BIOLOGICAL CRYSTALLOGRAPHY

ISSN 1399-0047

Structure of struthiocalcin-1, an intramineral protein from *Struthio camelus* eggshell, in two crystal forms

Rayana R. Ruiz-Arellano,[a] Francisco J. Medrano,[b] Abel Moreno[a] and Antonio Romero[b]*

[a]Instituto de Química, Universidad Nacional Autónoma de México, 04510 Mexico City, Mexico, and [b]Biología Físico-Química, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain. *Correspondence e-mail: romero@cib.csic.es

Received 17 December 2014
Accepted 20 January 2015

Edited by R. McKenna, University of Florida, USA

Keywords: biomineralization; intramineral proteins; eggshell; struthiocalcin-1; C-type lectin fold.

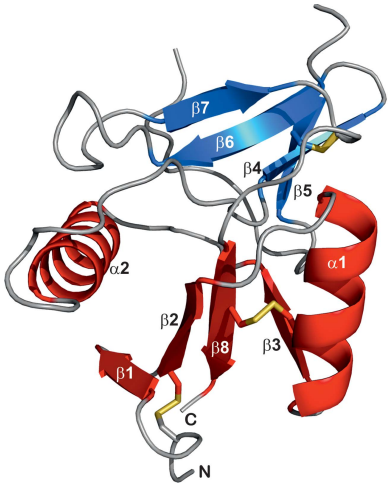
PDB references: SCA-1, 4uww; 4uxm

Supporting information: this article has supporting information at journals.iucr.org/d

Biomineralization is the process by which living organisms produce minerals. One remarkable example is the formation of eggshells in birds. Struthiocalcins present in the ostrich (*Struthio camellus*) eggshell matrix act as biosensors of calcite growth during eggshell formation. Here, the crystal structure of struthiocalcin-1 (SCA-1) is reported in two different crystal forms. The structure is a compact single domain with an α/β fold characteristic of the C-type lectin family. In contrast to the related avian ovocleidin OC17, the electrostatic potential on the molecular surface is dominated by an acidic patch. Scanning electron microscopy combined with Raman spectroscopy indicates that these intramineral proteins (SCA-1 and SCA-2) induce calcium carbonate precipitation, leading to the formation of a stable form of calcite in the mature eggshell. Finally, the implications of these two intramineral proteins SCA-1 and SCA-2 in the nucleation of calcite during the formation of eggshells in ratite birds are discussed.

1. Introduction

Much of the beauty of nature is reflected as much in the existence of gems (such as sapphires, diamonds, emeralds, rubies *etc.*) as in the attractive minerals such as calcite, quartz, pyrite *etc.* which have grown for millions of years deep under the Earth's surface. However, the most extraordinary minerals in terms of physico-chemical properties are those that form complex structures, namely those composed of minerals and biological macromolecules, usually called biominerals, present in living organisms (bones, teeth, stromatolites, eggshells, nacre shells *etc.*). Biominerals are formed by a combination of chemical, biochemical and biophysical processes. In general, the term biomineralization refers to the processes through which living organisms produce minerals under both physiological and pathological conditions (Mann, 1995; Kniep, 2010). Biominerals appear to be strongly dependent on their biological constituents, which regulate and control mineralization (Cölfen, 2010). Pathological mineralization leads to the formation of unwanted materials such as renal and bile stones, atherosclerotic plaques and dental calculus. There is therefore great interest in unravelling the mechanisms underlying the formation of both desirable and undesirable biominerals in living organisms. As stated by Cölfen, one of the most amazing characteristic of biominerals is their hierarchical architecture. Thus, the organic and inorganic components are arranged in building blocks over several length scales, ranging from nanostructures to perfect macroscopic structures with



particular shapes such as eggshells (in birds), horns (in animals), nails, teeth and bones (in mammals) (Cölfen, 2010).

The term biomineralization was coined at the beginning of the 1980s and provided a new source of inspiration for the production of tailored materials for a variety of biomedical and technological applications (Meldrum, 2003; Cölfen, 2010). Biomineralization processes are involved in the formation of bones and teeth, in optical/acoustic/magnetic sensors and even in pathological aspects such as cardiovascular calcifications of hydroxyapatite, renal diseases, pancreatic diseases and gallbladder stones. Among the nonpathological cases, one remarkable example is the formation of eggshells in birds. The eggshell is the only rigid part of an avian egg; it is made up of calcium carbonate and some biological macromolecules and has a highly specialized function. The question of the role of individual intramineral proteins in avian eggshell calcification is an important subject of investigation owing to the lack of information regarding the role of individual matrix proteins in avian eggshell calcification. Also, these types of proteins have never been studied as a biological sensor for calcium or carbonate (Reyes-Grajeda *et al.*, 2004). Lakshminarayanan and coworkers performed a comparative study of the structure–function relationships of the avian eggshell matrix proteins ansocalcin and ovocleidin-17 (OC17) (Lakshminarayanan *et al.*, 2005). These two proteins belong to the C-type lectin family and share a high degree of similarity, although they interact differently with the growing calcium carbonate crystals, suggesting different roles in eggshell calcification.

Mann and Siedler determined the amino-acid sequence of OC17, a major protein in hen (*Gallus gallus*) eggshell (Mann & Siedler, 1999). Later, the sequence of ansocalcin from goose (*Anser anser*) eggshell matrix was determined (Lakshminarayanan *et al.*, 2002). The sequences of struthiocalcin-1 and struthiocalcin-2 (SCA-1 and SCA-2) present in the ostrich (*Struthio camelus*) eggshell matrix have also been elucidated (Mann & Siedler, 2004), as have those of two proteins from emu (*Dromaius novaehollandiae*), dromaiocalcin-1 and dromaiocalcin-2, and two proteins from rhea (*Rhea americana*), rheacalcin-1 and rheacalcin-2 (Mann & Siedler, 2006). The first three-dimensional protein structure of an intramineral protein (ovocleidin, a protein from hen eggshell) was published in 2004 (Reyes-Grajeda *et al.*, 2004). Later, a comparison of different intramineral proteins from avian eggshells acting as biosensors for carbonate ions was published (Marín-García *et al.*, 2008). Although the preliminary X-ray characterization of the intramineral protein SCA-1 has been reported previously, its structure has not yet been determined (Reyes-Grajeda *et al.*, 2007).

Here, we report the purification, characterization, crystallization and three-dimensional structure of SCA-1. In addition, we carried out scanning electron microscopy to study the influence of SCA-1 and SCA-2 in the formation of spherical-shaped calcite crystals. The characterization of these calcium carbonate nanoparticles was performed by Raman spectroscopy using four different substrates. Finally, the role of these two intramineral proteins in eggshell formation in ratite birds is also discussed.

2. Materials and methods

2.1. Protein isolation and purification

The first part of the isolation and purification of SCA-1 and SCA-2 used a modification of the method developed by Mann & Siedler (1999). Ostrich eggshells were washed with water and then treated with 150 m*M* EDTA for 45 min to facilitate the extraction of the membrane and the outer layer. The clean eggshell was ground into a fine powder in a porcelain mortar and then extracted with 10%(*v*/*v*) acetic acid (20 ml per gram of eggshell) at 4°C overnight with constant stirring. After clearing the solution by centrifugation (40 000*g*), the supernatant was concentrated by ultrafiltration using Amicon Ultra centrifugal filter units (Millipore; 3 kDa molecular-weight cutoff). The solution was then dialyzed against 5 × 10 volumes of 5%(*v*/*v*) acetic acid. The dialyzed solution was then subjected to fractional precipitation with a 35–40% saturated solution of ammonium sulfate (Sigma–Aldrich). After 24 h, the solution was centrifuged at 64 500*g* for 30 min at 4°C, the supernatant was discarded and the pellet containing the protein fraction was resuspended in 5%(*v*/*v*) acetic acid and dialyzed in order to remove the remaining ammonium sulfate. Finally, the proteins were concentrated and injected onto a Superdex 75 gel-filtration column (GE Healthcare, Sweden) equilibrated with 50 m*M* sodium citrate buffer pH 4.0 containing 150 m*M* NaCl. The SCA proteins eluted as two different single peaks. The pooled peak fractions were collected and concentrated to 4 mg ml^{-1} using Amicon Ultra centrifugal filter units (Millipore). The concentration of both proteins was determined using the calculated extinction coefficient according to Gill & von Hippel (1989). The purified products were also analyzed by SDS–PAGE.

2.2. Crystallization

Initial crystallization screening was performed in 96-well sitting-drop plates (Swissci MRC, Molecular Dimensions, Suffolk, England) using a Cartesian Honeybee System (Genomic Solutions, Irvine, USA) and the JBScreen from Jena Bioscience (Jena, Germany), which comprises 240 unique reagent conditions. The nanodrops were of 0.4 µl in volume, consisting of 0.2 µl protein solution at 4.0 mg ml^{-1} and an equal volume of reservoir solution, and were equilibrated against 50 µl reservoir solution. Two sets of trials were performed in parallel and were placed at 277 and 295 K, respectively. Well formed diffraction-quality crystals were only obtained from SCA-1 at 277 K in three different crystal forms; no crystals appeared at room temperature. Crystals from the first condition (form 1) appeared with a plate-like morphology in 12%(*v*/*v*) MPD, 100 m*M* Tris–HCl pH 8.5, 50 m*M* $MgCl_2$. They belonged to space group $P2_12_12_1$, with unit-cell parameters a = 55.8, b = 68.8, c = 97.7 Å. These crystals showed significant pseudo-translational symmetry, which hampered complete structure solution and particularly the refinement process. The second crystal form (form 2) grew in 15%(*v*/*v*) MPD, 5% PEG 4K, 100 m*M* imidazole–HCl pH 8.0 and also belonged to space group $P2_12_12_1$, with unit-cell parameters a = 32.52, b = 55.67, c = 71.72 Å (Table 1). The third crystal

form (form 3) was obtained under similar crystallization conditions using MPD as the main precipitant but in the presence of glycerol as an additive [20%(*v*/*v*) MPD, 4%(*v*/*v*) glycerol, 200 m*M* NaCl]. In this case they grew in the orthorhombic space group $P2_12_12$, with unit-cell parameters $a = 55.81$, $b = 72.44$, $c = 37.05$ Å (Table 1). Crystals of adequate size appeared after 8–12 d and stopped growing within three weeks. A heavy-atom derivative was prepared by soaking crystal form 2 with 5 m*M* *p*-chloromercuribenzoate (PCMB) in reservoir solution for 24 h.

2.3. Data collection and processing

Crystals from the latter two conditions were gradually transferred to the respective reservoir solution containing higher concentrations of MPD in two successive steps. The final cryoprotectant solution contained 40%(*v*/*v*) MPD. For data collection, the crystals were flash-cooled in liquid nitrogen. Complete diffraction data sets were collected at 100 K using a PILATUS 6M detector on the BL13-XALOC beamline at the ALBA synchrotron, Barcelona (Spain) within a Long Term Proposal. Diffraction data were processed using *XDS* (Kabsch, 2010) and integrated and scaled using *SCALA* (Evans, 2011) within the *CCP*4 suite of programs (Winn *et al.*, 2011). A summary of the data-collection and processing statistics is given in Table 1.

2.4. Structure determination and refinement

The structure of SCA-1 was solved by single-wavelength anomalous diffraction using a single derivative of crystal form 2 at the Hg L_{III} edge peak wavelength ($\lambda = 1.0056$ Å), as revealed by the presence of a prominent peak in a previous X-ray absorption near-edge structure (XANES) fluorescence scan. The position of a single Hg site was determined with *SHELXD* (Sheldrick, 2010). No minor sites were found. Refinement of this site and subsequent phasing using *SHARP* (Bricogne *et al.*, 2003) yielded an electron-density map of excellent quality exhibiting clear solvent boundaries and suitable for model building. After density modification, a nearly complete model comprising 120 residues could be built using *ARP*/*wARP* (Langer *et al.*, 2008). This initial model was improved by successive cycles of refinement performed with *REFMAC* (Murshudov *et al.*, 2011) and alternate rounds of manual model building using *Coot* (Emsley *et al.*, 2010). Ordered solvent molecules were included during the final stages of the refinement and were visually inspected.

The structure of native SCA-1 in crystal form 3 was solved by molecular replacement with *MOLREP* (Vagin & Teplyakov, 2010) using the refined coordinates of crystal form 2 and a 1.54 Å resolution data set. The final refined model was obtained as described above. The missing residues could not be traced owing to a lack of interpretable electron density. Details of the model refinement are given in Table 1. Stereochemical validation of the final model was performed with *MolProbity* (Chen *et al.*, 2010). Electrostatic properties were studied using the *APBS* package (Unni *et al.*, 2011). Figures were prepared with *PyMOL* (DeLano, 2002). The final coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 4uww and 4uxm.

Table 1
X-ray data-collection and structure-refinement statistics.

Values in parentheses are for the outer shell.

	Native	Hg derivative
Data collection		
Space group	$P2_12_12$	$P2_12_12_1$
Unit-cell parameters (Å)	$a = 55.94$, $b = 72.60$, $c = 37.10$	$a = 32.67$, $b = 55.91$, $c = 72.06$
Mosaicity (°)	0.156	0.133
Wavelength (Å)	1.07153	1.00562
Resolution range (Å)	44.31–1.50 (1.58–1.50)	44.20–1.50 (1.58–1.50)
No. of reflections	313057 (42802)	384562 (54423)
No. of unique reflections	24808 (3546)	21799 (3076)
Wilson *B* factor (Å^2)	20.2	12.2
Completeness (%)	100 (100)	99.8 (99.8)
Multiplicity	12.6 (12.1)	15.4 (11.3)
$\langle I/\sigma(I)\rangle$	10.8 (1.6)	11.1 (3.9)
R_{merge}†	0.102 (1.885)	0.089 (0.747)
$R_{p.i.m.}$ (within I^+/I^-)‡	0.030 (0.559)	0.033 (0.258)
$CC_{1/2}$	0.99 (0.623)	0.99 (0.89)
SAD phasing		
Hg-atom sites		1
Anomalous correlation		2.45
Phasing power		1.86
Overall figure of merit		0.478
Anomalous multiplicity		8.7 (6.4)
Refinement statistics		
Resolution range (Å)	44.31–1.50	44.17–1.50
No. of reflections	23635	21798
R_{work}/R_{free}	0.213/0.232	0.188/0.213
No. of atoms		
Total	1153	1225
Protein	1079	1082
Water	74	132
Ligands	0	11
Average *B* factors (Å^2)		
Overall	28.6	18.40
Protein	28.0	17.25
Water	37.9	27.52
Ligands	—	21.55
R.m.s. deviations		
Bond lengths (Å)	0.021	0.017
Bond angles (°)	1.724	1.605
Ramachandran plot		
Favoured (%)	96.1	96.0
Allowed (%)	1.5	1.9
Outliers (%)	0.0	0.8
PDB code	4uww	4uxm

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the value of the *i*th intensity measurement and $\langle I(hkl)\rangle$ is the mean value of all measurements of $I(hkl)$. ‡ $R_{p.i.m.} = \sum_{hkl}\{1/[N(hkl)-1]\}^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

2.5. Calcite crystal-growth experiments by scanning electron microscopy (SEM)

In vitro crystallization assays were analyzed using scanning electron microscopy (SEM; VEGA3 SB, TESCAN, USA). Crystals were first grown on 18 mm glass plates and then covered with a gold layer to improve conductivity (Stiles *et al.*, 2008); scanning electron micrograph images were obtained using a 12–18 keV voltage.

2.6. Characterization of calcite crystals by Raman spectroscopy

Raman spectra were obtained using a DXR Raman microscope (Thermo Scientific) equipped with a diode-pumped, solid-state (DPSS) 532 nm laser. An 18 mm glass plate with the crystals was placed on a CaF_2 recipient on the holder of the microscope.

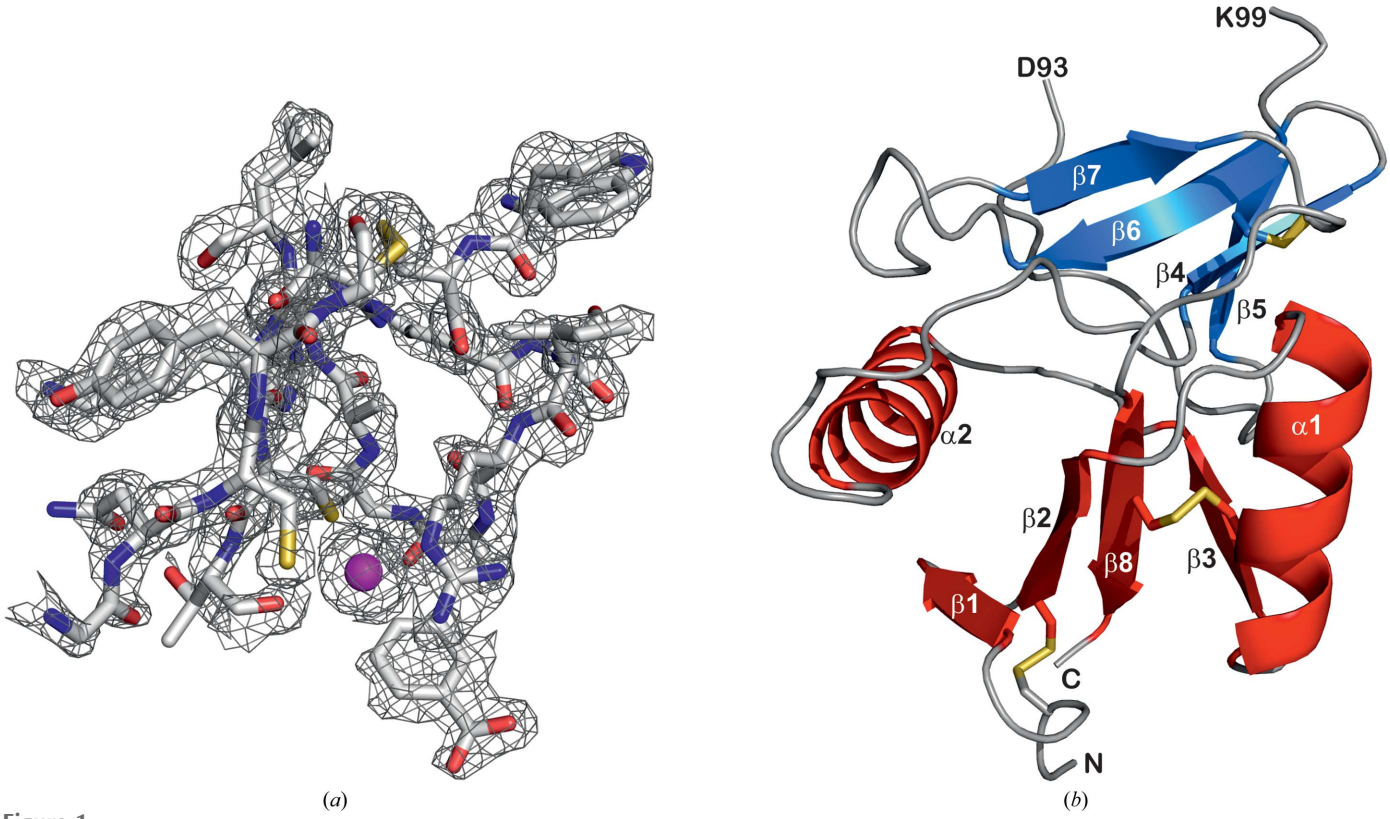
3. Results and discussion

3.1. Model building and overall structure

In previous work, orthorhombic crystals of SCA-1 were obtained in solutions containing MPD as a precipitant (Reyes-Grajeda *et al.*, 2007). These crystals share their space group and unit-cell parameters with those obtained in the present work corresponding to crystal form 1. However, the data showed a strong peak in the self-Patterson function at (0, 0, 0.45), suggesting a pseudo-translation of ∼1/2 along the *c* axis. This pseudo-translational symmetry hampered the complete structure solution and particularly the refinement process, with inappropriately high *R* factors (R_{work} = 34%; R_{free} = 46%).

In order to avoid model bias, the crystal structure of SCA-1 was determined by the single-wavelength anomalous (SAD) diffraction method using data to about 2.0 Å resolution from a mercury benzoate derivative obtained using crystal form 2 (see §2). The final model of SCA-1 includes residues Asp1–Asp93, Pro98–Ala132 and 88 solvent molecules (Table 1). The stereochemical quality of the SCA-1 model, with unambiguous electron density in the 1.5 Å resolution experimental map, was very good (Fig. 1*a*). Residues Asp94–Tyr97, which correspond to an acidic patch of the molecule in a flexible loop region, were not built owing to a lack of electron density.

The approximate dimensions of SCA-1 are 35 × 36 × 40 Å and structurally it is a compact domain of 132 residues involving an α/β fold (Fig. 1*b*) characteristic of the C-type lectin-like domain (Cummings & McEver, 2009; Gabius *et al.*, 2011). The molecule can be divided into two separate regions: a β subdomain and a mixed α/β subdomain containing the amino-terminal and carboxy-terminal residues. As shown in Fig. 1(*b*), the bottom half of the structure contains the two α-helices α1 (Trp24–Ile34) and α2 (Ser46–Tyr59), which are oriented nearly perpendicular to each other and surround a central β-sheet composed of four antiparallel β-strands: β1 (Trp7–Phe10), β2 (Asn13–Phe18), β8 (Phe126–Thr131) and β3 (Ala38–Leu40). The upper side of the molecule is made up by the second four-stranded antiparallel β-sheet, β7 (Trp115–Asp118), β6 (His102–Leu106), β4 (Leu72–Trp75) and β5 (Val78–Trp81), which covers the α/β subdomain. Three



(*a*) (*b*)

Figure 1
Overall structure of SCA-1. (*a*) Experimental electron-density map of the *p*-chloromercuribenzoate derivative at 1.5 Å resolution (contoured at 1σ) after solvent flattening and averaging calculated with *SHARP*. The benzoate moiety and protein residues are shown in stick mode. The Hg atom is shown as a pink sphere. (*b*) Cartoon representation of SCA-1. The structure involves an α/β fold characteristic of the C-type lectin-like domain, in which the two halves of the structure are coloured red (a mixed α/β region containing the N- and C-termini in the bottom part of the molecule) and cyan (a β-strand in the upper half) and the loops are in grey. The disulfide bonds are represented in ball-and-stick mode and are coloured yellow.

disulfide bonds, Cys3–Cys14, Cys31–Cys128 and Cys103–Cys120 (Fig. 1*b*), which are conserved in the long-form C-type lectin family (Cummings & McEver, 2009), help to stabilize the structure. As a result, strands $\beta1$ and $\beta2$ are connected through the Cys3–Cys14 disulfide bond, whereas Cys31–Cys128 joins helix $\alpha1$ and strand $\beta8$ both located in the lower lobe of the molecule. In the top half of the structure, Cys103–Cys120 connects strand $\beta4$ to the preceding $\beta8$ loop.

3.2. Structural comparisons

A comparison of the primary structure of SCA-1 with those of other intramineral proteins, as well as with that of the rat mannose-binding protein (rMBP), a representative member of the C-type lectin domain family (Cummings & McEver, 2009), is shown in Fig. 2. Despite the low sequence identity to other members of this family, sharing only 21% identity with rMBP and 32% with OC17, all of the common secondary-structural elements from the β and the mixed α/β subdomains are also found in SCA-1 (Fig. 3). This is not uncommon, as the C-type lectin fold has been found in a huge number of proteins with multiple functions in the cell that are not necessarily restricted to sugar binding (Cummings & McEver, 2009). In this regard, the C-type lectin fold is used by proteins to bind other entities (Zelensky & Gready, 2003) such as lipids, inorganic substrates (for example, $CaCO_3$ in the intramineral proteins) or even ice in the case of the antifreeze glycoproteins (Antson *et al.*, 2001). Hence, the C-type lectin fold can be regarded as an evolutionarily ancient structure with multipurpose uses in providing the scaffold for different substrate specificity.

There are, however, several structural differences in the connecting secondary-structure elements such as loops and turns. On examining the superposed structures of the intramineral proteins SCA-1 and OC17 with lithostathine, significant differences were found in the loop connecting the second α-helix ($\alpha2$) and strand $\beta4$ as well as in the loop connecting strands $\beta5$ and $\beta6$, where different lengths, conformations and charged surfaces are observed (Fig. 3). In particular, the $\alpha2$–$\beta4$

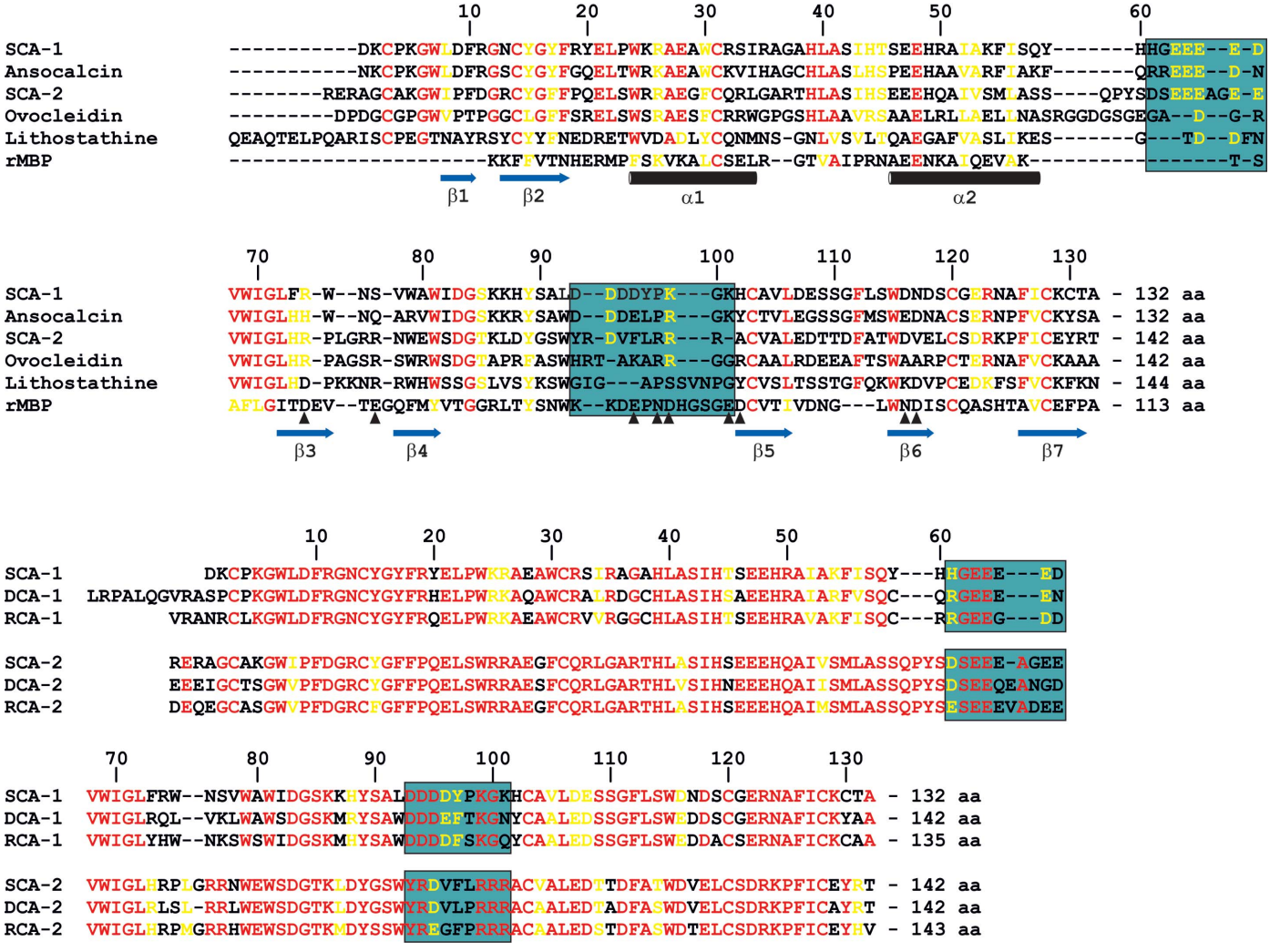


Figure 2
Sequence alignment of SCA-1 with the related proteins ovocleidin (Reyes-Grajeda *et al.*, 2004), lithostathine (Bertrand *et al.*, 1996) and rMBP (Ng *et al.*, 2002) performed using the *ClustalW* server (http://www.ebi.ac.uk/Tools/msa/clustalo/) and represented with *ESPript* (Robert & Gouet, 2014). Conserved residues are highlighted in yellow (>50% identity) or red (100% identity). Numbering and secondary-structural elements are based on SCA-1. The amino acids that are involved in Ca^{2+} binding in rMBP are indicated by black triangles.

loop, which is located at the interface between the two lobes of the structure, is notably shorter in SCA-1 than in the intramineral OC17 (Reyes-Grajeda *et al.*, 2004). This loop (Gln58–Glu65) is five residues shorter than in OC17 (Fig. 2), in which the five-residue insertion caps the bottom half of the α/β moiety (Fig. 3*a*). In fact, this loop has a similar length in lithostathine (Bertrand *et al.*, 1996) and rMBP (Ng *et al.*, 2002), furnishing an open cavity in this region that is occupied in rMBP by the N-terminal collagenous helix domain involved in oligomerization (Weis & Drickamer, 1994). Besides, this loop provides an important characteristic feature of SCA-1, the presence of an acidic patch made up of four Glu residues and one Asp residue (Glu63-Glu64-Glu65-Glu66-Asp67) clustering together at the base of the interface between the lobes of the structure (Fig. 3*b*) with potential implications for inorganic substrate binding in the biomineralization process (Mann, 2001).

Another relevant difference was located in the loop connecting strands $\beta5$ and $\beta6$ (Fig. 3*a*). This $\beta5$–$\beta6$ loop sits on top of the second β-sheet near the carbohydrate-recognition domain in rMBP, where the Ca^{2+} ions play an important role in the preservation of loop stability (Fig. 3*d*). The conformation of this loop, comprising residues Asp94–Pro98, was disordered in SCA-1 (Fig. 3*a*); it was therefore modelled to proceed with further discussion (Fig. 3*b*). Superposition of the structures showed that this loop can adopt different conformations

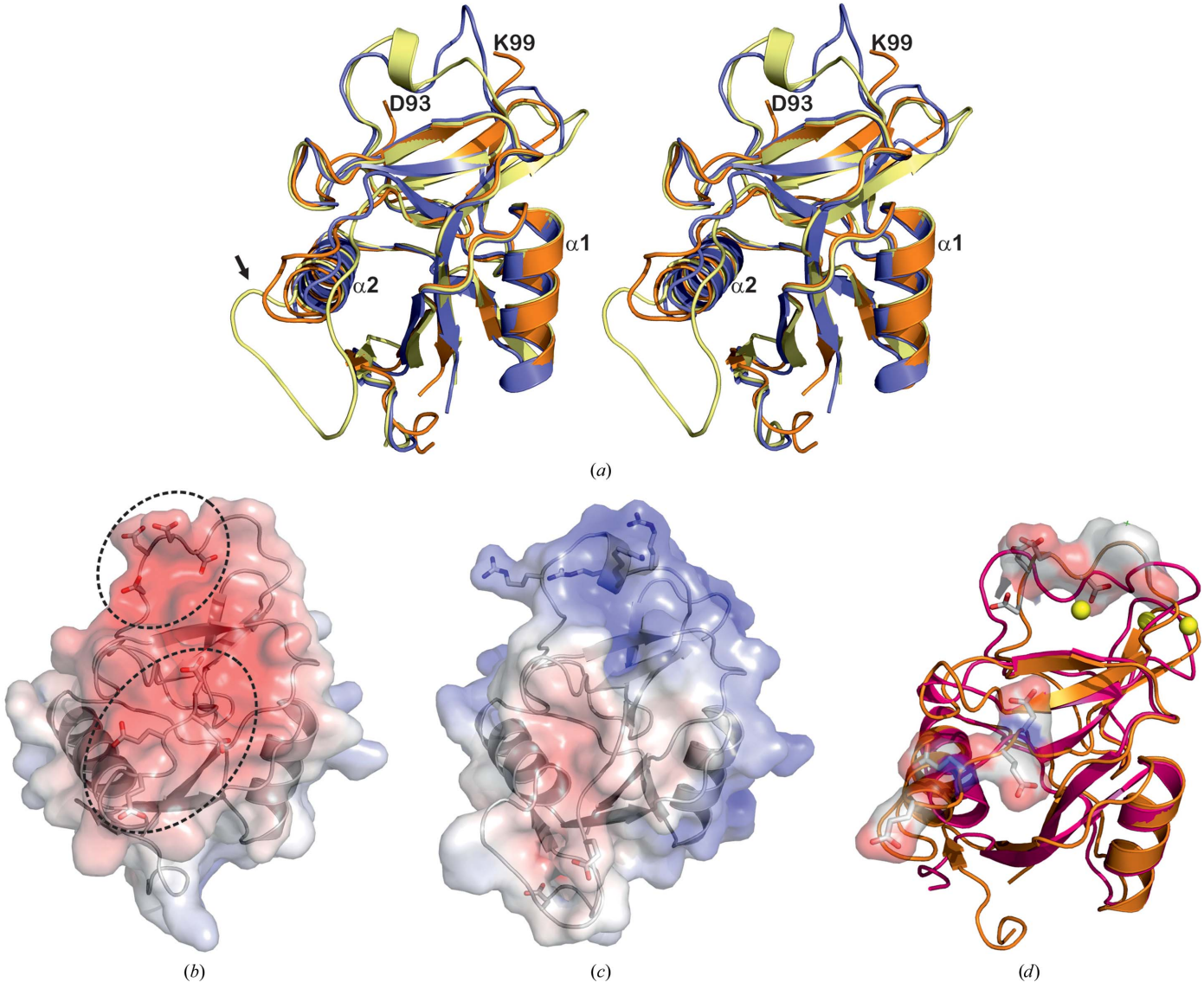


Figure 3
Structural comparison of SCA-1 with related proteins containing the C-type lectin-like domain. (*a*) Stereoview of the superposition of SCA-1 (orange), OC17 (yellow) and lithostathine (blue), showing the high structural similarity in the protein core. The main differences are found in distinct loop regions and turns, where the structural changes are markedly increased. (*b*, *c*) Ribbon diagrams in the same orientation of the C-type lectin domains from SCA-1 (*b*) and OC17 (*c*) with the electrostatic surface potential shown as a transparent surface with negatively charged residues depicted in red and positively charged residues in blue. The electrostatic surface potentials are contoured at -10 to $10kT\,e^{-1}$. (*d*) Superposition of SCA-1 (orange) and rMBP (pink; PDB entry 1kwu; Ng *et al.*, 2002), with the Ca^{2+} ions as yellow spheres. The two acidic regions of SCA-1 are highlighted by a transparent surface potential.

(Fig. 3*a*), but only in the case of OC17 (Reyes-Grajeda *et al.*, 2004) can a 3_{10}-helix be envisioned. It is worth mentioning that although this loop is conserved in length, its sequence variability (Fig. 2) leads to charged surfaces with very different potential distributions, highlighting the overall negative charge of SCA-1 in contrast to the highly positive charge of OC17 (Figs. 3*b* and 3*c*, upper lobe). It remains to be explained how these structural differences at the location of the charged residues would impact on the selection of the specific protein in the different growth phases of eggshell production (Lammie *et al.*, 2005).

3.3. Influence of SCA-1 on the crystal growth of calcium carbonate

Previously, we studied the effect of several intramineral proteins from ratite birds on the crystallization of calcium carbonate (Ruiz-Arellano & Moreno, 2014). It was observed that SCA-1 induced the formation of spherical-shaped crystals. Calcium carbonate crystals of this shape have also been observed in the formation of monodisperse biphasic vaterite–calcite microtablets (Tas, 2009). In these studies, the effect of the growth substrate on the crystal habit of calcium carbonate was not determined. Here, we have studied the growth of calcium carbonate crystals in the presence of SCA-1 on the surface of four different substrates (siliconized glass cover slides, hen eggshell inner membrane, muscovite and highly ordered pyrolytic graphite) by SEM, as shown in Fig. 4. Spherical-shaped crystals with a diameter ranging from 30 to 40 µm grew on siliconized glass cover slides (Fig. 4*a*). Small spherulites with a diameter ranging from 10 to 20 µm were obtained using the hen eggshell inner membrane as the substrate (Fig. 4*b*). Stack of layers of octahedral crystals of 10 µm in size were obtained on muscovite (a hydrophilic mica substrate) growing on the (100) direction (Fig. 4*c*). Rounded aggregates of 100 µm in size grew on the hydrophobic substrate HOPG (highly ordered pyrolytic graphite; Fig. 4*d*). These aggregates are very similar to those observed *in vivo* at the beginning of the mammillary cone layer, located between the internal shell membrane and the palisade layer of the eggshell. These results indicate that both the substrate and the protein influence the crystal habit. Thus, through nonspecific adsorption to the surface of the crystal, the protein blocks its growth in all directions, favouring the formation of rounded aggregates. This suggests that the formation of the aggregates occurs before the formation of the nuclei, which is followed by transformation of the crystals to a more stable form, as reported previously (Lakshminarayanan *et al.*, 2005). This also suggests that these intramineral proteins induce calcium carbonate precipitation leading to the formation of a stable form of calcite in the mature eggshell.

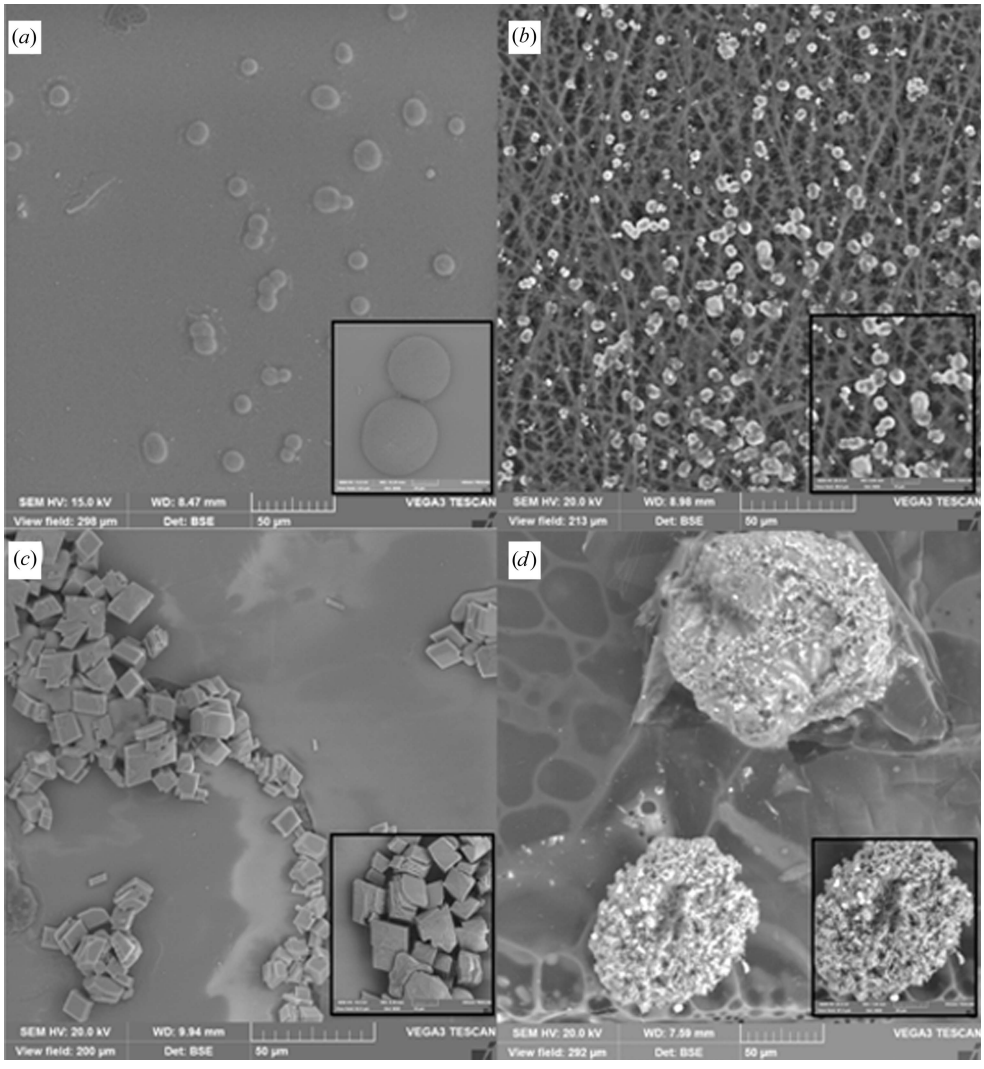


Figure 4
SEM micrographs of calcite crystals grown in the presence of SCA-1 (100 µg ml^{-1}) on four different substrates: (*a*) siliconized glass cover slides, (*b*) hen eggshell inner membrane, (*c*) muscovite and (*d*) highly ordered pyrolytic graphite (HOPG).

Fig. 5 shows SEM micrographs of the transverse section of hen eggshell (Figs. 5*a* and 5*b*) and ostrich eggshell (Figs. 5*c* and 5*d*). The rounded structures in the lower part of the mammillary cone layer (insets in Figs. 5*c* and 5*d*) resemble the spherical aggregates obtained *in vitro* when calcite crystals are grown on HOPG (Fig. 5*d*).

We used Raman spectroscopy to characterize the calcium carbonate crystals that grew in the presence of SCA-1. This technique was used owing to the small amounts of sample that could be obtained in crystallization experiments with calcium carbonate in the presence of the protein. Fig. 6 shows the characterization of these spherical nanoparticles and microparticles of calcium carbonate. Our results show that the aggregates obtained in the presence of SCA-1 corresponded only to the polymorph calcite. All of the main vibrations at 1435, 1085, 711, 280 and 159 cm^{-1} (related to CO_2^-, C═O and C—O—O groups) of a commercial sample of the polymorph calcite were observed in the samples obtained in the presence of the protein. This result indicates that calcite is the only polymorph of the three common anhydrous polymorphs of calcium carbonate (calcite, vaterite and aragonite) obtained in the presence of SCA-1.

All of these experiments showed that the intramineral proteins and the substrate should play an important role in the nucleation and growth of calcium carbonate crystals in the mammillary cone structures of the eggshell, leading to the formation of the mature egg.

3.4. Surface features and mechanistic implications

The electrostatic potential on the molecular surface of SCA-1, depicted in Fig. 7(*a*), reveals an acidic patch at the top of the molecule plus an extended interface between the two lobes. Negatively charged residues, namely Glu63, Glu64, Glu65, Glu66 and Asp67, located at the lobe interface, together with Asp93, Asp94, Asp95 and Asp96 at the upper side of the molecule, form a cluster of exposed acidic residues on the surface of SCA-1. In contrast to the ostrich intramineral SCA-1, the related avian ovocleidin OC17, the three-dimensional structure of which is known (Reyes-Grajeda *et al.*, 2004), displays a regular distribution of positively charged residues at the upper side region, including Arg103, Lys106, Arg108, Arg109 and Arg117 (Fig. 7*b*).

Interestingly, the ostrich calcified eggshell includes two different C-type lectin-like proteins, SCA-1 and SCA-2, as major components, as is also found in emu (*D. novaehollandiae*; dromaiocalcin-1 and dromaiocalcin-2) and in rhea (*R. americana*; rheacalcin-1 and rheacalcin-2) (Mann & Siedler, 2006), in contrast to chicken (OC17; Mann & Siedler, 1999) and goose (ansocalcin; Lakshminarayanan *et al.*, 2005) which have only one associated protein in the eggshell matrix. Although both proteins have been characterized (Mann & Siedler, 2004), the mechanisms of action and their role in eggshell calcification are still unknown.

It has been suggested that SCA-1 and ansocalcin are orthologues (Mann & Siedler, 2004); they share about 65% sequence identity, an identical length and the presence of a free unpaired cysteine (Fig. 2). Furthermore, the presence of two islands of conserved acidic amino acids in SCA-1 (Glu63-Glu64-Glu65-Glu66-Asp67 and Asp93-Asp94-Asp95-Asp96) and ansocalcin (Glu63-Glu64-Glu65-Asp66-Asn67 and Asp93-Asp94-Asp95-Glu96) reinforce the idea that the proteins are highly homologous. Unlike SCA-1, SCA-2 is more closely related to avian OC17 than to goose ansocalcin,

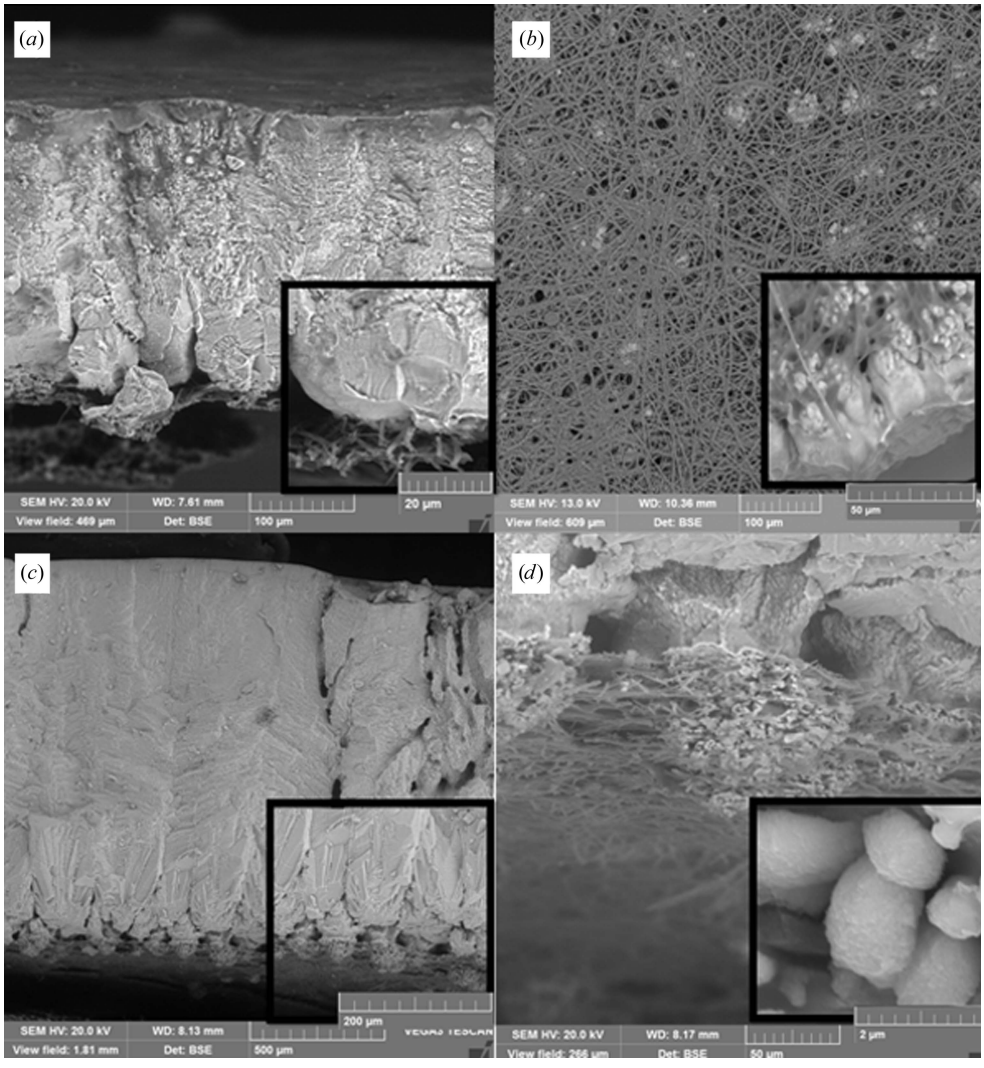


Figure 5
SEM micrographs of transverse cut imaging of two types of eggshell: (*a*) hen eggshell, (*b*) hen eggshell organic membrane, (*c*) ostrich eggshell and (*d*) the lower part of the organic membrane of the ostrich eggshell.

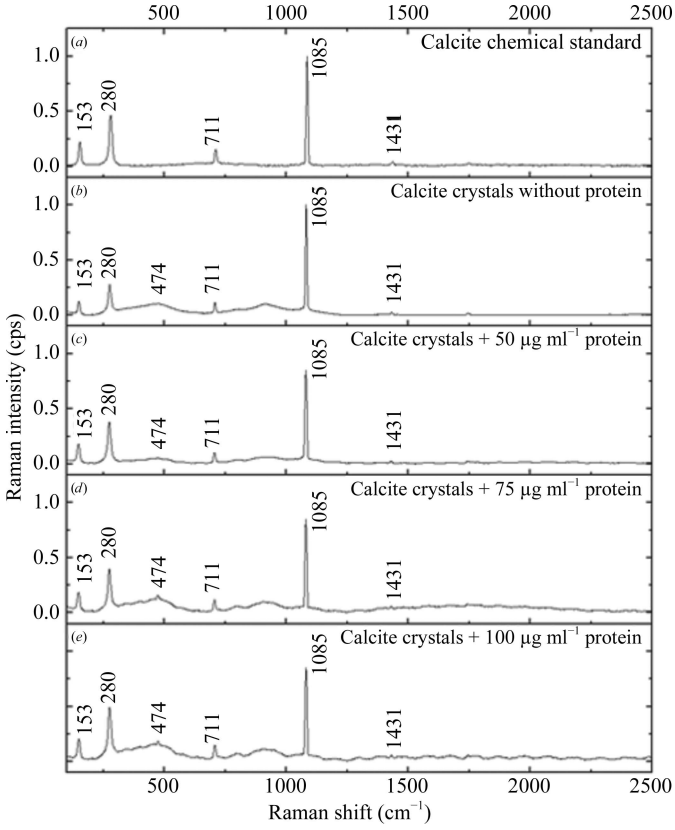


Figure 6
Characterization of spherical nanoparticles and microparticles of calcite by Raman spectroscopy. (*a*) Commercial sample of calcite, (*b*) calcite without SCA-1, (*c*) calcite grown with 50 µg ml^{-1} protein, (*d*) calcite grown with 75 µg ml^{-1} protein and (*e*) calcite grown with 100 µg ml^{-1} protein. All vibrations corresponded to the anhydrous polymorph calcite.

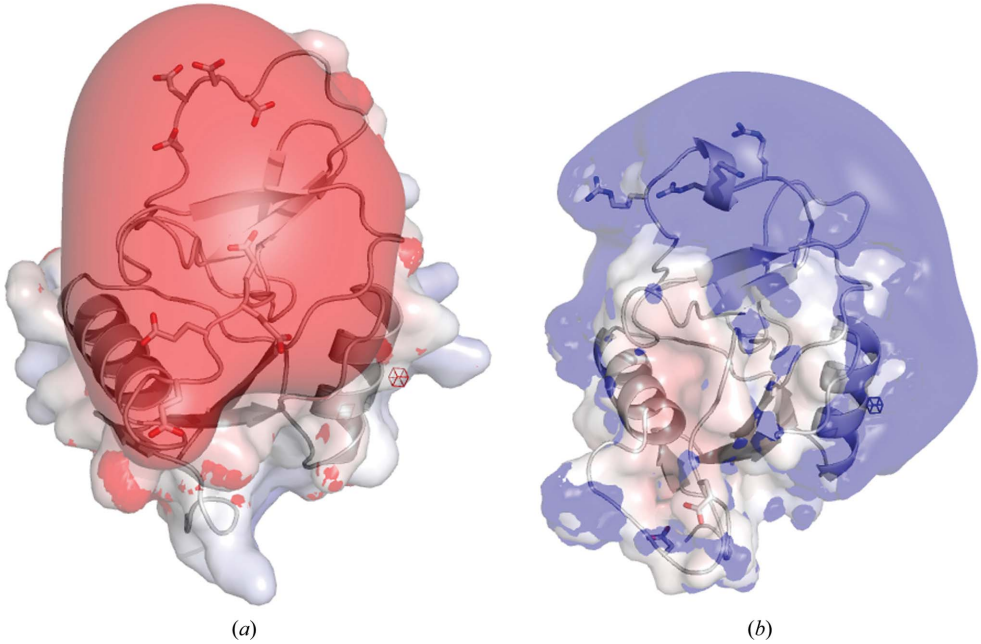


Figure 7
Electrostatic potential isosurface maps of (*a*) struthiocalcin-1 and (*b*) ovocleidin. Blue and red surfaces correspond to positive (+2kT e^{-1}) and negative (−2kT e^{-1}) potentials.

although it has common features with both. Firstly, two of the phosphorylated serines in SCA-2 are located at the same positions as in avian OC17. Secondly, SCA-2 has the three highly conserved disulfide bonds with no free cysteines, and the sequence length is identical to that of OC17 (Fig. 2). Thirdly, the charge distribution of SCA-2 must be highly polarized, with areas of negatively charged residues (Glu69-Glu70-Glu71-Ala-72-Gly73-Glu74-Glu75) present in both struthiocalcins, along with positively charged residues (Arg103, Arg108, Arg109 and Arg110) as in OC17.

There is very little knowledge about the biological function of the C-type lectin-like proteins present in eggshell. Current information shows the presence of two of those proteins in three orders of palaeognathous birds (Palaeognathae; Zhou *et al.*, 2014). This group is mainly made up of four types of large flightless birds (ostrich, emu, rhea and cassowary), one type of small flightless bird (kiwi) and one type of small bird capable of flight (tinamou). The information concerning this type of protein in the other group of birds (Neognathae), which is made up of most non-extinct birds, is very sparse; only 54 genomes are available at present. Using chicken ovocleidin (Reyes-Grajeda *et al.*, 2004) to search for similar proteins in this group showed its presence in three more species (*A. anser*, *Falco peregrinus* and *F. cherrug*). Interestingly, these birds showed only one protein present in the egg (*A. anser*) or the genome (*G. gallus*, *F. peregrinus* and *F. cherrug*). It is difficult to say anything accurately with this little information, but it is tempting to speculate that the presence of one or two proteins of this type is an evolutionary trait in birds. Thus, palaeognathous birds will have two proteins and neognathous birds will have one. Also, taking into account the presence of two types of small birds in this group, it could be said that the presence of two proteins might be necessary for the formation of large eggs, while small eggs require only one protein. It is also noteworthy to point out that of the 54 known genomes of neognathous birds, only three encoded a protein similar to ovocleidin. This raises the question: is there any other protein that fulfills the function of this C-type lectin-like protein? The answer might be that there is one protein of this type in the eggs of these birds which has very low sequence similarity or that there are other types of proteins that fulfill this function.

4. Conclusions

The three-dimensional structure of SCA-1 has been determined at 1.50 Å resolution and its overall structure is similar to that of the

C-type lectin-like fold. A remarkable difference is that unlike ovocleidin, the surface of SCA-1 is decorated with acidic residues that are concentrated at the top of the molecule plus an extended interface between the two lobes. The structural features of the struthiocalcins may have an impact on the mechanical properties of eggshell and explain the presence of two different C-type lectin-like proteins in palaeognathous birds (Palaeognathae), and our work places a new emphasis on the role of charged residues mediating crystal growth by controlling the size, shape and orientation of calcite.

Acknowledgements

This work was supported by grants BFU2011-24615 and CSD2009-00088 from the Spanish Ministry of Science and Innovation, the Regional Government of Madrid (S2010/BMD-2353) (to AR) and CONACYT project No. 175924 and PAPIIT Project No. IN201811 from the DGAPA-UNAM (to AM). RRR-A thanks CONACYT (registration No. 289778) for the scholarship for a fellowship into the PhD Program in Biomedical Sciences of Universidad Nacional Autónoma de Mexico. These experiments were performed on the PROXIMA1 beamline at SOLEIL, France and the XALOC beamline at ALBA, Spain. The authors cordially thank the staff members of the SOLEIL and ALBA synchrotrons for technical assistance.

References

Antson, A. A., Smith, D. J., Roper, D. I., Lewis, S., Caves, L. S., Verma, C. S., Buckley, S. L., Lillford, P. J. & Hubbard, R. E. (2001). *J. Mol. Biol.* **305**, 875–889.
Bertrand, J. A., Pignol, D., Bernard, J.-P., Verdier, J.-M., Dagorn, J.-C. & Fontecilla-Camps, J. C. (1996). *EMBO J.* **15**, 2678–2684.
Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. (2003). *Acta Cryst.* D**59**, 2023–2030.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Cölfen, H. (2010). *Nature Mater.* **9**, 960–961.
Cummings, R. D. & McEver, R. P. (2009). *Essentials of Glycobiology*, 2nd ed., edited by A. Varky, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart & M. E. Etzler, ch. 31. New York: Cold Spring Harbor Laboratory Press.
DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Evans, P. R. (2011). *Acta Cryst.* D**67**, 282–292.
Gabius, H.-J., André, S., Jiménez-Barbero, J., Romero, A. & Solís, D. (2011). *Trends Biochem. Sci.* **36**, 298–313.
Gill, S. C. & von Hippel, P. H. (1989). *Anal. Biochem.* **182**, 319–326.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kniep, R. (2010). *Nachr. Chem.* **58**, 419–423.
Lakshminarayanan, R., Joseph, J. S., Kini, R. M. & Valiyaveettil, S. (2005). *Biomacromolecules*, **6**, 741–751.
Lakshminarayanan, R., Kini, R. M. & Valiyaveettil, S. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 5155–5159.
Lammie, D., Bain, M. M. & Wess, T. J. (2005). *J. Synchrotron Rad.* **12**, 721–726.
Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nature Protoc.* **3**, 1171–1179.
Mann, S. (1995). *J. Mater. Chem.* **5**, 935–946.
Mann, S. (2001). *Biomineralization: Principles and Concepts in Bioinorganic Materials Chemistry*. New York: Oxford University Press.
Mann, K. & Siedler, F. (1999). *Biochem. Mol. Biol. Int.* **47**, 997–1007.
Mann, K. & Siedler, F. (2004). *Biochim. Biophys. Acta*, **1696**, 41–50.
Mann, K. & Siedler, F. (2006). *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **143**, 160–170.
Marín-García, L., Frontana-Uribe, B. A., Stojanoff, V., Serrano-Posada, H. J. & Moreno, A. (2008). *Cryst. Growth Des.* **8**, 1340–1345.
Meldrum, F. C. (2003). *Int. Mater. Rev.* **48**, 187–224.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Ng, K. K.-S., Kolatkar, A. R., Park-Snyder, S., Feinberg, H., Clark, D. A., Drickamer, K. & Weis, W. I. (2002). *J. Biol. Chem.* **277**, 16088–16095.
Reyes-Grajeda, J. P., Marín-García, L., Stojanoff, V. & Moreno, A. (2007). *Acta Cryst.* F**63**, 987–989.
Reyes-Grajeda, J. P., Moreno, A. & Romero, A. (2004). *J. Biol. Chem.* **279**, 40876–40881.
Robert, X. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320–W324.
Ruiz-Arellano, R. R. & Moreno, A. (2014). *Cryst. Growth Des.* **14**, 5137–5143.
Sheldrick, G. M. (2010). *Acta Cryst.* D**66**, 479–485.
Stiles, P. L., Dieringer, J. A., Shah, N. C. & Van Duyne, R. P. (2008). *Annu. Rev. Anal. Chem.* **1**, 601–626.
Tas, A. C. (2009). *Int. J. Appl. Ceram. Technol.* **6**, 53–59.
Unni, S., Huang, Y., Hanson, R. M., Tobias, M., Krishnan, S., Li, W. W., Nielsen, J. E. & Baker, N. A. (2011). *J. Comput. Chem.* **32**, 1488–1491.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Weis, W. I. & Drickamer, K. (1994). *Structure*, **2**, 1227–1240.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Zelensky, A. N. & Gready, J. E. (2003). *Proteins*, **52**, 466–477.
Zhou, Q., Zhang, J., Bachtrog, D., An, N., Huang, Q., Jarvis, E. D., Gilbert, M. T. P. & Zhang, G. (2014). *Science*, **346**, 1246338.