Guinea fowl eggshell quantitative proteomics yield new findings related to its unique structural characteristics and superior mechanical properties

Nathalie Le Roy, Lucie Combes-Soia, Aurélien Brionne, Valérie Labas, Alejandro B. Rodriguez-Navarro, Maxwell T. Hincke, Yves Nys, Joël Gautron

Guinea fowl eggshell quantitative proteomics yield new findings related to its unique structural characteristics and superior mechanical properties.

Nathalie Le Roy[1], Lucie Combes-Soia[2], Aurélien Brionne[1], Valérie Labas[2,3], Alejandro B. Rodriguez-Navarro[4], Maxwell T. Hincke[5], Yves Nys[1] and Joël Gautron[1*]

[1] BOA, INRA, Université de Tours, 37380 Nouzilly, France

[2] UMR PRC, INRA 85, CNRS 7247, Université de Tours, IFCE, 37380 Nouzilly, France

[3] CIRE, Pôle d'Analyse et d'Imagerie des Biomolécules, INRA, CHRU de Tours, Université de Tours, 37380 Nouzilly, France.

[4] Departmento de Mineralogia y Petrologia, Universidad de Granada, 18071 Granada, Spain

[5] Departments of Innovation in Medical Education, and Cellular and Molecular Medicine, University of Ottawa, Ottawa, K1H 8M5, Canada

* corresponding author: joel.gautron@inra.fr; Tel: +33247427540

Abstract

The Guinea fowl eggshell is a bioceramic material with the remarkable mechanical property of being twice as strong as the chicken eggshell. Both eggshells are composed of 95% calcite and 3.5% organic matrix, which control its structural organization. Chicken eggshell is made of columnar calcite crystals arranged vertically. In the Guinea fowl, the same structure is observed in its inner half, followed by a dramatic change in crystal size and orientation in the outer region. Guinea fowl eggshell is thicker than chicken eggshell. Both structure and shell thickness confer a superior resistance to breakage compared to eggshells of other bird species. To understand the underlying mechanisms controlling the structural organization of this highly resistant material, we used quantitative proteomics to analyze the protein composition of the Guinea fowl eggshell organic matrix at key stages of the biomineralization process. We identified 149 proteins, which were compared to other bird eggshell proteomes and analyzed their potential functions. Among the 149 proteins, 9 are unique to Guinea fowl, some are involved in the control of the calcite precipitation
(Lysozyme, Ovocleidin-17-like, Ovocleidin-116 and Ovalbumin), 61 are only found in the zone of microstructure shift and 17 are more abundant in this zone.

Significance

The avian eggshell is a critical physical barrier to protect the contents of this autonomous reproductive enclosure from physical and microbial assault. The Guinea fowl (*Numida meleagris*) eggshell exhibits a unique microstructure (texture), which confers exceptional mechanical properties compared to eggshells of other species. In order to understand the mechanisms that regulate formation of this texture in the Guinea fowl eggshell, we performed comparative quantitative proteomics at key stages of shell mineralization and particularly during the dramatic shift in shell microstructure. We demonstrate that the Guinea fowl eggshell proteome comprises 149 proteins, of which 61 were specifically associated with the change in size and orientation of calcite crystals. Comparative proteomics analysis with eggshell of other bird species leads to new insights into the biomineralization process. Moreover, our data represents a list of organic compounds as potential additives to regulate material design for industrial fabrication of ceramics. This information also provides molecular markers for efficient genomic selection of chicken strains to lay eggs with improved shell mechanical properties for enhanced food safety.

Introduction

The avian egg is a closed chamber that allows autonomous development of the embryo in an external environment. During this process, the eggshell physically protects the embryo against pathogen penetration, allows gaseous exchange and serves as a calcium source for the developing embryo [1,2]. The eggshell is a

bioceramic structure made of 95% calcium carbonate ($CaCO_3$) in the calcite polymorph, contains occluded organic components (3.5%) and 1.5% water [1–3]. Shell biomineralization is initiated in the distal oviduct (isthmus), where eggshell membranes are deposited to support the calcified structure that is subsequently formed in the uterine segment. Shell formation is temporally controlled, and in chickens four main steps can be identified during the 17 h process (from 5 h to 22 h post-ovulation) [4–6]. The first one is the nucleation and stabilization of amorphous calcium carbonate (ACC) on the outer eggshell membrane fibers and at specific nucleation sites (mammillary knobs). The ACC is then transformed into calcitic aggregates at these sites (2^{nd} step). Calcite crystals rapidly grow to form larger crystal units (3^{rd} step). The interaction with eggshell organic matrix components inhibits calcite crystal faces parallel to the c-axis, thus causing elongated crystal growth in this direction. Calcite crystals growing with their c-axis nearly perpendicular to the surface block the growth of adjacent crystals with less favorable orientations resulting in the development of columnar calcite units. Finally, mineralization is terminated and a thin proteinaceous layer (cuticle) is deposited on the shell surface (4^{th} step). Numerous *in vitro* experimental evidence has demonstrated that eggshell fabric is controlled by organic matrix components. This organic matrix determines which calcium carbonate polymorph phase (i.e., calcite) is selected/precipitated, in addition to facilitating their nucleation and influencing the size and morphology of crystals [4]. The role of organic matrix is also shown *in vivo*

when exploring genetic and physiological factors influencing changes in chicken eggshell microstructure and/or its mechanical properties [1,7,8]. Across a large number of avian species, a strong correlation is observed between egg mass or eggshell thickness and shell strength [9,10], indicating that the solidity of shells is similar in a majority of birds, once corrected for eggshell thickness. However, this correlation does not apply to Guinea fowl (*Numida meleagris*) eggshell, which has a greater breaking strength than expected compared to that of other birds [11]. Part of the greater strength in Guinea fowl shell results from higher shell thickness (eggshell weight represents 15% and 10% of the total egg weight in Guinea fowl and in chicken, respectively)
[12,13]. The length of the shell deposition process is 2.5 h longer in Guinea fowl than in chicken, resulting in a thicker eggshell (500 µm in Guinea fowl *vs* 350 µm in chicken) [13,14]. However, the amount of material deposited in the shell cannot solely explain the superior mechanical properties of the Guinea fowl eggshell [13]. Indeed, when Guinea fowl and chicken eggshell of similar thickness are compared, that of Guinea fowl is two times stronger than its chicken correspondant [15]. Guinea fowl eggshell possesses a particular structure that confers greater resistance to fracture [15–17]. In other biomineralization models such as molluskan shells, the different types of textures (nacre, prismatic and cross-lamellar layers [18,19]) are known to exhibit a difference in their strength in addition to the effect of shell thickness [20-22]. Guinea fowl eggshell mineralization initially follows the same pattern

described in chickens resulting in the characteristic columnar structure of most birds. However, a sharp change in the size and orientation of crystals occurs at 11 hours p.o. when one-third of the final eggshell thickness has been deposited [13,17]. At this point, the large columnar calcite units break up into smaller crystal units with varying crystallographic orientations forming a microstructure with an intricate interlacing of calcite crystals. These changes in the structure and organization of Guinea fowl eggshell from the macro to micro are described in detail in an accompanying paper (Rodriguez-Navarro AB et al., unpublished). That study underlines the bilayer microstructure of Guinea fowl eggshell, and that the shift in eggshell fabric (texture) between these layers is accompanied by changes in the distribution and amount of intra-crystalline organic matter. Thus, the organic matrix is predicted to firstly induce the initial structural shift between these layers, followed by a secondary nucleation event involving smaller crystals with increasing misorientation. This particular organization is responsible for the exceptional mechanical properties of Guinea fowl eggshell by comparison to that of other birds.

In such a context, our objectives were to characterize the Guinea fowl eggshell proteome, and then to use quantitative proteomics to analyze the change in protein profile throughout the steps of eggshell deposition prior to, during and after the main shifts in the size and orientation of calcite crystals. This study identified for the first time 149 Guinea fowl matrix proteins that we compared to proteins previously identified in eggshells of four other bird species and one

non-avian reptile. We have quantified these proteins and herein describe a protein toolkit associated with the modification in calcite crystal size and orientation that confers remarkable strength on the Guinea fowl eggshell.

Material and methods

Ethical statement, animals handling and housing

All experiments, including all animal-handling protocols were handled and carried out in accordance with the European Communities Council Directives concerning the practice for the care and use of animals for Scientific Purposes and the French Ministry on Animal experimentation under the supervision of authorized scientists (authorization #1940, delivered on 2016, Feb, 29th by the French ministry of Education and Research). Birds were kept in the experimental unit PEAT 1295 of INRA, which has permission to rear birds and to perform euthanasia of experimental animals (decree N° B37-175-1 of August 28th 2012 delivered by the "Préfécture d'Indre et Loire" following inspection by the Direction of Veterinary Services). Our experimental protocol was approved by an ethical committee (Comité d'éthique de Val de Loire, officially registered under number 19 of the French National Ethics Committee for Animal Experimentation) under agreement number 00159.02.

Collection of egg samples

Twenty-nine Guinea fowl hens (40-47 weeks of age) were individually caged and subjected to a cycle of 14 h of light and 10 h of darkness. Each bird was fed with a layer mash *ad libitum* as recommended by the Institut National de la Recherche Agronomique (INRA). Each cage was equipped with a device for automatic recording of oviposition (time of egg laying). Eggs were sampled by euthanasia of animals at one of five stages: 4 h (eggshell=5), 10 h (eggshell=6), 11 h (eggshell=6), 12 h (eggshell=6) and 18 h (eggshell=6) post-ovulation (p.o., *i.e.* time after the previous egg) when the nucleation sites appear and early mineralization starts (4 h p.o.), or just before (10 h p.o.), during (11-12 h p.o.) and after the shift in crystal orientation (18 h p.o.). Eggs were removed from the oviduct, broken and the forming eggshells were washed with water to remove egg white residues. Eggshells were air dried and stored at -20°C until the extraction of proteins.

Extraction of organic matrix proteins from eggshells

The organic matrix was extracted from each eggshell according to the protocol previously described [6,23]. Briefly, eggshells were washed in 154 mM NaCl buffer containing protease inhibitors (2.5 mM benzamidine-HCl, 50 mM amino-n-caproic acid, 0.5 mM N-ethylmaleimide and 1 mM phenylmethylsulfonyl fluoride) and then ground under liquid nitrogen. Eggshell powder was completely demineralized by shaking with 20% acetic acid overnight at 4°C. The resulting turbid suspensions were placed in dialysis tubing (cut-off 3,500

Da), dialyzed against demineralized water for 24 hours at 4°C and then lyophilized. Eggshell extracts were shaken overnight at 4°C using the following solution: 4 M guanidine hydrochloride, 5 mM benzamidine hydrochloride, 0.1 M ε-amino-n-caproic acid, 10 mM EDTA, 50 mM sodium acetate and 1 mM phenylmethylsulfonyl fluoride. Samples were then dialyzed (cut-off 3,500 Da) against 0.5 M sodium acetate pH 7.4 for 24 hours at 4°C. Samples were centrifuged for 10 min, 2000xg at 4°C and supernatants were retained. Proteins were quantified using the BioRad DC Protein Assay kit II, with bovine serum albumin as standard.

Electrophoresis and Western blotting

Proteins were prepared in the Laemmli/β-mercapto-ethanol buffer and boiled 5 min at 90°C for electrophoresis for mass spectrometry analyses and Western blotting. We used biological replicates for each stage: 4 h (eggshell=5), 10 h (eggshell=6), 11 h (eggshell=6), 12 h (eggshell=6) and 18 h (eggshell=6) post-ovulation, totaling 29 eggshells as explained in Figure 1. To test the homogeneity of eggshell samples at each stage (step #1 in Figure 1), each biological replicate per stage (15 □ g) was loaded on a 12.5% gel for rapid electrophoresis (10 min) in order to obtain a single band. After staining with Coomassie Blue, the single band per lane was excised. The 29 excised single bands were in-gel digested using trypsin (Sequencing grade, Roche diagnostics, Paris, France), and separately subjected to nanoLC-MS/MS analysis in triplicate

(technical replicates). The analysis of a single band does not allow identifying a maximum of proteins at each stage, so we pooled 16 µg of extracted proteins from each of 5 biological replicates at 4 h, and 13.33 µg of extracted proteins from each 6 of the biological replicates at 10 h, 11 h, 12 h and 18 h p.o. (step #2, Figure 1) to obtain 80 µg of equaly pooled biological replicates for each of the 5 post-ovulation stages. For each stage, the 80 µg was loaded on a 4-20% gel for SDS-PAGE and was entirely migrated (100 min), followed by staining with Coomassie Blue. Each of the 5 lanes was sectioned into 20 gel slices (step #2, Figure 1), totaling 100 bands. These bands were in-gel digested with bovine trypsin (Sequencing grade, Roche diagnostics, Paris, France) and then subjected to nanoLC-MS/MS analysis. For Western blot analyses, 10 µg of protein from each pooled biological replicate was electrophoresed on 12.5% gels and then transferred to nitrocellulose membranes (0.2 µm, GE Healthcare, Little Chalfont, England) for 1 h in 1X TBS (50 mM Tris, 150 mM NaCl, pH 7.4), 10% ethanol. Membranes were blocked 1 h at room temperature in Odyssey® Blocking Buffer (TBS) (LI-COR, Bad Homburg, Germany) and then incubated in blocking buffer for 1h at room temperature with anti-OVAL (1/5000; Sigma Aldrich, C6534) or anti-Ovocleidin-17 (1/5000;[24]) polyclonal antibodies. Membranes were washed three times 5 min in 1X TBS, 0.1% Tween 20 and then incubated for 1h in the dark in Odyssey® Blocking Buffer (TBS) with AlexaFluor® 680 goat anti-rabbit antibody (1/20000; Invitrogen, A21109). After three washes of 5 min in 1X TBS, 0.1% Tween 20 and one wash in 1X

TBS, membranes were scanned using the Odyssey® CLx Imaging System (LI-COR, Bad Homburg, Germany) with the 700 nm channel.

Nano-LC-MS/MS analyses

Peptides resulting from trypsin digestion of the 29 bands from individual biological replicates per stage and 100 bands from pooled biological replicates per stage were analyzed using the LTQ Orbitrap Velos Mass Spectrometer (Thermo Fisher Scientific, Bremen, Germany) coupled to an Ultimate® 3000 RSLC Ultra High Pressure Liquid Chromatographer (Dionex, Amsterdam, The Netherlands). Five microliters of each sample were loaded in three technical replicates onto a trap column for desalting. Subsequently, peptide separation was performed using a nano-column (Acclaim PepMap C18, 75 µm inner diameter, 50 cm long, 3 µm particle size, 100 Å pores). The mobile phases consisted of (A) 0.1% formic acid, 97.9% water, 2% acetonitrile (v/v/v) and (B) 0.1% formic acid, 15.9% water, 84% acetonitrile (v/v/v). The gradient consisted of 4–55% B for 90 min at 300 nl/min flow rate. Data were acquired using Chromeleon Software (v6.8 SR11, Dionex, Amsterdam, The Netherlands) and Xcalibur software (v2.1; Thermo Fisher Scientific, San Jose, CA). The LTQ Orbitrap Velos instrument was operated in a positive data-dependent mode. Resolution of the Orbitrap was set to R = 60,000. In the scan range of m/z 300–1800, the 20 most intense peptide ions with charge states ≥2 were fragmented

using Collision Ion Dissociation (CID). Dynamic exclusion was activated for 30 s with a repeat count of 1. Polydimethylcyclosiloxane ions (m/z, 445.1200025, $(Si(CH_3)_2O)_6$) were used as a lock mass for internal calibration.

The raw mass spectrometric data have been deposited to PRIDE *via* the ProteomeXchange Consortium (https://www.ebi.ac.uk/pride/archive/) with the project accession PXD011333 and project DOI 10.6019/PXD011333.

Protein identification and validation

Mass spectrum ion searches were performed using MASCOT search engine (v2.3, Matrix Science, London, UK). Interrogation was made against the Chordata section of the non-redundant NCBI database (January 2018). Enzyme specificity was set to trypsin with 2 missed cleavages, using carbamidomethylcysteine, oxidation of methionine and N-terminal protein acetylation as variable modifications. The tolerance of the ions was set to 5 ppm for parent and 0.8 Da for fragment ion matches. Mascot results obtained from the target database searches were subjected to Scaffold v4.8.2 software (Proteome Software Inc., Portland, USA). Peptide Decoy was selected to detect false-positive results. Peptide identifications were accepted if they could be established at greater than 95% probability by the Peptide Prophet algorithm [25]. Protein identifications were accepted if they could be established at greater than 95% probability and contained at least two unique peptides. Protein

probabilities were assigned by the Protein Prophet algorithm [26]. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide identity were grouped into clusters. Keratin, bovine albumin and trypsin protein identifications were deleted from our dataset as they result from external contaminants (human keratin) and the digestion procedure (bovine trypsin).

Statistical analyses

Six eggshells were used per biomineralization stages (five stages).

Electrophoretic profile of eggshell organic matrix were compared within each stage and one individual from stage 4 h p.o. was removed because of its too different electrophoretic profile. The six or five individual samples were pooled for each stage. Individual and pooled samples were analyzed by Mass spectrometry in triplicate.

The quantity of extracted organic matrix from each eggshell and the eggshell weights at each stage were statistically analyzed using the Kruskal-Wallis test and pairwise comparison, with the significance threshold at $p<0.05$, as they are twenty-nine independent samples.

After nanoLC-MS/MS analysis of all samples (individual biological replicates and pooled biological replicates), relative quantification of identified proteins

inside each sample was performed using Exponentially Modified Protein Abundance Index (emPAI) and Normalized Spectral Abundance Factor (NSAF). The relative quantification of protein abundance between the five conditions, in the absence of labeling ("label-free"), was carried out according to the Weighted Spectral count (WS) and the eXtracted Ion Chromatogram (XIC) methods. The WS method enumerates the number of weighted spectra assigned for each protein as defined by Scaffold and the XIC is a quantification of proteins according to the area under the spectrometer peaks. The analysis of variance for each identified protein at the five stages of eggshell biomineralization was performed with WS (threshold ≥5 spectra) and XIC values using the appropriate one-way ANOVA test, followed by a pairwise Tukey postHoc test to determine significant differences in protein abundance between the five stages ($p<0.05$). Accession numbers of identified eggshell proteins from four birds and one crocodile proteomes [6,27–31] were listed and protein amino acid sequences were recovered from NCBI (https://www.ncbi.nlm.nih.gov), UniProt (https://www.uniprot.org) and NODAI (http://www.nodai-genome.org) databases. Each proteome dataset was locally blasted against the *Numida meleagris* proteome using BLAST+ suite 2.6.0+ [32] on Genotoul server (http://bioinfo.genotoul.fr).

Results

Guinea fowl eggshell proteome

Forming eggshell samples were collected at five calcification stages, around the time of shift in the eggshell ultrastructure and crystallographic texture of Guinea fowl eggshells. These were the initial stage corresponding to the first events of mineral deposition at 4 hours post-ovulation (4 h p.o.); the growth of large columnar calcite units just prior to the shift of crystal size (10 h p.o.); during the period when the microcrystalline layer begins to be deposited at 11 h p.o., and at 12 h p.o.. And finally, the last stage (18 h p.o.) was collected at a time when the growth of the microcrystalline layer is ongoing and secondary nucleation events continue to occur.

For this study, we used biological replicates: 5 eggshells at 4 h and 6 eggshells for each 10 h, 11, h, 12 h and 18 h p.o. stages. Data from mass-spectrometry analysis of the single band per biological replicate (Figure 1) were found to be homogenous and consequently validated our methodology for the quantification of proteins at different stages using pooled biological replicates (Supplementary data 1, sheet #1 and sheet #2). Indeed, in each stage, biological replicates were similar as major proteins were identified in the 5 or 6 biological replicates (Supplementary data 1, sheet #1).

The profiles of the matrix proteins extracted at these 5 key steps of Guinea fowl eggshell formation and pooled in equally quantity per stage are shown on a 4-20% SDS-PAGE gel (Figure 1). Proteins smaller than 25 kDa were less abundant at 4 h p.o. than at the other stages. The main changes between protein profiles of shell collected from 10 to 18 hours concerned the intensity of the

electrophoretic bands. To further analyze these modifications, the entire lane at each calcification stage was sectioned into 20 single bands. To determine the variability of sample populations at each stage, individual shell samples (5 or 6 per stage) were briefly run on polyacrylamide gels (without fractionation) allowing single bands containing all proteins to be excised. All gel slices were digested using trypsin and the resulting peptides were analyzed by nanoLC-MS/MS, in triplicate. Mass spectra from peptides were assigned using the NCBI Chordata database.

NanoLC-MS/MS analysis of pooled eggshells at the five post-ovulation stages allowed us to identify 162 clusters of proteins, which were assigned using the non-redundant NCBI database for chordate taxa. After manual removal of redundancy, enzymatic trypsin used for digestion and contaminating proteins (bovine serum albumin, keratins and casein; Supplementary data 2), 149 non-redundant proteins were identified, which represents the first description of *Numida meleagris* eggshell proteome (Supplementary data 3, sheet #1).

We used the Exponentially Modified Protein Abundance Index (emPAI), which is based on unique peptide count for each protein, to classify the relative abundance of the different proteins within individual stages (4, 10, 11, 12 and 18 h p.o.; Supplementary data 3, sheet #2 to 5). Table 1 reports the 15 most abundant proteins classified by magnitude of abundance within each stage. Dromaiocalcin-1-like (DCA-1-like), Ovocleidin-17-like (OC-17-like), Extracellular Fatty Acid-Binding protein (ExFAB), Ovomucoid (OVM),

Ovotransferrin (OVT), Ovalbumin isoform X2 (OVA X2), Lysozyme C (LYZ C), Cystatin C (CST3), Alpha-1-acid glycoprotein 2-like (AGP-like) and Clusterin (CLU) were 10 proteins systematically present in the 15 most abundant proteins, whatever the stage. Hemopexin (HPX) was in the top 15 proteins except for the last stage (18 h p.o.), when the growth of the new crystalline microstructure layer is stabilized a few hours after the mineral structural shift. Ovalbumin related protein Y (OVALY), Hyaluronan and proteoglycan link protein 3 (HAPLN3) and Hemoglobin subunit Alpha-A (HBAA) were in the 15 most abundant proteins only at 4 h p.o., when the first events of mineral deposition occurred. Conversely, Ovocleidin-116-like (OC-116-like/MEPE-like) was in the top 15 of the most abundant proteins for all stages except at 4 h p.o., the onset of shell calcification. Small Basic Protein-1-like (SBP-1-like) and Beta-2-Microglobulin (B2M) were in top 15 most abundant proteins only at 18 h p.o.. Avidin-like (AVD-like) was one of the 15 most abundant proteins at 4, 10 and 11 h p.o.. Ovocalyxin-32-like (OCX-32-like/RARRES1-like) and EGF-like repeat and discoidin I-like domain-containing protein 3 (EDIL3) appeared to be present in the top 15 proteins at the stages bracketing the crystal shift (10-12 h p.o.). Finally, Albumin (ALB) was in list of 15 most abundant proteins just after the main structural shift (12 h and 18 h p.o.).

Comparison with the eggshell proteome of other archosaurian species

To date, four bird eggshell proteomes have been studied quite extensively; 699 proteins were identified in chicken (*Gallus gallus*) [6,27], 697 proteins in turkey (*Meleagris gallopavo)* [28], 622 proteins in quail (*Coturnix japonica*) [29] and 475 proteins in zebra finch (*Taeniopygia guttata*) [30]. Recently, the first proteome of the crocodile eggshell, from *Crocodylus siamensis*, was published with a total of 58 proteins [31]. These five proteomes were compared with the Guinea fowl eggshell proteome (present study) using the blastp method. The number of identified proteins varied largely between the studies, as those concerning crocodile and Guinea fowl eggshells are less extensive, possibly due to incomplete annotation of these diverse genomes. Results of this analysis showed that *N. meleagris* current eggshell proteome shared 86.6%, 62.4%, 75.8% and 61.7% of its proteins with the chicken, quail, turkey, and zebra finch, respectively whereas 14.1% (21 proteins) were shared with the proteome of the crocodile eggshell (Figure 2A; Supplementary data 4). Among the 149 proteins identified in the Guinea fowl eggshell proteome, 9 of these proteins were not detected in the five other archosaurian species. The majority of proteins (119) were shared between Guinea fowl and at least one bird species, and among them 54 were shared by all 5 bird species (Supplementary data 4). However, only 21 proteins overlapped with those of the crocodile species (Figure 2B).

The 9 proteins unique to Guinea fowl eggshell (Table 2) were Dromaiocalcin-1-like (DCA-1-like), Phosphatidylethanolamine-binding protein 4 (PEBP4), Small basic protein1-like (SBP-1-like), Uncharacterized protein At5g39570-like

(LOC110408336), ribonuclease UK114 (RIDA), Cell growth regulator with EF-hand domain protein 1 (CGREF1), granulins (GRN), Growth/differentiation factor 6 (GDF6) and Sushi repeat-containing protein SRPX (SRPX) (Table 2, Figure 2B). Among these 9 proteins, Dromaiocalcin-1-like possessed the highest emPAI value (from 42.3 at 4 h p.o. to 809.1 at 18 h p.o.). The 8 other proteins exhibited lower emPAI.

Proteins identified and variation of their abundances at various stages during the Guinea fowl eggshell mineralization

In a first approach, we used the list of identified proteins based on unique peptide count (emPAI) to determine the specificity of proteins during the shell deposition. We first observed an increase in the diversity of matrix proteins in the shell explaining the increasing number of identified proteins from 4 h p.o. (51 proteins), to the shift period (119, 113 and 114 at 10, 11 and 12 h p.o., respectively) (Figure 3A). This observation is in agreement with the fact that the expulsed eggshells contain the constituents of the earlier stages when the egg is expulsed at a later time point. We observed a decreased number of identified proteins at the later stage (62 at 18 h p.o.), like in the analysis of biological replicates for which up to 16 proteins at 4 h, 32 proteins at 10 h, 28 proteins at 11 h, 29 proteins at 12 h and 18 proteins at 18 h were identified (Supplementary data 1, sheet #1). This decrease was probably due to deposition of highly abundant proteins at this stage that hinder the detection of fainter proteins. For

instance, at 18 h p.o. the average of emPAI values of biological replicates for DCA-1-like protein is 147.6, while it is 115.8, 124.1 and 113.8 at 10 h, 11 h and 12 h p.o., respectively (Supplementary data 1, sheet #2). Venn diagrams (Figure 3B) showed the distribution of identified proteins according to the stage of shell formation. We found 25 proteins common to all five stages of eggshell formation. Seven proteins were specific to the onset of calcification (4 h p.o.) (Olfactomedin-4 (OLFAL4), Sulfhydryl oxidase 1 (QSOX1), Lipocalin-like (LCNL1-like), Catalase (CAT), Hypothetical protein ASZ78_007562 (LOC110389945), Vitelline membrane outer layer protein 1 (VMO1) and Annexin A8 (ANXA8)). Sixty one proteins were only present at 10-12 h p.o. stages (Supplementary data 5, sheet #1) corresponding to the change in the size and orientation of crystals and only four (Soluble scavenger receptor cysteine-rich domain-containing protein SSC5D-like, Glypican-1 isoform X1, Lymphocyte antigen 86 and Reticulocalbin-2 isoform 1 precursor) appeared to be associated with the steady accumulation of the new crystalline layer (18 h p.o.).

In a second approach, we extended the qualitative approach described above using quantitative methods to determine the abundance of proteins at each stage of shell mineralization. The amount of eggshell matrix was highest at 18 hours and significantly different than at other stages (Figure 4). Two quantitative proteomic methods, the Weighted Spectral count (WS) and the eXtracted Ion Chromatogram (XIC), allowed us to determine the relative abundance between

stages of the 149 proteins constituting the Guinea fowl eggshell proteome (105 proteins according to WS threshold and 149 proteins using XIC methods). The number of proteins with significant differences in abundance according to the calcification stage was 89 and 86 with WS and XIC methods, respectively, and 64 of them were differentially abundant using both methods (Supplementary data 6). The mass spectrometry quantification was validated by Western blot analyses using antibodies to chicken Ovocleidin-17 and Ovalbumin that cross-reacted with their Guinea fowl orthologs (Supplementary data 7).

Qualitative (unique petide count) and quantitative (Weighted Spectra) approaches were finally used to identify proteins associated with the main structural shift between microstructurally distinct eggshell layers in which the size of crystals is greatly reduced (shift period). We identified 61 proteins unique to the shift region of the eggshell and 17 proteins present at various stages but significantly overabundant at this site (Supplementary data 5, sheet #1 and sheet #2). P-values were added in both tables for the five stages of shell calcification. Proteins were classified in three groups according to their functions identified in the gene ontology section of Uniprot (www.uniprot.org). Two groups corresponded to proteins with functions ascribed to shell biomineralization (direct or indirect roles), the third group corresponding to other proteins with functions not related to biomineralization. This analysis highlighted 78 proteins associated to the shift period. Clusterin (CLU) and Ovocalyxin-32-like (OCX-32-like/RARRES1-like) are the most abundant

during the main structural shift and were previously described in chicken eggshell. In contrast, Growth/differentiation factor 6 (GDF6) and sushi repeat-containing protein (SRPX) are specific to Guinea fowl eggshell and unique to the structural shift.

Discussion

The Guinea fowl eggshell proteome in evolutionary context

The aim of this study was to characterize and quantify Guinea fowl eggshell matrix proteins in order to decipher how they are related to the particular ultrastructure of this shell. The present study is the first proteomic investigation of Guinea fowl eggshell during shell deposition and allowed us to compare the proteins involved in its biomineralization with eggshell of other archosaurian species (*i.e.* four birds and one crocodile).

The nanoLC-MS/MS approach allowed us to identify 149 non-redundant proteins. In other eggshell proteomes, 699, 622, 697 and 475 proteins were identified in chicken, quail, turkey and zebra finch, respectively [6,27-30]. The Guinea fowl genomic sequence and transcriptomics databases are quite recent, and its gene annotation is based on the chickem genome. Moreover, this preliminary draft has not yet been thoroughly reviewed for accuracy (Cedric Cabault and Alain Vignal, personal communication). This limits the power of bioinformatics to accurately assign all the MS spectra to predicted protein fragments. In such a context, there was reduced ability to identify all protein

constituents of the Guinea fowl eggshell proteome. Regarding other known archosaurian eggshell proteomes, a total of 14 common proteins were identified in the four birds (*Gallus gallus*, *Coturnix japonica*, *Meleagris gallopavo* and *Taeniopygia guttata*), the crocodilian species (*Corcodylus sinensis*) and the Guinea fowl eggshell proteome (Figure 2B and 2C). This suggests a common set of at least 14 proteins recruited to the process of eggshell formation by the archosaurian ancestor, about 219 Myrs ago [33]. These proteins are OVOT, HAPLN3, OVST, A2ML1, LOXL2, HIS H2B, ACTB, HBAA, MUC5AC, SDF4, TUBB, Ezrin, SERPINE2 and VTG1. They can be classified into three main groups: protease inhibitors (A2ML1, OVST and SERPINE2), proteins already known to be involved in shell mineralization (OVOT, HAPLN3, LOXL2, MUC5AC and SDF4) and proteins with intracellular functions (HIS H2B, ACTB, HBAA and TUBB). OVOT, HAPLN3 and LOXL2 have been identified as overabundant proteins during the first events of eggshell biocalcification in chicken [6]. MUC5AC was previously identified in the organic matrix proteome of invertebrate metazoan $CaCO_3$ skeletons and in gallstone formation in humans [34,35]. These observations demonstrate that essential functions that these proteins play during eggshell biomineralization were already present in the archosaurian ancestor. As the protocols for organic matrix extraction and technologies for proteomics used for these 6 proteomes were different, and database annotation is still incomplete for various species, we assume that the eggshell proteomes are still incomplete and that the set of

common proteins in archosaurs is likely larger. However, we are likely to have identified the most abundant members of this biomineralization toolkit.

The avian group (Aves; Figure 2C) emerged more than 91 Myrs ago and is divided into Palaeognathae (ostrich, emu, rhea…) and Neognathae (chicken, Guinea fowl, zebra finch…), which appeared about 85 Myrs ago [36]. Among the 119 proteins shared by Guinea fowl and at least one other bird, 54 proteins are common to all Neognathae species considered in this study, which have been consistently recruited in this taxa for the shell formation process [28-30]; in contrast, 12 proteins were not found in the Neoaves eggshell proteome (zebra finch). Among the 119 proteins, we noticed the presence of four proteins related to the biomineralization process (Nucleobindin-2 (NUCB2), extracellular serine/threonine protein kinase (FAM20C), Milk Fat Globule-EGF factor 8 protein (MFGE8) and Calbindin D-28K (CALB1)). NUCB2 is a calcium-binding protein (two EF-hand domains) highly suspected to be involved in the chicken eggshell biomineralization process [6,37,38]. Its role as pro-osteogenic activator in bones of mammals [39] reinforces this hypothesis. FAM20C also named DMP4 is involved in the regulation of biomineralization as previously shown in the teeth and bone formation [40,41] and its transcript was previously shown as overexpressed in the chicken uterus [37]. Recently, the role of this protein kinase was demonstrated in the biomineralization process of the pearl oyster *Pinctada fucata* [42], suggesting an ancient origin of its recruitment to participate in biomineralization. MFGE8 possesses calcium-binding domains

and was identified as a strong candidate involved in chicken eggshell mineralization [6,37]. CALB1, a calcium-binding protein, is present in the chicken uterine cells as the main cytosolic calcium transporter during eggshell deposition, but this protein is known to be intracellular. Consequently, its presence in the shell could be the result of passive accumulation and not an active process related to shell deposition [43–46]. The 50 other proteins are involved in catalytic and binding functions.

Finally, this comparative study singles out 9 proteins, which are unique to the Guinea fowl eggshell when compared to other Neognathae (Table 2; Figure 2C). Two of them possess putative calcium-binding domains: uncharacterized protein At5g39570-like (LOC110408336) with 12.7% aspartic acid residues, and Cell Growth Regulator with EF-hand domain protein 1 (CGREF1). Highly acidic-proteins are important proteins in molluskan shell matrix, and strongly interact with calcium due to their high negative charge [47]. LOC110408336 is an acidic protein that appears to be unique to Guinea fowl eggshell. The role of this particular protein should be investigated in relation with the particular shell structure in order to determine its potential role in the crystal shift, particularly since its stretches of poly-aspartic acid resemble the mineral-interacting motifs of osteopontin (SPP1). The C-type lectin (CTL) Dromaiocalcin-1-like (DCA-1-like), a homolog of Ovocleidin-17 (OC-17), belongs to this group of Guinea fowl unique proteins. The presence of pairs of CTL paralogs as major constituents of the eggshell proteome has been observed in the Pekin duck [48],

and in ostrich, emu and rhea [49,50]. However, proteins only identified in the Guinea fowl eggshell are important candidates in relation to the remarkable features of its shell, and could constitute a specific toolkit to build the unique ultrastructure of this bioceramic in a highly controlled way.

The first events and massive deposition of Guinea fowl eggshell

Proteins and proteoglycans that constitute the shell organic matrix are believed to play a key role in shell formation. This controlled biomineralization process occurs in a confined space (lumen of the uterus) where ionic concentrations (calcium and bicarbonates) are highly supersaturated relative to calcite [51]. Previous work explored early shell mineralization mechanisms in chicken eggshell and highlighted the importance of the formation of a transient amorphous mineral phase (ACC) at the initial stage of eggshell mineralization [5]. The recent investigation of the structural organization at early stages of Guinea fowl eggshell mineralization from 4 h to 5 h p.o. showed that initial process of calcification is identical to that observed in chickens (Rodriguez-Navarro AB et al., unpublished). Consequently, it is not surprising to observe, among the 15 most abundant Guinea fowl proteins at this stage, 10 proteins analogous to that observed in chicken at the same stage (OVAL, OVM, OVOT, HPX, LYZ C, OC-17-like, HAPLN3, CST3, HBAA and CLU) (Table 1). OVAL, LYZ, OC-17, HAPLN3 are proteins previously identified as overabundant at the initiation stage of chicken eggshell mineralization and are

known to play a direct role on the $CaCO_3$ precipitation (Figure 5) [6]. Lysozyme can also play a role in the transformation of ACC into calcite crystals [52-54], but does not stabilize the ACC phase [55]. OVAL locally increases the concentration of calcium ions, to create nucleation sites, to modify the morphology of calcite crystals and can promote ACC [6,54–57]. OVOT is thought to influence $CaCO_3$ nucleation and crystal growth in chicken [58], in addition to its well-known antibacterial role as iron chelator [59,60]. A similar role might be ascribed to HPX, which is a heme transporter able to bind metal ions [61,62]. It is overabundant at 4 h p.o. as in the chicken eggshell, nevertheless, its role in $CaCO_3$ formation is not yet confirmed. Proteoglycans are major actors in biocalcification and have been detected in eggshells [63–66]. These macromolecules, such as MEPE (OC-116-like) combine a protein core with negatively charged dermatan sulfate [67], which highly interact with calcium [68], suggesting a role in eggshell mineralization. HAPLN3, which possesses proteoglycan-binding properties, is also overabundant at the onset of calcification in both Guinea fowl and chicken eggshell [6,38,69].

Other proteins are indirectly related to the calcification process, as they are involved in the regulation of proteins driving mineralization. We also observed overabundant proteins belonging to this group in both chicken and Guinea fowl eggshell at the initiation stage of mineralization. OVM, a trypsin inhibitor [70], and CST3, a cysteine protease inhibitor [71] would have an indirect role in the calcification process by controlling the activity of eggshell matrix proteins,

either by inhibiting protein degradation or by modifying the maturation of precursor proteins [37,72,73].

Another protein, Lysyl oxidase like 2 (LOXL2), is not in the 15 major proteins at 4 h p.o. but is overabundant at this stage compared to other stages, as in the chicken [6]. This protein is responsible for collagen cross-linking and is localized in the eggshell membranes of the chicken egg [74,75]. Altogether, these results suggest that the primary events of transformation of ACC into calcite are associated with a similar group of proteins, which have a putative role in the precipitation and/or stabilization of $CaCO_3$, in eggshell membrane formation, and/or to participate in antimicrobial defense in both chicken and Guinea fowl.

Beside these proteins in common in both species, this study also revealed proteins that were overabundant in Guinea Fowl eggshell only at the onset of calcification. This is the case of Avidin-like (AVD-like), a heparin/biotin-binding protein, which plays an antimicrobial role [76–78]. We paid a particular interest in Cell Growth regulator with EF-hand domain protein-1 (CGREF1), which is one of the nine proteins only observed in Guinea fowl eggshell, because it possesses an EF-hand motif that would bind calcium ion with high affinity. Its concentration in the shell decreases during the shift (10-12 h p.o.) and it is not present in the shell at 18 h p.o.. Because of its species specificity, its likely calcium-binding property, and its overabundance at the onset of shell calcification, we propose that CGREF1 could be an important candidate for the

first nucleation events associated to mammillary knob mineralization during initial shell calcification.

Unique proteomic signature of the Guinea fowl eggshell main structural shift

In Guinea fowl eggshell, the result of the matrix-mineral interaction around 11 h p.o. is a new organization of mineral during which the size and orientation of calcite crystal units changes abruptly, which is responsible for its exceptional mechanical properties. We paid particular attention to this mineralization period in order to determine intra-crystalline matrix components and key molecular actors involved at this period. We observed an increase in the organic matrix content from 10 to 12 h p.o.. This result is in agreement with those observed in unpublished study of Rodriguez-Navarro AB et al., who observed a higher content of shell matrix during the main structural shift. Using proteomics, we identified 61 proteins only present during the shift period (from 10 h to 12 h p.o.; Supplementary data 5, sheet #1) and among them, 27 have their maximum of abundance at 11 h p.o. (Supplementary data 5, sheet #1). Additionally, 17 proteins were identified as differentially abundant, according to WS and/or XIC results, with an increased abundance from 4 h to the shift period and then the stabilization or the decrease in abundance at 18 h p.o. (Supplementary data 5, sheet #2). These proteins represent potential candidates of the secondary nucleation events and consequently we describe below in more detail their potential functions related to biomineral deposition.

We identified calcium-binding proteins during the shift (Supplementary data 5, sheet #1 and sheet #2; Figure 5). Nephronectin isoform X1 (NPNT X1), protein S100-A6, Calbindin D-28K (CALB1), Annexin A1 and A2 (ANXA1 and ANXA2) are only detected during the shift. ANXA1 and 2, only detected at 11 h p.o., and are known to induce the mineralization of dentin and endochondral bone [79,80]. ANXA2 was previously detected in the eggshell proteome of the chicken [6,81] and in the chorioallantoic membrane during embryonic development [82]. Protein S100-A6 was shift specific and overabundant at 11 h p.o.. This protein possesses two EF-hand domains suggesting that it binds calcium ions [83]. We also noticed in this group the presence of NPNT X1 and CALB1. Both Ca^{2+}-binding proteins were differentially abundant from 10 h to 12 h p.o.. CALB1 is the essential uterine transporter of calcium ions during eggshell mineralization and its abundance increased until 12 h p.o but its role in the shell is not clear. Cadherin-2 (CDH2) and EDIL3 X2 were not shift specific (Supplementary data 5, sheet #2) but were overabundant during the crystal shift period. CDH2 is a cell adhesion protein able to bind calcium ions [84–86] and was previously detected in the chicken eggshell proteome at 16h p.o. [6]. EDIL3 X1, which possesses three EGF-like domains, is also present in the chicken eggshell proteome, where its levels increase during the initial phase of shell mineralization process (from the first ACC deposit to the formation of large crystal units) [27,73]. In Guinea fowl, EDIL3 X1 was overabundant during the

main structural shift period then decreased at 18 h p.o. after the secondary nucleation events has occurred.

Proteoglycans, such as OC-116/MEPE [67], are strong candidates to regulate the biomineralization process. Tsukushin (TSKU) is a proteoglycan suspected to have a role in the eggshell biomineralization because of its high expression in the chicken uterus during eggshell formation [72]. The heparan sulfate proteoglycan, Glypican-4 (GPC4), appeared at 11 h and increased at 12 h p.o.. It was previously identified as an overabundant protein with high emPAI value during the active growth phase in the chicken eggshell [6]. Proteins that bind proteoglycans were also identified at the stage when crystal size and orientation change. They are Receptor-type tyrosine-protein phosphatase S (PTPRS) and Pleiotrophin (PTN), Vitronectin (VTN) and Follistatin-related protein 1 (FSTL1).

During the shift period, we also observed in Guinea fowl, the presence of proteins involved in the regulation of other proteins. These proteins are known to play an important role in the mineralization process by controlling the activity of proteins driving calcification [6]. Of particular interest is the extracellular serine/threonine protein kinase FAM20C, overabundant during the shift, which phosphorylates proteins related to $CaCO_3$ precipitation such as Osteopontin (SPP1/OPN) and Ovocleidin-116 (OC-116/MEPE) [40]. These last two proteins exhibit high affinity for calcium and belong to the small integrin-binding ligand N-linked glycoproteins (SIBLINGs) family. SPP1 was identified in the chicken,

quail and turkey eggshell and is proposed to play a critical role in chicken biomineralization [28,29,87–90]. We did not detect Guinea fowl SPP1 in the eggshell matrix during this study; however, the protein was immunolocalized in Guinea fowl eggshell (Rodriguez-Navarro AB et al., unpublished). Indeed, phosphoproteins require special proteomic approaches, and SPP1 was previously detected in chicken eggshell using a phosphoproteome methodology [87] not used in the present study. Olfactomedin-like protein 3 (OLFML3), was also detected during the shift period with a maximum at 11 h p.o.. It is an inhibitor of chordin (CHD) known to be a bone morphogenetic protein (BMP) antagonist [91–94]. We did not observe BMP in the Guinea fowl proteome but it is noteworthy that GDF6, a BMP ligand known to play a role in bones and joints formation of limbs [95,96], is specific to the period of the change in crystal size and orientation during the shift with an overabundance peak at 11 h p.o., when the first crystals composing the new layer are deposited. We suggest a regulation of chordin by OLFML3, which indirectly regulates GDF6 activity specifically during the change in the size and orientation of crystals. GDF6 was not identified in the eggshell proteome of other species (Figure 2), thus it can be considered as a key protein of the secondary nucleation events in Guinea fowl. Protein NOV homolog is another candidate that was only detected during the shift period; its function is associated with cartilage formation in mammals [97,98].

In addition to these functions, we notice the presence of molecular chaperone proteins essential for the proper folding of proteins, such as Clusterin (CLU) and Torsin-1B (TOR1B). CLU, overabundant during the shift (Supplementary data 5, sheet #2; Figure 5), and TOR1B specific to the shift (Supplementary data 5, sheet #1) have been suspected to prevent organic matrix aggregation during eggshell mineralization [6,49,99]. As observed in chicken eggshell, CLU abundance decreases at the end of mineralization; the abundance of TOR1B increases during the active calcification stage in chicken eggshell while it is not detected at the same timing in Guinea fowl.

Other potential candidates are Pigment epithelium-derived factor (SERPINF1) and Ovocalyxin-32 (OCX-32/RARRES1). SERPINF1 belongs to the serpin family but does not exhibit serine protease inhibitory activity [100]. SERPINF1 enhances osteoblast differentiation and is involved in mineral deposition and bone homeostasis [101]. OCX-32 is an eggshell protein [18], for which SNPs have been associated with eggshell quality traits [102].

The calcium carbonate biomineralization process requires bicarbonate in addition to calcium ions. All bird eggshell proteomes contain carbonic anhydrases (CAs) such as CA2 and CA4, which provide carbonate ions by catalysis of the reversible hydration of CO_2 [6,27–30]. CA4 is overabundant at 11 h p.o. in Guinea fowl eggshell, while it is more abundant at 16h post-ovulation in chicken eggshell [6]. Even if CA4 is not a shift-specific protein, it

can play an essential role in bicarbonate supply during the change of calcite crystal size and orientation.

Finally, other proteins, which are not shift-specific, can play a role in the main shift toolkit. Indeed, we noticed the presence of two key proteins with the highest abundance at this critical step in Guinea fowl eggshell formation: Dromaiocalcin-1-like (DCA-1-like) and Ocvocleidin-17-like (OC-17-like). Orthologs of these two paralogous C-type lectins are also present in other bird eggshells (Pekin duck, ostrich, emu and rhea) [48,50]. They are the most abundant proteins in the Guinea fowl shell and their abundance increases with the calcification process and reaches their maximum from 10 h to 18 h p.o.. This study identified for the first time DCA-1-like protein in Neognathae eggshell. OC-17-like has been characterized in chicken (OC-17), Pekin duck, zebra finch and goose (ansocalcin) [24,30,48,103–105]. In chicken, OC-17 is predicted to play a role in the eggshell mineralization process, and molecular simulation modeling suggests that it binds ACC to promote calcite formation [6,106–108]. In addition, in goose, *in vitro* crystallization experiments showed the influence of ansocalcin upon calcite crystal morphology [104]. In Guinea fowl these two homologous proteins could also play a role in eggshell formation by binding calcite and/or ACC [109]. These proteins are overabundant at the shift and during the formation of new layer in Guinea fowl eggshell, and consequently we suggest a predominant role for DCA-1-like and OC-17-like in the change of crystal size and orientation, and the formation of the new crystalline layer.

Conclusions

The present work is the first Guinea fowl eggshell proteomic study, which allowed the identification of 149 proteins. Comparison of the Guinea fowl eggshell proteome with that of other bird species leads to the identification of 9 proteins that are only present in that of Guinea fowl. Among them, S100-A6 and GDF6 were only found during the shift of crystal size and orientation. We hypothesize that they could play a role in shell mineralization and especially by controlling crystal size and/or orientation, for instance, by binding to calcite crystal faces and blocking growth which would modify the mineralization process. In addition to these two proteins, we also report that 61 proteins are more abundant during the secondary nucleation events associated with the change in crystal orientation and the formation of the new layer. Additionally, we identify the most abundant proteins involved in the different phases of Guinea fowl shell formation (Figure 5), from the first events of biomineralization until the deposition of the new layer. Our study showed that proteins associated with early mineralization are similar in chicken and Guinea fowl, but also revealed candidate proteins, which may be involved in a dramatic shift in eggshell microstructure that is unique to Guinea fowl. The proteins more abundant during the shift were ANXA2, S100A6, CALB1, TSKU, FAM20C, GPC-4, DCA-1-like and GDF6. Further investigation is needed to determine the

functional implications of these proteins in the secondary nucleation events that are specific to the Guinea fowl eggshell biomineralization process.

Acknowledgments

This work was funded by the Agence Nationale de la Recherche (ANR-13-BSV6-0007-01) Impact Project. Nathalie Le Roy's post-doctoral position was funded by an Agreenskills fellowship (2016-2018). Authors gratefully thank Mr. P. Jamenot from the Gallor breeder company, which provided the Guinea Fowl hens. We thank Dr. Mylène Da Silva for her kind assistance with Scaffold software. We are grateful to the genotoul bioinformatics platform Toulouse Midi-Pyrenees (Bioinfo Genotoul) for providing help, computing and storage resources. We thank the two anonymous reviewers that greatly improved the quality of the manuscript.

Conflict of interest

The authors have no conflicts of interest to declare.

Author statement

Nathalie Le Roy: Investigation, Conceptualization, Formal analysis, Writing – Original Draft. Lucie Combes-Soia: Validation, Formal Analysis. Aurélien Brionne: Software, Formal Analysis. Valérie Labas: Methodology, Validation, Data Curration. Alejandro B. Rodriguez-Navarro: Writting – Review and Editing. Maxwell T. Hincke: Writting – Review and Editing. Yves Nys:

Conceptualization, Writting – Review and Editing. Joël Gautron: Conceptualization, Supervision, Investigation, Funding Acquisition, Writting – Review and Editing.

Supplementary data

Supplementary data 1: Table of proteomics dataset for each biological replicate for the 5 post-ovulation stages of eggshell mineralization in *Numida meleagris*. Unique peptide counts are indicated for each technical replicate (tech. rep.) of each biological replicate (eggshell) in sheet #1. Average, standard deviation and variance of Exponentially Modified Protein Abundance Index (emPAI), Normalized Spectral Abundance Factor (NSAF), Weighted Spectral count (WS) and eXtracted Ion Chromatogram (XIC) values for proteins identified in biological replicates are indicated in sheet #2.

Supplementary data 2: List of the 13 proteins, identified as external contaminants and molecular variants, removed from the eggshell proteome of Guinea fowl (*Numida meleagris*).

Supplementary data 3: List of the 149 non-redundant proteins identified in the Guinea fowl eggshell at each stage of biomineralization. Mean of unique peptide count, emPAI, NSAF, WS, XIC, associated clusters of the WS heat

map, percent coverage, and protein identification probability are indicated. A list of proteins per stage is also provided in this file in distinct worksheets.

Supplementary data 4: Comparison between the Guinea fowl eggshell proteome and five other eggshell proteome studies (chicken, turkey, quail, zebra finch and crocodile). Lines in orange indicate proteins only present in *Numida meleagris*, in green are indicated proteins present in all six species and in blue proteins only present in eggshell proteome of birds.

Supplementary data 5: Proteins detected during the shift period of the Guinea fowl eggshell. List of 61 proteins only detected during the shift period (10 h, 11 h and/or 12 h p.o.) (sheet #1) and list of 17 differentially overabundant proteins, which decrease or stabilize at 18 h post-ovulation, in both WS and XIC quantification methods during the shift (sheet #2). DA: differentially abundant; NS: not significant.

Supplementary data 6: Results of the ANOVA and Tukey post-hoc tests on the WS and XIC values for each protein in the Guinea fowl eggshell proteome.

Supplementary data 7: Immunodetection by Western blot of OC-17-like and OVAL in the matrix proteins extracted from Guinea fowl eggshells at

five stages of biomineralization. Lane 1 corresponds to the protein ladder (kDa). For each extract, 10 µg of proteins were subjected to electrophoretic separation and electro-transfer: stage 4 h (lane 2), 10 h (lane 3), 11 h (lane 4), 12 h (lane 5) and 18 h p.o. (lane 6). The positive control (lane 7) corresponds to 10 µg proteins extracted from fully formed chicken eggshell.

References

[1] Y. Nys, J. Gautron, J.M. Garcia-Ruiz, M.T. Hincke, Avian eggshell mineralization: Biochemical and functional characterization of matrix proteins, CR Palevol. 3 (2004) 549–562. doi:10.1016/j.crpv.2004.08.002.

[2] M.T. Hincke, Y. Nys, J. Gautron, K. Mann, B. Rodriguez-Navarro, M.D. McKee, The eggshell: Structure, composition and mineralization, Front. Biosci. 17 (2012) 1266–1280. doi:10.2741/3985.

[3] J. Gautron, M.T. Hincke, Y. Nys, Precursor matrix proteins in the uterine fluid change with stages of eggshell formation in hens, Connect. Tissue Res. 36 (1997) 195–210. doi:10.3109/03008209709160220.

[4] A. Hernandez-Hernandez, M.L. Vidal, J. Gomez-Morales, A.B. Rodriguez-Navarro, V. Labas, J. Gautron, Y. Nys, J.M. Garcia Ruiz, Influence of eggshell matrix proteins on the precipitation of calcium carbonate ($CaCO_3$), J. Cryst. Growth. 310 (2008) 1754–1759. doi:10.1016/j.jcrysgro.2007.11.170.

[5] A.B. Rodríguez-Navarro, P. Marie, Y. Nys, M.T. Hincke, J. Gautron,

Amorphous calcium carbonate controls avian eggshell mineralization: A new paradigm for understanding rapid eggshell calcification, J. Struct. Biol. 190 (2015) 291–303. doi:10.1016/j.jsb.2015.04.014.

[6] P. Marie, V. Labas, A. Brionne, G. Harichaux, C. Hennequet-Antier, A.B. Rodríguez-Navarro, Y. Nys, J. Gautron, Quantitative proteomics provides new insights into chicken eggshell matrix protein functions during the primary events of mineralisation and the active calcification phase, J. Proteomics. 126 (2015) 140–154. doi:10.1016/j.jprot.2015.05.034.

[7] A.B. Rodríguez-Navarro, C.S. Romanek, Mineral fabrics analysis using a low-cost universal stage for X-ray diffractometry, Eur. J. Mineral. 14 (2002) 987–992. doi:10.1127/0935-1221/2002/0014-0987.

[8] I.C. Dunn, Poultry breeding for egg quality: Traditional and modern genetic approaches, in: Y. Nys, M. Bain, F. Van Imerseel (Eds.), Improv. Saf. Qual. Egg Prod., Woodhead publishing in food science, technology and nutrition, Cambridge, UK, 2011: pp. 245–260.

[9] A. Ar, C.V. Paganelli, R.B. Reeves, D.G. Greene, H. Rahn, The avian egg: Water vapor conductance, shell thickness and functional pore area, Condor. 76 (1974) 153–158. doi:10.2307/1366725.

[10] A. Ar, H. Rahn, C.V. Paganelli, The avian egg: Mass and strength, Condor. 81 (1979) 331–337. doi:10.2307/1366955.

[11] M. Panheleux, M. Bain, M.S. Fernandez, I. Morales, J. Gautron, J.L. Arias, S.E. Solomon, M. Hincke, Y. Nys, Organic matrix composition and

ultrastructure of eggshell: A comparative study, Br. Poult. Sci. 40 (1999) 240–252. doi:10.1080/00071669987665.

[12] S. Alkan, T. Karsli, A. Galic, K. Karabag, Determination of phenotypic correlations between internal and external quality traits of Guinea fowl eggs, Kafkas Univ. Vet. Fak. Derg. 19 (2013) 861–867. doi:10.9775/kvfd.2013.8988.

[13] M. Panheleux, O. Kälin, J. Gautron, Y. Nys, Features of eggshell formation in guinea fowl: Kinetics of shell deposition, uterine protein secretion and uterine histology, Br. Poult. Sci. 40 (1999) 632–43. doi:10.1080/00071669987025.

[14] K.T. Song, S.H. Choi, H.R. Oh, A comparison of egg quality of pheasant, chukar, quail and guinea fowl, Asian-Australasian J. Anim. Sci. (2000) 986–990. doi:10.5713/ajas.2000.986.

[15] J. Petersen, C. Tyler, The strength of guinea fowl (*Numida meleagris*) egg shells, Br. Poult. Sci. (1966) 291–296. doi:10.1080/00071666608415635.

[16] A. Danilova, I.S. Shpits, Chicken and guinea fowl egg shell strength and chemical content, in: Proc. 15th World's Poult. Congr., New Orleans, 1974: pp. 202–203.

[17] A. Pérez-Huerta, Y. Dauphin, Comparison of the structure, crystallography and composition of eggshells of the guinea fowl and graylag goose, Zoology. 119 (2016) 52–63. doi:10.1016/j.zool.2015.11.002.

[18] A.G. Checca, E. Macias-Sanchez, E.M. Harper, J.H.E. Cartwright, Organic membranes determine the pattern of the columnar prismatic lyer of mollusc shells, Proc. R. Soc. B. 283 (2016). doi:10.1098/rspb.2016.0032.

[19] O.B.A. Agbaje, R. Wirth, L.F.G. Morales, K. Shrai, M. Kosnik, T. Watanabe, D.E. Jacob, Architecture of crossed-lamellar bivalve shells: The southern giant clam (*Tridacna derasa,* Roding, 1798), R. Soc. Open sci. 4 (2017). doi:10.1098/rsos.170622.

[20] D. Chateigner, C. Hedegaard, H.-R. Wenk, Mollusc shell microstructures and crystallographic textures, J. Struct. Geol. 22 (2000) 1723-1735. doi:10.1016/S0191-8141(00)00088-2.

[21] F. Song, Y.L. Bai, Effects of nanostructures on the fracture strength of the interfaces in nacre, J. Mater. Res. 18 (2003) 1741-1744. doi:10.1557/JMR.2003.0239.

[22] R. Avery, R.J. Etter, Microstructural differences in the reinforcement of a gastropod shell against predation, Mar. Ecol. Prog. Ser. 323 (2006) 159-170. doi:10.3354/meps323159.

[23] J. Gautron, M.T. Hincke, K. Mann, M. Panheleux, M. Bain, M.D. McKee, S.E. Solomon, Y. Nys, Ovocalyxin-32, a novel chicken eggshell matrix protein. isolation, amino acid sequencing, cloning, and immunocytochemical localization, J. Biol. Chem. 276 (2001) 39243–3952. doi:10.1074/jbc.M104543200.

[24] M.T. Hincke, C.P.W. Tsang, M. Courtney, V. Hill, R. Narbaitz,

Purification and immunochemistry of a soluble matrix protein of the chicken eggshell (ovocleidin 17), Calcif. Tissue Int. 56 (1995) 578–583. doi:10.1007/BF00298593.

[25] A. Keller, A.I. Nesvizhskii, E. Kolker, R. Aebersold, Empirical statistical model To estimate the accuracy of peptide identifications made by MS/MS and database search, Anal. Chem. 74 (2002) 5383–5392. doi:10.1021/ac025747h.

[26] A.I. Nesvizhskii, A. Keller, E. Kolker, R. Aebersold, A statistical model for identifying proteins by tandem mass spectrometry, Anal. Chem. 75 (2003) 4646–58. doi:10.1021/ac0341261.

[27] K. Mann, B. Maček, J. V. Olsen, Proteomic analysis of the acid-soluble organic matrix of the chicken calcified eggshell layer, Proteomics. 6 (2006) 3801–3810. doi:10.1002/pmic.200600120.

[28] K. Mann, M. Mann, The proteome of the calcified layer organic matrix of turkey (*Meleagris gallopavo*) eggshell, Proteome Sci. 11 (2013) 40. doi:10.1186/1477-5956-11-40.

[29] K. Mann, M. Mann, Proteomic analysis of quail calcified eggshell matrix: A comparison to chicken and turkey eggshell proteomes, Proteome Sci. 13 (2015) 22. doi:10.1186/s12953-015-0078-1.

[30] K. Mann, The calcified eggshell matrix proteome of a songbird, the zebra finch (*Taeniopygia guttata*), Proteome Sci. 13 (2015) 1–20. doi:10.1186/s12953-015-0086-1.

[31] I. Mikšík, S. Paradis, A. Eckhardt, D. Sedmera, Analysis of siamese crocodile (*Crocodylus siamensis*) eggshell proteome, Protein J. 37 (2018) 21–37. doi:10.1007/s10930-017-9750-x.

[32] C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, T.L. Madden, BLAST+: Architecture and applications, BMC Bioinformatics. 10 (2009) 421. doi:10.1186/1471-2105-10-421.

[33] Y. Chiari, V. Cahais, N. Galtier, F. Delsuc, Phylogenomic analyses support the position of turtles as the sister group of birds and crocodiles (Archosauria), BMC Biol. 10 (2012) 65. doi:10.1186/1741-7007-10-65.

[34] C. Sayers, J. Wyatt, R.D. Soloway, D.R. Taylor, M.D. Stringer, Gallbladder mucin production and calcium carbonate gallstones in children, Pediatr. Surg. Int. 23 (2007) 219–223. doi:10.1007/s00383-006-1867-5.

[35] A. Kasprzak, W. Malkowski, C. Helak-Łapaj, A. Seraszek, E. Kaczmarek, A. Adamek, M. Zabel, Polysaccharides and mucin 5AC (MUC5AC) expression in gallbladder mucosa of young patients with gallstones as evaluated by spatial visualization and quantification, Folia Histochem. Cytobiol. 48 (2010) 646–657. doi:10.2478/v10042-010-0053-8.

[36] S. Claramunt, J. Cracraft, Evolutionary Ecology: A new time tree reveals Earth history's imprint on the evolution of modern birds, Sci. Adv. 1 (2015) 1–14. doi:10.1126/sciadv.1501005.

[37] V. Jonchère, S. Réhault-Godbert, C. Hennequet-Antier, C. Cabau, V.

Sibut, L.A. Cogburn, Y. Nys, J. Gautron, Gene expression profiling to identify eggshell proteins involved in physical defense of the chicken egg, BMC Genomics. 11 (2010) 57. doi:10.1186/1471-2164-11-57.

[38] M. Rose-Martel, S. Smiley, M.T. Hincke, Novel identification of matrix proteins involved in calcitic biomineralization, J. Proteomics. 116 (2015) 81–96. doi:10.1016/j.jprot.2015.01.002.

[39] R. Li, Q. Wu, Y. Zhao, W. Jin, X. Yuan, X. Wu, Y. Tang, J. Zhang, X. Tan, F. Bi, J.N. Liu, The novel pro-osteogenic activity of NUCB21-83, PLoS One. 8 (2013) 4–9. doi:10.1371/journal.pone.0061619.

[40] V.S. Tagliabracci, J.L. Engel, J. Wen, S.E. Wiley, C.A. Worby, L.N. Kinch, J. Xiao, N. V Grishin, J.E. Dixon, Secreted kinase phosphorylates extracellular proteins that regulate biomineralization, Science. 336 (2012) 1150–1153. doi:10.1126/science.1217817.

[41] J. Cui, J. Xiao, V.S. Tagliabracci, J. Wen, M. Rahdar, J.E. Dixon, A secretory kinase complex regulates extracellular protein phosphorylation, Elife. (2015) 1–18. doi:10.7554/eLife.06120.

[42] J. Du, C. Liu, G. Xu, J. Xie, L. Xie, R. Zhang, Fam20C participates in the shell formation in the pearl oyster, *Pinctada fucata*, Sci. Rep. 8 (2018) 1–9. doi:10.1038/s41598-018-21797-w.

[43] Y. Nys, X. de Laage, Effects of suppression of eggshell calcification and of 1,25$(OH)_2$D3 on $Mg2^+$, Ca2+ and $Mg^{2+}HCO_3^-$ ATPase, alkaline phosphatase, carbonic anhydrase and CaBP levels--I. The laying hen

uterus, Comp. Biochem. Physiol. A. Comp. Physiol. 78 (1984) 839–844.

[44] Y. Nys, K. Baker, D.E.M. Lawson, Estrogen and a calcium flux dependent factor modulate the calbindin gene expression in the uterus of laying hens, Gen. Comp. Endocrinol. 87 (1992) 87–94. doi:10.1016/0016-6480(92)90153-B.

[45] A. Bar, Differential regulation of calbindin in the calcium-transporting organs of birds with high calcium requirements, J. Poult. Sci. 46 (2009) 267–285. doi:10.2141/jpsa.46.267.

[46] V. Jonchère, A. Brionne, J. Gautron, Y. Nys, Identification of uterine ion transporters for mineralisation precursors of the avian eggshell, BMC Physiol. 12 (2012) 10. doi:10.1186/1472-6793-12-10.

[47] B.-A. Gotliv, N. Kessler, J.L. Sumerel, D.E. Morse, N. Tuross, L. Addadi, S. Weiner, Asprich: A novel aspartic acid-rich protein family from the prismatic shell matrix of the bivalve *Atrina rigida*, ChemBioChem. 6 (2005) 304–314. doi:10.1002/cbic.200400221.

[48] F. Zhu, F. Zhang, M. Hincke, Z.T. Yin, S.R. Chen, N. Yang, Z.C. Hou, iTRAQ-Based quantitative proteomic analysis of duck eggshell during biomineralization, Proteomics. 19 (2019) e1900011. doi:10.1002/pmic.201900011.

[49] K. Mann, F. Siedler, Ostrich (*Struthio camelus*) eggshell matrix contains two different C-type lectin-like proteins. Isolation, amino acid sequence, and posttranslational modifications, Biochim. Biophys. Acta - Proteins

Proteomics. 1696 (2004) 41–50. doi:10.1016/j.bbapap.2003.09.006.

[50] K. Mann, F. Siedler, Amino acid sequences and phosphorylation sites of emu and rhea eggshell C-type lectin-like proteins, Comp. Biochem. Physiol. - B Biochem. Mol. Biol. 143 (2006) 160–170. doi:10.1016/j.cbpb.2005.11.003.

[51] Y. Nys, J. Zawadzki, J. Gautron, A. Mills, Whitening of brown-shelled eggs: Mineral composition of uterine fluid and rate of protoporphyrin deposition, Poult. Sci. 70 (1991) 1236–1245. doi:10.3382/ps.0701236.

[52] M.T. Hincke, J. Gautron, M. Panheleux, J. Garcia-Ruiz, M.D. McKee, Y. Nys, Identification and localization of lysozyme as a component of eggshell membranes and eggshell matrix, Matrix Biol. 19 (2000) 443–453. doi:10.1016/S0945-053X(00)00095-0.

[53] C. Jimenez-Lopez, A. Rodriguez-Navarro, J.M. Dominguez-Vera, J.M. Garcia-Ruiz, Influence of lysozyme on the precipitation of calcium carbonate: A kinetic and morphologic study, Geochim. Cosmochim. Acta. 67 (2003) 1667–1676. doi:10.1016/S0016-7037(02)01275-9.

[54] K. Zhao, M. Wang, X. Wang, C. Wu, H. Xu, J.R. Lu, Crystal growth of calcite mediated by ovalbumin and lysozyme: Atomic force microscopy study, Cryst. Growth Des. 13 (2013) 1583–1589. doi:10.1021/cg301820w.

[55] S.E. Wolf, J. Leiterer, V. Pipich, R. Barrea, F. Emmerling, W. Tremel, Strong stabilization of amorphous calcium carbonate emulsion by ovalbumin: Gaining insight into the mechanism of “Polymer-Induced

Liquid Precursor" processes, J. Am. Chem. Soc. 133 (2011) 12642–12649. doi:10.1021/ja202622g.

[56] V. Pipich, M. Balz, S.E. Wolf, W. Tremel, D. Ju, Nucleation and growth of $CaCO_3$ mediated by the egg-white protein ovalbumin: A time-resolved in situ study using small-angle neutron scattering, J. Am. Chem. 130 (2008) 6879–6892. doi:10.1021/ja801798h.

[57] X. Wang, C. Wu, K. Tao, K. Zhao, J. Wang, H. Xu, D. Xia, H. Shan, J.R. Lu, Influence of ovalbumin on $CaCO_3$ precipitation during in vitro biomineralization, J. Phys. Chem. B. 114 (2010) 5301–5308. doi:10.1021/jp1008237.

[58] J. Gautron, M.T. Hincke, M. Panheleux, J.M. Garcia-Ruiz, T. Boldicke, Y. Nys, Ovotransferrin is a matrix protein of the hen eggshell membranes and basal calcified layer, Connect. Tissue Res. 42 (2001) 255–267. doi:10.3109/03008200109016840.

[59] P. Valenti, G. Antonini, C. Von Hunolstein, P. Visca, N. Orsi, E. Antonini, Studies of the antimicrobial activity of ovotransferrin, Int. J. Tissue React. 5 (1983) 97–105.

[60] F. Baron, S. Jan, F. Gonnet, M. Pasco, J. Jardin, B. Giudici, M. Gautier, C. Guérin-Dubiard, F. Nau, Ovotransferrin plays a major role in the strong bactericidal effect of egg white against the Bacillus cereus group, J. Food Prot. 77 (2014) 955–62. doi:10.4315/0362-028X.JFP-13-473.

[61] G. Grieninger, T.J. Liang, G. Beuving, V. Goldfarb, S.A. Metcalfe, U.

Muller-Eberhard, Hemopexin is a developmentally regulated, acute-phase plasma protein in the chicken, J. Biol. Chem. 261 (1986) 15719–15724.

[62] M.R. Mauk, F.I. Rosell, B. Lelj-Garolla, G.R. Moore, A.G. Mauk, Metal ion binding to human hemopexin, Biochemistry. 44 (2005) 1864–1871. doi:10.1021/bi0481747.

[63] D.A. Carrino, J.E. Dennis, T.M. Wu, J.L. Arias, M.S. Fernandez, J.P. Rodriguez, D.J. Fink, A.H. Heuer, A.I. Caplan, The avian eggshell extracellular matrix as a model for biomineralization, Connect. Tissue Res. 35 (1996) 325–9. doi:10.3109/03008209609029207.

[64] M.S. Fernandez, M. Araya, J.L. Arias, Eggshells are shaped by a precise spatio-temporal arrangement of sequentially deposited macromolecules, Matrix Biol. 16 (1997) 13–20. doi:10.1016/S0945-053X(97)90112-8.

[65] T. Nakano, N. Ikawa, L. Ozimek, Extraction of glycosaminoglycans from chicken eggshell, Poult. Sci. 80 (2001) 681–684. doi:10.1093/ps/80.5.681.

[66] T. Nakano, N. Ikawa, L. Ozimek, Galactosaminoglycan composition in chicken eggshell, Poult. Sci. 81 (2002) 709–714. doi:10.1093/ps/81.5.709.

[67] M.T. Hincke, J. Gautron, P. Charles, W. Tsang, M.D. Mckee, Y. Nys, C.P.W. Tsang, Molecular cloning and ultrastructural localization of the core protein of a molecular cloning and ultrastructural localization of the core protein of an eggshell matrix proteoglycan, ovocleidin-116, J. Biol. Chem. 274 (1999) 32915–32923. doi:10.1074/jbc.274.46.32915.

[68] G.K. Hunter, K.S. Wong, J.J. Kim, Binding of calcium to

glycosaminoglycans: An equilibrium dialysis study, Arch. Biochem. Biophys. 260 (1988) 161–167. doi:10.1016/0003-9861(88)90437-7.

[69] A.P. Spicer, A. Joo, R.A. Bowling, A hyaluronan binding link protein gene family whose members are physically linked adjacent to chrondroitin sulfate proteoglycan core protein genes. The missing links, J. Biol. Chem. 278 (2003) 21083–21091. doi:10.1074/jbc.M213100200.

[70] R.E. Feeney, F.C. Stevens, D.T. Osuga, The specificities of chicken ovomucoid and ovoinhibitor, J. Biol. Chem. 238 (1963) 1415–1418.

[71] M. Abrahamson, A. Ritonja, M. Brown, A. Grubb, W. Machleidt, A.J. Barrett, Identification of the probable inhibitory reactive sites of the cysteine protease inhibitors human cystatin C and chicken cystatin, J. Biol. Chem. 202 (1987) 9688–9694.

[72] A. Brionne, Y. Nys, C. Hennequet-Antier, J. Gautron, Hen uterine gene expression profiling during eggshell formation reveals putative proteins involved in the supply of minerals or in the shell mineralization process, BMC Genomics. 15 (2014) 220. doi:10.1186/1471-2164-15-220.

[73] P. Marie, V. Labas, A. Brionne, G. Harichaux, C. Hennequet-Antier, Y. Nys, J. Gautron, Quantitative proteomics and bioinformatic analysis provide new insight into protein function during avian eggshell biomineralization, J. Proteomics. 113 (2015) 178–193. doi:10.1016/j.jprot.2014.09.024.

[74] K. Kaweewong, W. Garnjanagoonchorn, W. Jirapakkul, S. Roytrakul,

Solubilization and identification of hen eggshell membrane proteins during different times of chicken embryo development using the proteomic approach, Protein J. 32 (2013) 297–308. doi:10.1007/s10930-013-9487-0.

[75] T.A.E. Ahmed, H.P. Suso, M.T. Hincke, In-depth comparative analysis of the chicken eggshell membrane proteome, J. Proteomics. 155 (2017) 49–62. doi:10.1016/j.jprot.2017.01.002.

[76] H.B. White, W.H. Orth, R.W. Schreiber, C.C. Whitehead, Availability of avidin-bound biotin to the chicken embryo, Arch. Biochem. Biophys. 298 (1992) 80–83. doi:10.1016/0003-9861(92)90096-F.

[77] W.C. Kett, R.I.W. Osmond, L. Moe, S.E. Skett, B.F. Kinnear, D.R. Coombe, Avidin is a heparin-binding protein. Affinity, specificity and structural analysis, Biochim. Biophys. Acta - Gen. Subj. 1620 (2003) 225–234. doi:10.1016/S0304-4165(02)00539-1.

[78] N. Guyot, V. Labas, G. Harichaux, M. Chessé, J.C. Poirier, Y. Nys, S. Réhault-Godbert, Proteomic analysis of egg white heparin-binding proteins: Towards the identification of natural antibacterial molecules, Sci. Rep. 6 (2016) 1–11. doi:10.1038/srep27974.

[79] E.E. Golub, Role of matrix vesicles in biomineralization, Biochim. Biophys. Acta. 1790 (2009) 1592–1598. doi:10.1016/j.bbagen.2009.09.006.Role.

[80] R.E. Wuthier, Matrix vesicles: Structure, composition, formation and function in calcification, Front. Biosci. 16 (2011) 2812–2902.

doi:10.2741/3887.

[81] C.M.M. Cordeiro, M.T. Hincke, Quantitative proteomics analysis of eggshell membrane proteins during chick embryonic development, J. Proteomics. 130 (2016) 11–25. doi:10.1016/j.jprot.2015.08.014.

[82] K. Matschke, L. Da Silva-Azevedo, R. Hlushchuk, V. Djonov, O. Baum, Annexins as cell-type-specific markers in the developing chicken chorionallantoic membrane, Cell Tissue Res. 323 (2006) 395–404. doi:10.1007/s00441-005-0112-1.

[83] C.-W. Cheng, A. Rifai, S.-M. Ka, H.-A. Shui, Y.-F. Lin, W.-H. Lee, A. Chen, Calcium-binding proteins annexin A2 and S100A6 are sensors of tubular injury and recovery in acute renal failure, Kidney Int. 68 (2005) 2694–2703. doi:10.1111/j.1523-1755.2005.00740.x.

[84] B. Nagar, M. Overduin, M. Ikura, J.M. Rini, Structural basis of calcium-induced E-cadherin rigidification and dimerization, Nature. 380 (1996) 360–364. doi:10.1038/380360a0.

[85] E. Parisini, J.M.G. Higgins, J. huan Liu, M.B. Brenner, J. huai Wang, The crystal structure of human E-cadherin domains 1 and 2, and comparison with other cadherins in the context of adhesion mechanism, J. Mol. Biol. 373 (2007) 401–411. doi:10.1016/j.jmb.2007.08.011.

[86] N. Vunnam, S. Pedigo, Sequential binding of calcium leads to dimerization in neural cadherin, Biochemistry. 50 (2011) 2973–2982. doi:10.1021/bi101872b.

[87] K. Mann, J. V Olsen, B. Macek, F. Gnad, M. Mann, Phosphoproteins of the chicken eggshell calcified layer, Proteomics. 7 (2007) 106–115. doi:10.1002/pmic.200600635.

[88] M.T. Hincke, Y.-C. Chien, L.C. Gerstenfeld, M.D. McKee, Colloidal-gold immunocytochemical localization of osteopontin in avian eggshell gland and eggshell, J. Histochem. Cytochem. 56 (2008) 467–476. doi:10.1369/jhc.2008.950576.

[89] Y.C. Chien, M.T. Hincke, H. Vali, M.D. McKee, Ultrastructural matrix-mineral relationships in avian eggshell, and effects of osteopontin on calcite growth *in vitro*, J. Struct. Biol. 163 (2008) 84–99. doi:10.1016/j.jsb.2008.04.008.

[90] D. Athanasiadou, W. Jiang, D. Goldbaum, A. Saleem, K. Basu, M.S. Pacella, C.F. Böhm, R.R. Chromik, M.T. Hincke, A.B. Rodríguez-Navarro, H. Vali, S.E. Wolf, J.J. Gray, K.H. Bui, M.D. McKee, Nanostructure, osteopontin, and mechanical properties of calcitic avian eggshell, Sci. Adv. 4 (2018) 1–14. doi:10.1126/sciadv.aar3219.

[91] J. Larraín, D. Bachiller, B. Lu, E. Agius, Piccolo, D. Robertis, BMP-binding modules in chordin: A model for signalling regulation in the extracellular space, Development. 127 (2000) 821–830. doi:10.1016/j.drugalcdep.2008.02.002.A.

[92] R. Jasuja, B.L. Allen, W.N. Pappano, A.C. Rapraeger, D.S. Greenspan, Cell-surface heparan sulfate proteoglycans potentiate chordin antagonism

of bone morphogenetic protein signaling and are necessary for cellular uptake of chordin, J. Biol. Chem. 279 (2004) 51289–51297. doi:10.1074/jbc.M408129200.

[93] R.M. Harland, A protein scaffold plays matchmaker for Chordin, Cell. 134 (2008) 718–719. doi:10.1016/j.cell.2008.08.024.

[94] D. Umulis, M.B. O'Connor, S.S. Blair, The extracellular regulation of bone morphogenetic protein signaling, Development. 136 (2009) 3715–3728. doi:10.1242/dev.031534.

[95] S.H. Settle, R.B. Rountree, A. Sinha, A. Thacker, K. Higgins, D.M. Kingsley, Multiple joint and skeletal patterning defects caused by single and double mutations in the mouse Gdf6 and Gdf5 genes, Dev. Biol. 254 (2003) 116–130. doi:10.1016/S0012-1606(02)00022-2.

[96] V.B. Indjeian, G.A. Kingman, F.C. Jones, C.A. Guenther, J. Grimwood, J. Schmutz, R.M. Myers, D.M. Kingsley, Evolving new skeletal traits by cis-regulatory changes in bone morphogenetic proteins, Cell. 164 (2016) 45–56. doi:10.1016/j.cell.2015.12.007.

[97] H. Kawaki, S. Kubota, A. Suzuki, N. Lazar, T. Yamada, T. Matsumura, T. Ohgawara, T. Maeda, B. Perbal, K.M. Lyons, M. Takigawa, Cooperative regulation of chondrocyte differentiation by CCN2 and CCN3 shown by a comprehensive analysis of the CCN family proteins in cartilage, J. Bone Miner. Res. 23 (2008) 1751–1764. doi:10.1359/jbmr.080615.

[98] D. Janune, S. Kubota, T. Nishida, H. Kawaki, B. Perbal, S. Iida, M.

Takigawa, Novel effects of CCN3 that may direct the differentiation of chondrocytes, FEBS Lett. 585 (2011) 3033–3040. doi:10.1016/j.febslet.2011.08.024.

[99] P.J. McLean, H. Kawamata, S. Shariff, J. Hewett, N. Sharma, K. Ueda, X.O. Breakefield, B.T. Hyman, TorsinA and heat shock proteins act as molecular chaperones: Suppression of α-synuclein aggregation, J. Neurochem. 83 (2002) 846–854. doi:10.1046/j.1471-4159.2002.01190.x.

[100] T. Kawaguchi, S.-I. Yamagishi, M. Sata, Structure-function relationships of PEDF, Curr. Mol. Med. 10 (2010) 302–311. doi:10.2174/156652410791065255.

[101] M. Baxter-Holland, C.R. Dass, Pigment epithelium-derived factor: A key mediator in bone homeostasis and potential for bone regenerative therapy, J. Pharm. Pharmacol. 70 (2018) 1127–1138. doi:10.1111/jphp.12942.

[102] J.E. Fulton, M. Soller, A.R. Lund, J. Arango, E. Lipkin, Variation in the ovocalyxin-32 gene in commercial egg-laying chickens and its relationship with egg production and egg quality traits, Anim. Genet. 43 (2012) 102–113. doi:10.1111/j.1365-2052.2012.02384.x.

[103] K. Mann, F. Siedler, The amino acid sequence of Ovocleidin-17, a major protein of the avian eggshell calcified layer, Biochem. Mol. Biol. Int. 47 (1999) 997–1007.

[104] R. Lakshminarayanan, R.M. Kini, S. Valiyaveettil, Investigation of the role of ansocalcin in the biomineralization in goose eggshell matrix, Proc.

Natl. Acad. Sci. USA. 99 (2002) 5155–5159. doi:10.1073/pnas.072658899.

[105] Q. Zhang, L. Liu, F. Zhu, Z. Ning, M.T. Hincke, N. Yang, Z. Hou, Integrating *de novo* transcriptome assembly and cloning to obtain chicken ovocleidin-17 full-length cDNA, PLoS One. 9 (2014). doi:10.1371/journal.pone.0093452.

[106] J.P. Reyes-Grajeda, A. Moreno, A. Romero, Crystal structure of ovocleidin-17, a major protein of the calcified *Gallus gallus* eggshell: Implications in the calcite mineral growth pattern, J. Biol. Chem. 279 (2004) 40876–40881. doi:10.1074/jbc.M406033200.

[107] C.L. Freeman, J.H. Harding, D. Quigley, P.M. Rodger, Simulations of ovocleidin-17 binding to calcite surfaces and its implications for eggshell formation, J. Phys. Chem. C. 115 (2011) 8175–8183. doi:10.1021/jp200145m.

[108] C.L. Freeman, J.H. Harding, D. Quigley, P.M. Rodger, Structural control of crystal nuclei by an eggshell protein, Angew. Chemie - Int. Ed. 49 (2010) 5135–5137. doi:10.1002/anie.201000679.

[109] A. Wallace, J. Schiffbauer, Proteins from the past, Elife. 5 (2016) 1–4. doi:10.7554/eLife.17092.

Figure captions

Figure 1: Experimental design of the nanoLC-MS/MS analysis on eggshell samples at five post-ovulation stages in *Numida meleagris*. Number of Guinea fowls (n) and sampled eggshells (E, i.e. biological replicates) are indicated. A scheme represents the eggshell at each five stages of mineralization. Step #1 represents the rapid migration on 12.5% SDS-PAGE of 15 µg of matrix proteins extracted from each biological replicate, the excision of a single band/biological replicate (as indicated by the red dashed box), their trypsin digestion and nanoLC-MS/MS analysis. Step #2 represents the entire migration on 4-20% SDS-PAGE of 80 µg of pooled matrix proteins extracted from biological replicates (16 µg/biological replicate at 4 h p.o. and 13.33 µg/biological replicate for each 10 h, 11 h, 12 h and 18 h p.o. stages). Twenty gel slices were excised from each lane and their corresponding numbers are indicated on the right of the gel. Excised bands (as indicated by the red dashed box) were digested with trypsin and analyzed using nanoLC-MS/MS. ACC: amorphous calcium carbonate; E: eggshell biological replicate; MW: molecular weight in kDa (left of the gel).

Figure 2: Comparison of *Numida meleagris* eggshell proteins with those of other archosaurian species. Comparison of Guinea fowl, four other birds and

one crocodile species eggshell proteomes (A) [6,27-31]. The number of proteins identified by proteomics approaches is indicated in color below the name of the corresponding species. Common proteins between two species are indicated at the intersection of the two blocks. Venn diagram represents the shared and specific proteins between species (B). Unique proteins to the *N. meleagris* eggshell proteome are indicated in the green circle and are listed with their short name, gene ID and presence in the different stages of eggshell deposition. Yellow stars indicate proteins with putative biomineral interactions. The number of proteins common to the different species observed in (B) is reported on the phylogenetic tree representing species relationships according to tree of life web project (www.tolweb.org) (C). Divergent times indicated at Archosauria, Aves and Neognathae nodes come from [36].

Figure 3: Distribution of proteins identified in the eggshell soluble matrix *of Numida meleagris* at 4 h, 10-12 h and 18 h post-ovulation. Histogram represents the number of identified proteins (A) and Venn diagrams (B) indicate the distribution of identified proteins at 4 h, 10-12 h and 18 h p.o. stages and specifically during the shift of eggshell biomineralization. Proteins only present at 4 h and 18 h p.o. are listed with their short name and gene ID.

Figure 4: Quantity of matrix proteins per gram of eggshell and eggshell weight for each stage. Results of the Kruskal-Wallis statistical test with

pairwise comparison between stages are indicated with different letters ($p<0.05$).

Figure 5: Schematic depiction of the most abundant eggshell proteins during shell deposition in *Numida meleagris*. The most abundant proteins, according to their emPAI value, are listed at each stage. Differentially abundant proteins are colored in orange, according to their WS and/or XIC values. Proteins absent from the chicken eggshell proteome at the corresponding stage are underlined. Framed proteins are absent in other Neognathae birds. Proteins specific to the shift that are putatively involved in the biomineralization process are in italics and a yellow star indicates predicted calcium-binding properties. Data for chicken are extracted from [6]. Scanning electronic microscopy images (left) illustrate the ultrastructure of the Guinea fowl eggshell at three stages of shell formation.

Table 1: List of the 15 most abundant proteins at the five stages of the Guinea fowl eggshell biomineralization. Average of emPAI values are indicated for each protein.

4 h p.o.			10 h p.o.			11 h p.o.			12 h p.o.			18 h p.o.		
Protein	**Accession N°**	**emPAI average**	**Protein**	**Accession N°**	**emPAI average**	**Protein**	**Accession N°**	**emPAI average**	**Protein**	**Accession N°**	**emPAI average**	**Protein**	**Accession N°**	**emPAI average**
OVALX2	XP_021241976.1	261,85	**DCA-1-like**	XP_021238988.1	768,89	**DCA-1-like**	XP_021238988.1	655,14	**DCA-1-like**	XP_021238988.1	738,67	**DCA-1-like**	XP_021238988.1	809,11
ExFABX2	XP_021269747.1	144,99	**ExFABX2**	XP_021269747.1	77,31	**OC-17-like**	XP_021238987.1	148,38	**OC-17-like**	XP_021238987.1	123,12	**OC-17-like**	XP_021238987.1	121,68
OVM	XP_021266564.1	108,99	**OC-17-like**	XP_021238987.1	39,00	**OVALX2**	XP_021241976.1	29,44	**ExFABX2**	XP_021269747.1	62,65	**ExFABX2**	XP_021269747.1	40,52
OV	XP_0212	83,	**OV**	XP_0212	19,	**OC-**	XP_0212	23,	**OC-**	XP_0212	14,	**OC-**	XP_0212	15,

OT	4948 2.1	22	**ALX2**	4197 6.1	10	**116-like /MEPE-like**	5177 9.1	01	**116-like /MEPE-like**	5177 9.1	88	**116-like /MEPE-like**	5177 9.1	22
HPX	XP_0212 3953 7.1	66,03	**LYZC**	P00 704. 2	17,71	**OVM**	XP_0212 6656 4.1	22,46	**OVM**	XP_0212 6656 4.1	10,39	**LYZC**	P00 704. 2	4,78
LYZC	P00 704. 2	43,43	**OVM**	XP_0212 6656 4.1	14,52	**ExFABX2**	XP_0212 6974 7.1	20,06	**LYZC**	P00 704. 2	9,34	**OVOT**	XP_0212 4948 2.1	3,19
DCA-1-like	XP_0212 3898 8.1	42,31	**OC-116-like /MEPE-like**	XP_0212 5177 9.1	10,34	**LYZC**	P00 704. 2	19,82	**OVALX2**	XP_0212 4197 6.1	6,14	**Meleagrin-like**	XP_0212 4810 0.1	2,65
OC-17-like	XP_0212 3898 7.1	27,71	**OVOT**	XP_0212 4948 2.1	8,47	**OVOT**	XP_0212 4948 2.1	19,20	**OVOT**	XP_0212 4948 2.1	6,14	**ALB**	XP_0212 4969 0.1	2,40
OVAL	XP_0212 4420	26,0	**OCX-32-**	XP_0212 4934	3,86	**OCX-32-**	XP_0212 4934	5,14	**CLU**	XP_0212 4772	3,16	**SBP-1-lik**	XP_0212 4810	1,85

Y	3.1	4	**like /RARRES-like**	1.1		**like /RARRES-like**	1.1			5.1		**e**	1.1		
HAPLN3	XP_021263163.1	17,94	**AGP-like**	XP_021269241.1	3,56	**AVD-like**	XP_021236736.1	5,13	**OCX-32-like /RARRES-like**	XP_021249341.1	3,10	**CST3**	XP_021248732.1	1,80	
AVD-like	XP_021236736.1	17,14	**EDIL3X2**	XP_021235778.1	3,37	**AGP-like**	XP_021269241.1	4,07	**ALB**	XP_021249690.1	2,49	**CLU**	XP_021247725.1	1,36	
AGP-like	XP_021269241.1	8,62	**HPX**	XP_021239537.1	3,28	**EDIL3X2**	XP_021235778.1	3,96	**EDIL3X2**	XP_021235778.1	2,35	**B2M**	XP_021262402.1	1,28	
CST3	XP_021248732.1	7,51	**AVD-lik**	XP_021236736.1	2,98	**CLU**	XP_021247725.1	3,86	**CST3**	XP_021248732.1	2,25	**OVAL**	XP_021241976.1	1,18	

HBAA	XP_021267466.1	5,99	**e** **CST3**	XP_021248732.1	2,52	**HPX**	XP_021239537.1	3,24	**AGP-like**	XP_021269241.1	1,97	**X2** **OVM**	XP_021266564.1	1,17
CLU	XP_021247725.1	5,32	**CLU**	XP_021247725.1	2,39	**CST3**	XP_021248732.1	3,17	**HPX**	XP_021239537.1	1,59	**AGP-like**	XP_021269241.1	1,16

Table 2: The nine proteins uniquely identified in the Guinea fowl eggshell compared to other Neognathae, and their relative abundance according to Weight Spectra (WS) average at each stage.

Identified Proteins (9)	Protein short name	Accession N° (NCBI)	Gene ID	MW (kDa)	WS average				
					4 h	10 h	11 h	12 h	18 h
dromaiocalcin-1-like	**DCA-1-like**	XP_021238988.1	110391424	21	354.6	2854.7	1656.0	2860.0	3604.5
phosphatidylethanolamine-binding protein 4 isoform X1	**PEBP4 X1**	XP_021230598.1	11038711 4	31	2.4	29.9	33.1	27.9	21.8
small basic protein 1-like	**SBP-1-like**	XP_021248101.1	110396639	7	0.0	2.6	3.4	2.4	5.1
uncharacterized protein At5g39570-like	**LOC110408336**	XP_021272630.1	110408336	30	0.5	1.2	2.0	5.1	12.2
ribonuclease UK114	**RIDA**	XP_021241925.1	110393399	14	3.5	26.6	13.8	12.6	1.4
cell growth regulator with EF hand domain protein 1	**CGREF1**	XP_021248455.1	110396884	26	10.6	5.0	0.6	1.3	0.0
granulins	**GRN**	XP_021233282.1	110388392	29	0.0	1.8	3.1	4.5	8.8

growth/differentiation factor 6	**GDF6**	OWK61448.1	1104001002	43	0.0	7.7	6.4	6.2	0.0
sushi repeat-containing protein SRPX isoform X1	**SRPX**	XP_015679515.1	1072947700	51	0.0	2.0	3.1	0.6	0.0

Highlights

- The Guinea fowl eggshell is twice as strong as chicken eggshell because of a dramatic shift in calcite crystal size and orientation.
- We analyzed the protein composition and abundance at five stages of Guinea fowl eggshell biomineralization.
- We identified 149 proteins in Guinea fowl eggshell using the nano-LC MS/MS approach.
- Sixty-one (61) proteins are unique to the zone of microstructural shift, and are potential candidates to regulate this change in calcite texture.

Graphical abstract

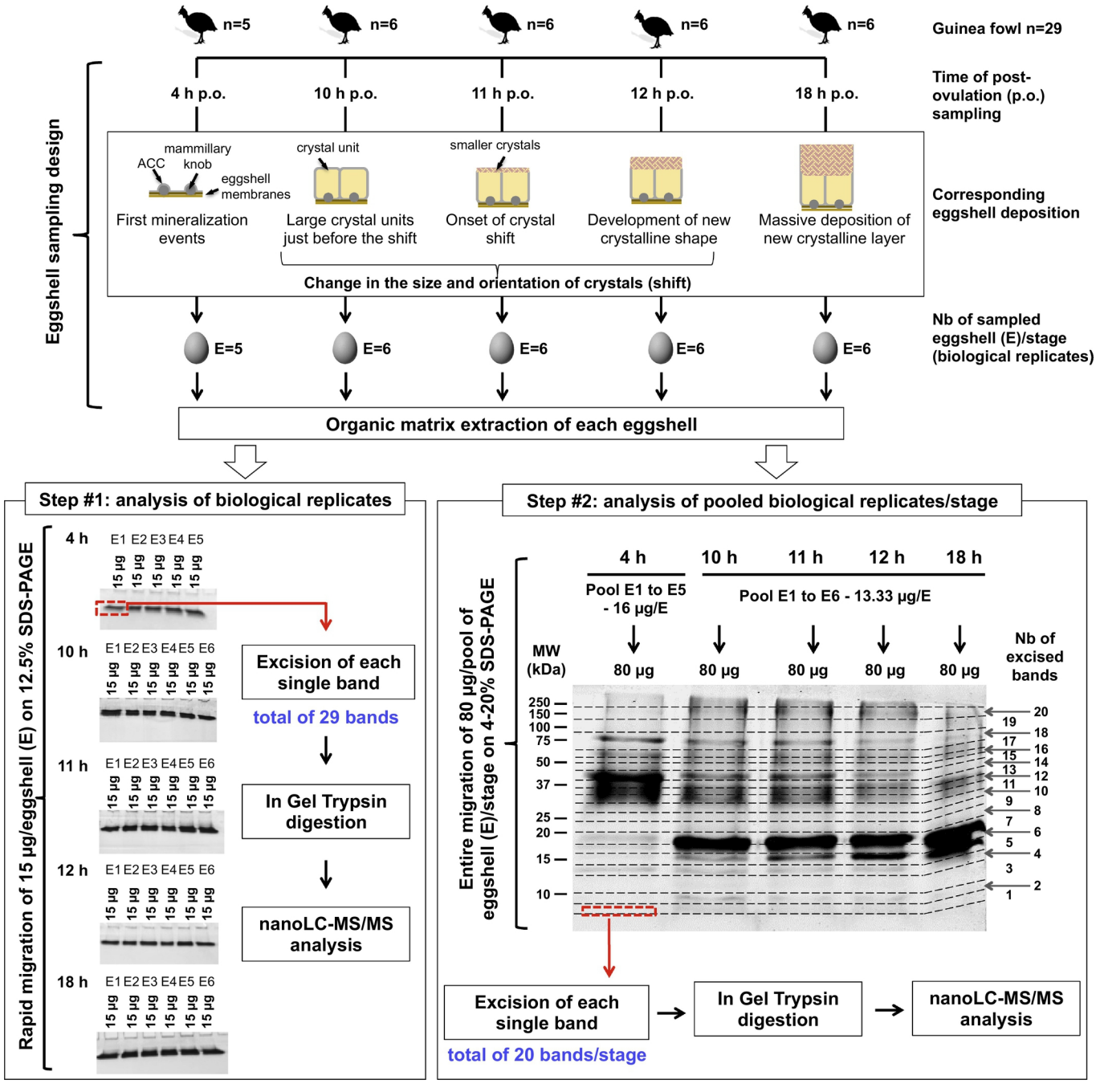

n=5
n=6
n=6
n=6
n=6
Guinea fowl n=29
Eggshell sampling design
4 h p.o.
10 h p.o.
11 h p.o.
12 h p.o.
18 h p.o.
Time of post-ovulation (p.o.) sampling
mammillary knob
ACC
eggshell membranes
crystal unit
smaller crystals
First mineralization events
Large crystal units just before the shift
Onset of crystal shift
Development of new crystalline shape
Massive deposition of new crystalline layer
Corresponding eggshell deposition
Change in the size and orientation of crystals (shift)
E=5
E=6
E=6
E=6
E=6
Nb of sampled eggshell (E)/stage (biological replicates)
Organic matrix extraction of each eggshell
Step #1: analysis of biological replicates
Rapid migration of 15 µg/eggshell (E) on 12.5% SDS-PAGE
4 h
10 h
11 h
12 h
18 h
E1 E2 E3 E4 E5
E1 E2 E3 E4 E5 E6
15 µg
Excision of each single band
total of 29 bands
In Gel Trypsin digestion
nanoLC-MS/MS analysis
Step #2: analysis of pooled biological replicates/stage
Entire migration of 80 µg/pool of eggshell (E)/stage on 4-20% SDS-PAGE
4 h
10 h
11 h
12 h
18 h
Pool E1 to E5 - 16 µg/E
Pool E1 to E6 - 13.33 µg/E
MW (kDa)
80 µg
Nb of excised bands
250
150
100
75
50
37
25
20
15
10
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
Excision of each single band
In Gel Trypsin digestion
nanoLC-MS/MS analysis
total of 20 bands/stage

Figure 1

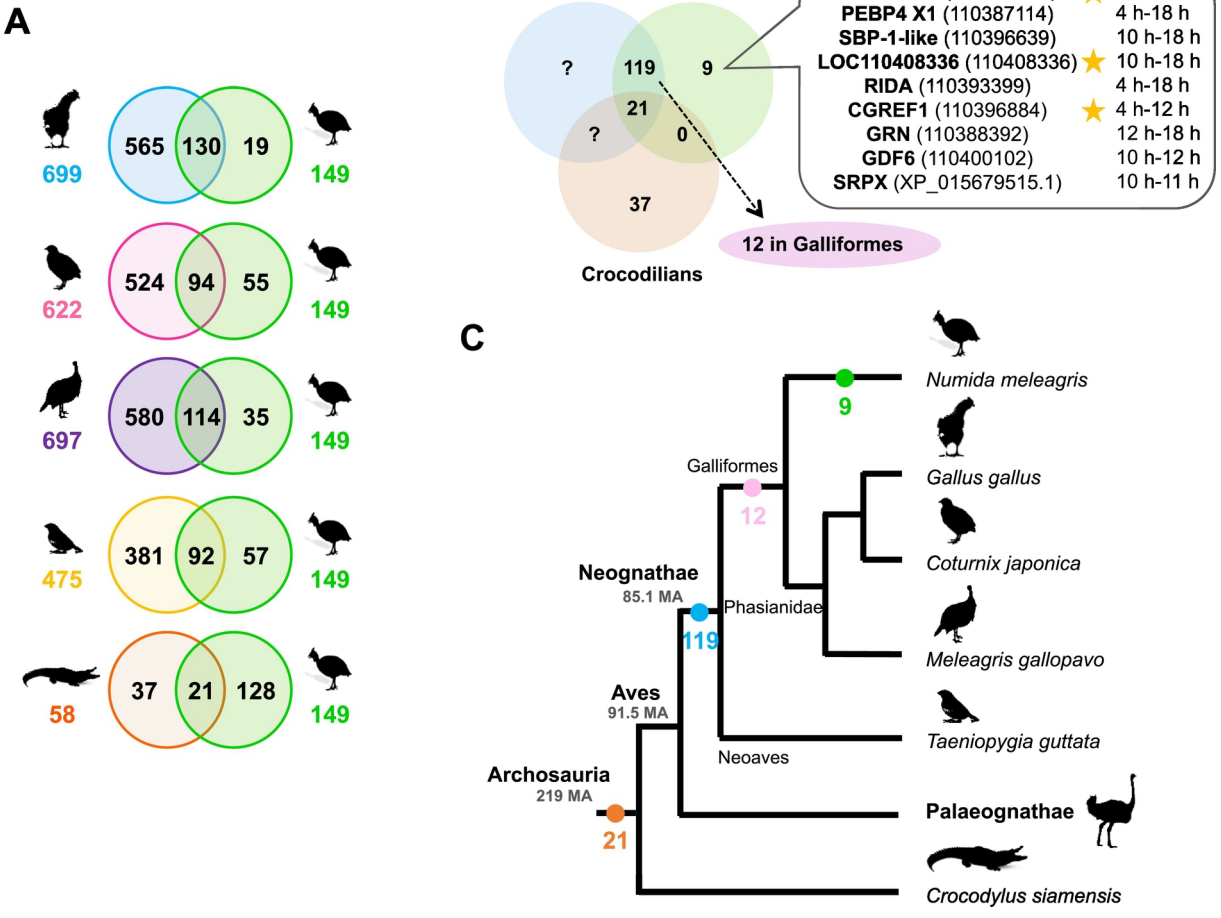

A
699
565
130
19
149
622
524
94
55
149
697
580
114
35
149
475
381
92
57
149
58
37
21
128
149
B
Neognathae
Guinea fowl
?
119
9
21
?
0
37
Crocodilians
12 in Galliformes
DCA-1-like (110391424) 4 h-18 h
PEBP4 X1 (110387114) 4 h-18 h
SBP-1-like (110396639) 10 h-18 h
LOC110408336 (110408336) 10 h-18 h
RIDA (110393399) 4 h-18 h
CGREF1 (110396884) 4 h-12 h
GRN (110388392) 12 h-18 h
GDF6 (110400102) 10 h-12 h
SRPX (XP_015679515.1) 10 h-11 h
C
Numida meleagris
9
Galliformes
12
Gallus gallus
Coturnix japonica
Neognathae
85.1 MA
Phasianidae
119
Meleagris gallopavo
Aves
91.5 MA
Taeniopygia guttata
Neoaves
Archosauria
219 MA
Palaeognathae
21
Crocodylus siamensis

Figure 2

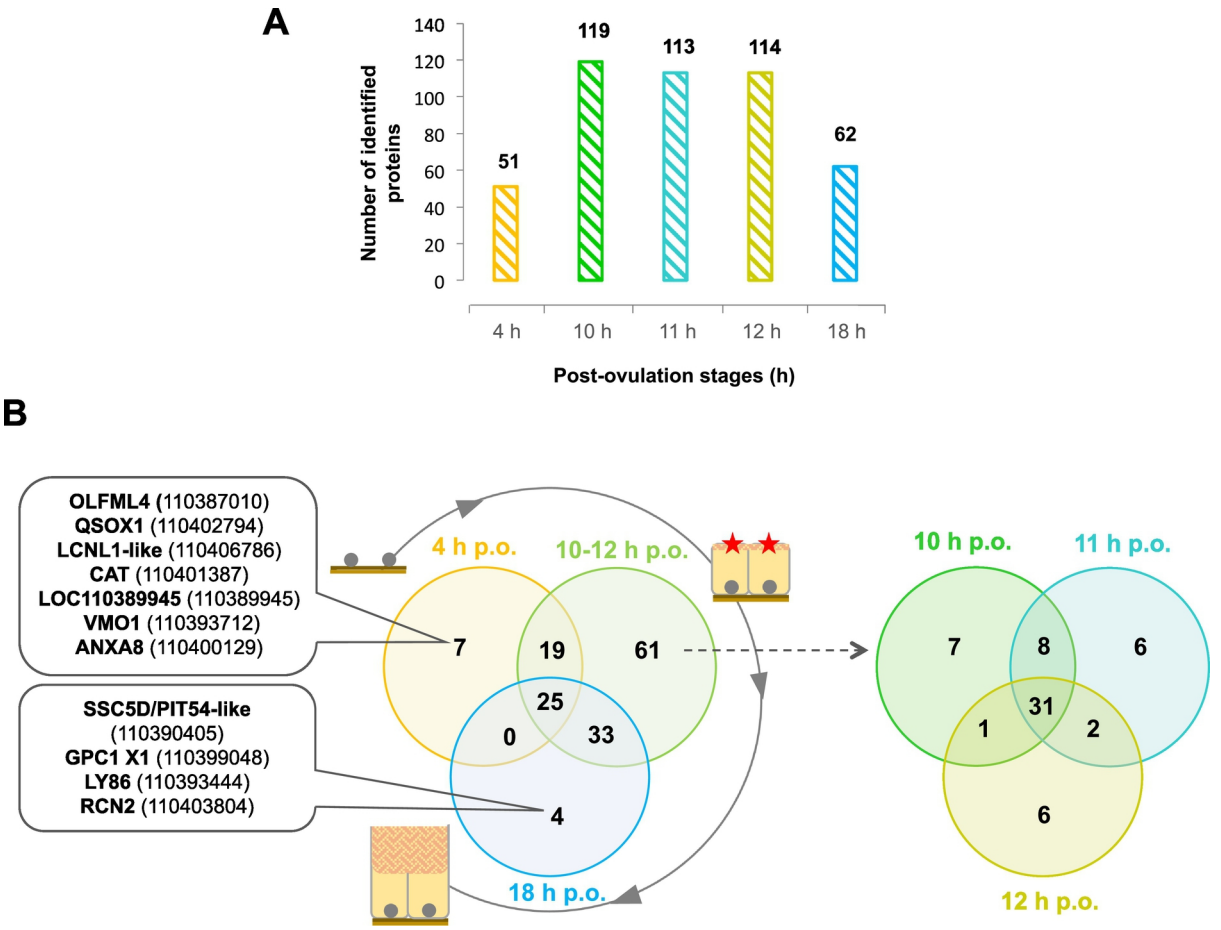

A
Number of identified proteins
140
120
100
80
60
40
20
0
51
119
113
114
62
4 h
10 h
11 h
12 h
18 h
Post-ovulation stages (h)
B
OLFML4 (110387010)
QSOX1 (110402794)
LCNL1-like (110406786)
CAT (110401387)
LOC110389945 (110389945)
VMO1 (110393712)
ANXA8 (110400129)
SSC5D/PIT54-like (110390405)
GPC1 X1 (110399048)
LY86 (110393444)
RCN2 (110403804)
4 h p.o.
10-12 h p.o.
18 h p.o.
7
19
61
25
0
33
4
10 h p.o.
11 h p.o.
12 h p.o.
7
8
6
31
1
2
6

Figure 3

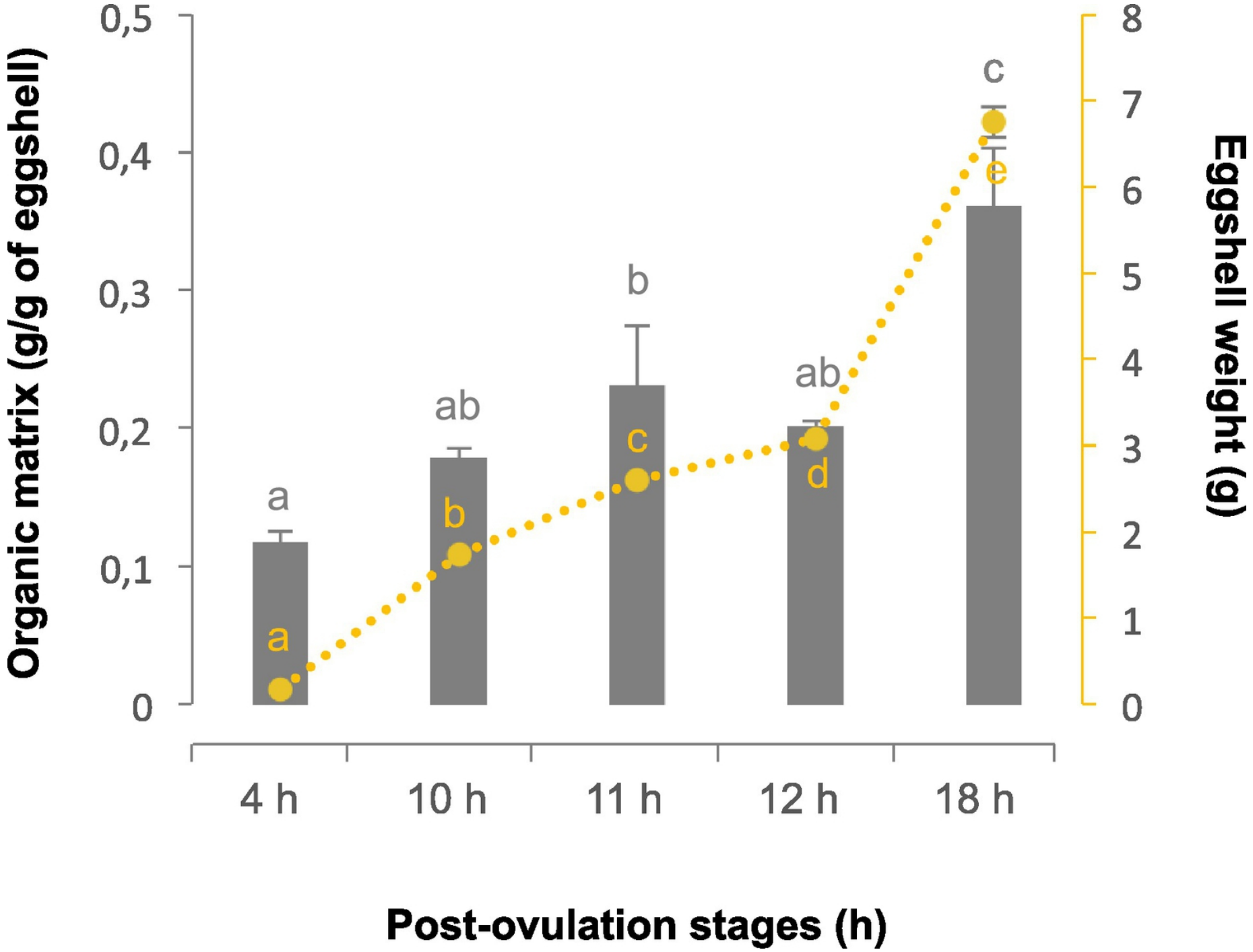
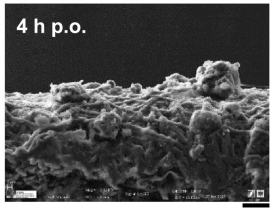
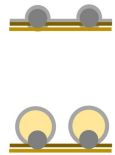

Organic matrix (g/g of eggshell)
0,5
0,4
0,3
0,2
0,1
0
Eggshell weight (g)
8
7
6
5
4
3
2
1
0
a
ab
b
ab
c
a
b
c
d
e
4 h
10 h
11 h
12 h
18 h
Post-ovulation stages (h)

Figure 4

First mineralization events





4 h p.o.

OVALX2, ExFAB, OVM, OVOT, HPX, LYZ C, DCA-1-like, OC-17-like, OVALY, HAPLN3, AVD-like, AGP-like, CST3, HBAA, CLU

Change in the size and orientation of crystals (shift)

10 h p.o.

DCA-1-like, ExFAB, OC-17-like, OVALX2, LYZ C, OVM, OC-116-like/MEPE-like, OVOT, OCX-32-like/RARRES1-like, AGP-like, EDIL3, HPX, AVD-like, CST3, CLU

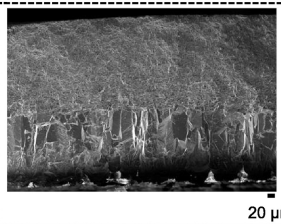
11 h p.o.

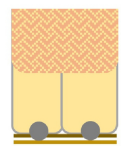
DCA-1-like, OC-17-like, OVALX2, OC-116-like/MEPE-like, OVM, ExFAB, LYZ C, OVOT, OCX-32-like/RARRES1-like, AVD-like, AGP-like, EDIL3, CLU, HPX, CST3

12 h p.o.

DCA-1-like, OC-17-like, ExFAB, OC-116-like/MEPE-like, OVM, LYZ C, OVALX2, OVOT, CLU, OCX-32-like/RARRES1-like, ALB, EDIL3, CST3, AGP-like, HPX

Massive deposition of new crystalline layer





18 h p.o.

DCA-1-like, OC-17-like, ExFAB, OC-116-like/MEPE-like, LYZ C , OVOT, Meleagrin-like, ALB, SBP-1-like, CST3, CLU, B2M, OVAL X2, OVM, AGP-like

NPNT ★
CALB1 ★
ANXA2 ★
S100-A6 ★
ANXA1 ★
LOC104050214
TSKU
SERPINF1
PTN
PTPRS
VTN
FSTL1

Figure 5