used.

An important question is the description of collagen in databases. The first point to stress is that all peptides assigned to collagens are specific to these proteins. Another issue is their assignment to specific animal species when collagen structure is relatively conservative. For this reason many peptides can be assigned to different animals (i.e. collagen from chicken, cow, human, dog, and rat but also *mammoth* or *Tyranosaurus rex* can have the same collagenous peptides). In the presented search we prefer assignment to chicken, however only one peptide specific to chicken was observed in the collagen alpha-2(I) chain. For this reason, the peptides were preferentially ranked against IPI00572548 (or SwissProt P02457) for collagen chain alpha 1(I) and IPI00813745 (or SwissProt P02467) for collagen chain alpha 2(I). However identified collagen type III does not contain any chicken-specific peptides.

Keratins are another case of sequence similarity. These proteins are present in many laboratory samples (as a common contaminant) so we concentrated on chicken keratins only (see methods for subtraction of human keratins). However we focus on the IPI-chicken database only (Table 1). The same approach was used by other researchers [5, 7, 21, 30]. When the SwissProt database search was performed a few other proteins were identified: P02769 serum albumin (*Bos taurus*) in all layers and P02091 haemoglobin subunit (*Rattus norvegicus*) in palisade layer.

Lysozyme C was identified as specific for chicken, however two peptides in the mammillary layer were found to be specific for duck (*Anas platyrhynchos*), when asparagine in position 103 was modified to aspartic acid, and glutamine at position 58 was modified to glutamic acid. These modifications can be connected to post-translational changes and sample preparation. For this reason we excluded this protein from the tables.

Individual layers differ in their protein composition, in the relative abundance of individual proteins (Table 3). Some proteins that are relatively abundant overall are missing in the cuticle layer (fibronectin 1, EDIL3, ovalbumin) or in a limited amount (collagens), but ovocalyxin-32 is preferentially present in the cuticle. On the other hand, some proteins are mainly present in the mammillary layer (ovalbumin, lysozyme C), but collagen is missing in this layer. A protein typical of the palisade layer (the thickest part of eggshell) is collagen I (both chains). Some proteins are relatively common in all three layers: ovocleidin-116, clusterin, ovocalyxin-36, Kunitz-like protease inhibitor, and ovocleidin-17. These findings (differences between individual layers) also indicate that the procedure for eggshell solubilization was appropriately selected.

It is also obvious that the relative abundance of insoluble (EDTA-insoluble) proteins differs significantly in comparison with the previously published profile of acid-soluble proteins [5, 6]. The insolubility of the proteins is probably caused by the cross-linking of these proteins [12]. Of course triple-helical collagen type I (which consists of two types of α-chains) is well known for its insolubility and resistance to cleavage. The theory of there being a specific protein network in the eggshell matrix is supported by incorporation of eggshell pigment (protoporphyrin IX) in the cuticle matrix. This pigment is not extractable without erosion of the cuticle [12].

Conclusion

EDTA-insoluble proteins in the chicken eggshell were identified. The protein profile differs significantly between the three layers cuticle, pallisade, and mammillary. The total number of proteins was 28, with seven proteins in the cuticle, 16 in the pallisade layer, and 23 in the mammillary layer. We identified for the first time collagens I (two chains) and III in the eggshell matrix, and Kunitz-like protease inhibitor as a major shell matrix protein. The dominant proteins are eggshell-specific proteins. Abundant proteins have antimicrobial functions (in addition to the eggshell-specific proteins, also the Kunitz-like protease inhibitor).

While some major insoluble proteins were identified we think that many minor proteins still could be identified. Another perspective is to determine posttranslational modifications of proteins, such as cross-links or glycosylation.

Table 3 Overview of protein distribution in various eggshell layers (proteins with at least two peptides are presented). The proteins were sorted according to their relative abundances using the exponentially modified protein abundance index (emPAI)

Cuticle layer			Palisade layer			Mammillary layer		
Database accession #	Protein name	emPAI	Database accession #	Protein Name	emPAI	Database accession #	Protein name	emPAI
IPI00572756	Ovocleidin-17	1.581	IPI00581368	LOC395256 Ovocleidin-116	1.649	IPI00581368	LOC395256 Ovocleidin-116	1.534
IPI00581368	LOC395256 Ovocleidin-116	0.943	IPI00573506	LOC419289 Ovocalyxin-36	0.911	IPI00572756	Ovocleidin-17	1.254
IPI00578622	RARRES1 Ovocalyxin-32	0.874	IPI00819356	CLU 49 kDa protein (clusterin)	0.778	IPI00600859	LYZ Lysozyme C	0.931
IPI00681344	LOC771972 similar to Kunitz-like protease inhibitor	0.840	IPI00572756	Ovocleidin-17	0.719	IPI00583974	LOC396058 Ovalbumin	0.884
IPI00573506	LOC419289 Ovocalyxin-36	0.655	IPI00681344	LOC771972 similar to Kunitz-like protease inhibitor	0.501	IPI00819356	CLU 49 kDa protein (clusterin)	0.874
IPI00819356	CLU 49 kDa protein (clusterin)	0.369	IPI00591158 IPI00588727	EDIL3 similar to EGF-like repeats and discoidin I-like domains 3	0.448	IPI00573506	LOC419289 Ovocalyxin-36	0.655
IPI00572548	LOC395532 Collagen alpha-1 (I) chain	0.042	IPI00590535	FN1 similar to fibronectin 1 isoform 1 preproprotein	0.357	IPI00591158 IPI00588727	EDIL3 similar to EGF-like repeats and discoidin I-like domains 3	0.572
			IPI00572786	NPNT similar to nephronectin short	0.194	IPI00681344	LOC771972 similar to Kunitz-like protease inhibitor	0.501
			P02465	Collagen alpha-2(I) chain OS = Bos taurus GN = COL1A2 PE=1 SV=2	0.164	IPI00685019 IPI00577021	IGLL1 24 kDa protein; IGGL1 Ig lambda chain C region	0.334
			IPI00822585	KRT4 hypothetical protein	0.151	IPI00588573	MFGE8 similar to Milk fat globule-EGF factor 8 protein	0.299
			IPI00571581	QSOX1 Sulfhydryl oxidase 1	0.150	IPI00571581	QSOX1 Sulfhydryl oxidase 1	0.277
			IPI00575989	TUBA1C Tubulin alpha-1 chain (Fragment)	0.136	IPI00574331	GKN2 hypothetical protein	0.274
			P02453	Collagen alpha-1(I) chain OS = Bos taurus GN = COL1A1 PE=1 SV=3	0.131	IPI00578016	DKK3 Dickkopf-related protein 3	0.259
			IPI00592753	KRT14 Keratin, type I cytoskeletal 14	0.105	IPI00578622	RARRES1 Ovocalyxin-32	0.233
			IPI00573727	NUCB2 Putative uncharacterized protein	0.096	IPI00590535	FN1 similar to fibronectin 1 isoform 1 preproprotein	0.219
			P04258	Collagen alpha-1(III) chain OS = Bos taurus GN = COL3A1 PE=1 SV=1	0.089	IPI00575989	TUBA1C Tubulin alpha-1 chain (Fragment)	0.212
			IPI00572548	LOC395532 Collagen alpha-1(I) chain	0.064	IPI00585021	SERPINBP1 similar to Ovalbumin-related protein Y	0.184
			IPI00585935	SEMA3G similar to FLJ00014 protein	0.063	IPI00822585	KRT4 hypothetical protein	0.151
						IPI00572786	NPNT similar to nephronectin short	0.142
						IPI00585935	SEMA3G similar to FLJ00014 protein	0.131
						IPI00585604	−189 kDa protein	0.106
						IPI00574055	PLOD1 Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1	0.099
						IPI00683271 IPI00578012	Ovotransferrin; LTF 78 kDa protein	0.099
						P04258	Collagen alpha-1(III) chain OS = Bos taurus GN = COL3A1 PE=1 SV=1	0.089

Acknowledgements Financial support from the Grant Agency of the Czech Republic (203/08/1428 and 203/09/0675) and the Centre for Cardiovascular Research 1M0510 are gratefully acknowledged.

References

1. Banks JF (1998) Protein analysis. In: Deyl Z, Mikšík I, Tagliaro F, Tesařová E (eds) Advanced chromatographic and electromigration methods in BioSciences. Elsevier, Amsterdam, pp 525–573
2. Nys Y, Hincke MT, Arias JL, Garcia-Ruiz JM, Solomon SE (1999) Avian eggshell mineralization. Poult Avian Biol Rev 10:143–166
3. Rose M, Hincke MT (2009) Protein constituents of the eggshell: eggshell-specific matrix proteins. Cell Mol Life Sci 66:2707–2719
4. Arias JL, Wiff JP, Fuenzalida V, Fernandez MS (2003) Molecular regulation of avian eggshell biomineralization. In: Kobayashi I, Ozawa H (eds) Biomineralization: formation, diversity, evolution and application. Tokai University Press, Japan, pp 13–17
5. Mann K, Macek B, Olsen JV (2006) Proteomic analysis of the acid-soluble organic matrix of the chicken calcified eggshell layer. Proteomics 6:3801–3810
6. Mann K, Olsen JV, Maček B, Gnad F, Mann M (2008) Identification of new chicken egg proteins by mass spectrometry-based proteomic analysis. World's Poult Sci J 64:209–218
7. Farinazzo A, Restuccia U, Bachi A, Guerrier L, Fortis F, Boschetti E, Fasoli E, Citterio A, Righetti PG (2009) Chicken egg yolk cytoplasmic proteome, mined via combinatorial peptide ligand libraries. J Chromatogr A 1216:1241–1252
8. Hernandez-Hernandez A, Gomez-Morales J, Rodriguez-Navarro AB, Gautron J, Nys Y, Garcia-Ruiz JM (2008) Identification of some active proteins in the process of hen eggshell formation. Cryst Growth Des 8:4330–4339
9. Mine Y, Oberle C, Kassaify Z (2003) Eggshell matrix proteins as defence mechanism of avian eggs. J Agric Food Chem 51:249–253
10. Wellman-Labadie O, Picman J, Hincke MT (2008) Antimicrobial activity of the Anseriform outer eggshell and cuticle. Comp Biochem Physiol B Biochem Mol Biol 149:640–649
11. Wellman-Labadie O, Picman J, Hincke MT (2008) Antimicrobial activity of cuticle and outer eggshell protein extracts from three species of domestic birds. Br Poultry Sci 49:133–143
12. Mikšík I, Sedláková P, Mikuliková K, Eckhardt A (2007) Proteins of insoluble matrix of avian (Gallus gallus) eggshell. Connect Tissue Res 48:1–8
13. Lakshminarayanan R, Kini MR, Valiyaveettil S (2002) Investigation of the role of ansocalcin in the biomineralization in goose eggshell matrix. PNAS 99:5155–5159
14. Lakshminarayanan R, Kini RM, Valiyaveettil S (2002) Supramolecular chemistry and self-assembly special feature: Investigation of the role of ansocalcin in the biomineralization in goose eggshell matrix. PNAS 99:5155–5159
15. Lakshminarayanan R, Valiyaveettil S, Rao VS, Kini RM (2003) Purification, characterization, and in vitro mineralization studies of a novel goose eggshell matrix protein, ansocalcin. J Biol Chem 278:2928–2936
16. Mann K (2004) Identification of the major proteins of the organic matrix of emu (Dromaius novaehollandiae) and rhea (Rhea americana) eggshell calcified layer. Br Poultry Sci 45:483–490
17. Kersey PJ, Duarte J, Williams A, Karavidopoulou Y, Birney E, Apweiler R (2004) The International Protein Index: an integrated database for proteomics experiments. Proteomics 4:1985–1988
18. Arias JL, Fernandez MS (2001) Role of extracellular matrix molecules in shell formation and structure. World's Poult Sci J 57:349–357
19. Pataridis S, Eckhardt A, Mikulíková K, Sedláková P, Miksik I (2009) Determination and quantification of collagen types in tissues using HPLC–MS/MS. Curr Anal Chem 5:316–323
20. Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, Mann M (2005) Exponentially Modified Protein Abundance Index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol Cell Proteomics 4:1265–1272
21. Mann K (2007) The chicken egg white proteome. Proteomics 7:3558–3568
22. Mikšík I, Sedláková P, Mikulíková K, Eckhardt A, Kašička V (2008) Comparison of CE–MS and LC–MS analyses of avian eggshell matrix proteins. Chromatographia 67:S89–S96
23. Gautron J, Nys Y (2007) Eggshell matrix proteins. In: Huopalahti R, López-Fandiño R, Anton R, Schade R (eds) Bioactive egg compounds. Springer, Berlin Heidelberg New York, pp 103–108
24. Myllyharju J, Kivirikko KI (2004) Collagens, modifying enzymes and their mutations in humans, flies and worms. Trends Genet 20:33–43
25. Viguet-Carrin S, Garnero P, Delmas P (2006) The role of collagen in bone strength. Osteoporos Internat 17:319–336
26. Fernandez MS, Moya A, Lopez L, Arias JL (2001) Secretion pattern, ultrastructural localization and function of extracellular matrix molecules involved in eggshell formation. Matrix Biol 19:793–803
27. Arias JL, Fernández MS, Dennis JE, Caplan AI (1991) Collagens of the chicken eggshell membranes. Connect Tissue Res 26:37–45
28. Arias JL, Nakamura O, Fernández MS, Wu JJ, Knigge P, Eyre DR, Caplan AI (1997) Role of Type X collagen on experimental mineralization of eggshell membranes. Connect Tissue Res 36:21–33
29. Wong M, Hendrix MJC, von der Mark K, Little C, Stern R (1984) Collagen in the egg shell membranes of the hen. Develop Biol 104:28–36
30. Mann K, Mann M (2008) The chicken egg yolk plasma and granule proteomes. Proteomics 8:178–191