Chapter 6

Molecular Architecture of Helicoidal Proteinaceous Eggshells

Stavros J. Hamodrakas[1]

1 Introduction

1.1 Helicoidal Architecture: a Brief Description

Several extracellural fibrous structures are known to have helicoidal architecture. Such structures include arthropod cuticles, vertebrate tendons, plant cell walls etc. The widespread occurrence of the helicoidal structure in spherical shells, such as eggshells, spore walls, cyst walls, and others, and its correlation with the mechanical strength it provides is intriguing.

Excellent reviews on the helicoidal architecture and its appearance in biological systems have been made by Bouligand (1972, 1978a, b) and Neville (1975, 1981, 1986). These works describe in a beautiful and most comprehensive way how helicoids are identified, how widespread they are, the basic molecular principles of their formation and their geometrical, physical, and biological properties.

A brief description of the helicoidal architecture is as follows (Fig. 1).

The helicoidal architecture consists of helicoidally arranged (like the steps of a spiral staircase) parallel planes or sheets of fibrils (mono- or polymolecular). Within individual planes the fibrils are oriented more or less parallel to each other. Between successive planes the fibril direction rotates progressively (through a constant angle — the so called helicoidal twist), thus, giving rise to a helix, with its axis perpendicular to the planes, a sort of multi-directional "plywood" (Fig. 1a).

The helix can be either left- (anticlockwise) or right- (clockwise) handed; most biological systems so far examined have been shown to be left-handed (reviewed in Neville 1975).

The pitch of a helicoid is defined as the minimum distance corresponding to a 360° rotation of the direction of the fibrils. Therefore, the fibrils, are in parallel alignment every half pitch, which is equivalent to one lamella (Fig. 1a).

A helicoidal organization is identified ultrastructurally, usually under a transmission electron microscope, as a lamellar structure, exhibiting in, suitably cut, oblique thin sections, parabolic arrays of fibrils ("arced" patterns), which constitute each lamella (Fig. 1b).

[1] Department of Biochemistry, Cell and Molecular Biology, and Genetics, University of Athens, Panepistimiopolis, Athens 157.01, Greece

Results and Problems in Cell Differentiation 19
Biopolymers
Case, S.T. (Ed.)

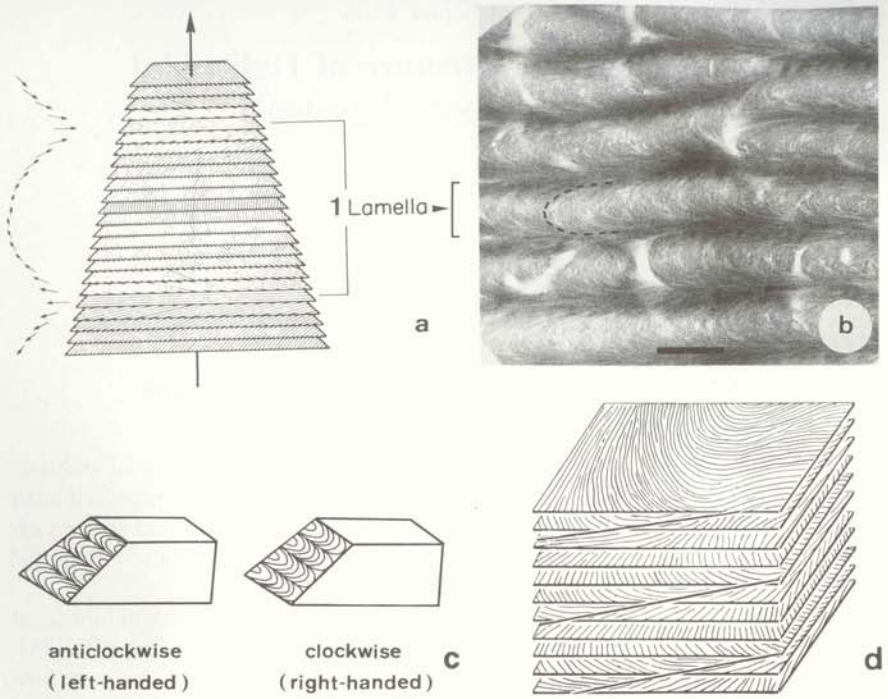


Fig. 1 a. Principle of helicoidal architecture. A helicoidal structure consists of parallel planes or sheets of fibrils. Within individual planes the fibrils are oriented, more or less parallel to each other. Between successive planes the fibril direction rotates progressively through a constant angle, the helicoidal twist, thus giving rise to a helix with its axis perpendicular to the planes. Fibrils are in parallel alignment every half-pitch of the helix (180° rotation of the fibril direction) which is equivalent to one lamella. In oblique thin sections typical parabolic arrays of fibrils ("arced" patterns) are seen to constitute each lamella. **b** Electron micrograph of an oblique section through a helicoidal proteinaceous eggshell of a silkmoth, *A. polyphemus*. Parabolic arrays of protein fibrils are seen in each lamella. *Bar* 0.4 μm. **c** The sense of rotation of a helicoidal structure can be deduced from the direction of the parabolic patterning on a known oblique section. An anti-clockwise (left-handed) helicoid is one in which successive planes of fibrils, twist anti-clockwise in a direction further away from the observer. **d** A generalized variant of the helicoidal model [see text; after Bouligand (1972)]

The sense of rotation (left or right) of a helicoid can be deduced from the direction of the parabolic pattern on a known oblique face (Fig. 1c).

A helicoidal structure should have as components rod-shaped, chiral (optically active) molecules (Bouligand 1972; Neville 1986).

A generalized variant of the helicoidal model may consist of a set of parallel surfaces of any shape, instead of planes, and not necessarily straight fibrils on each surface (Bouligand 1972; Fig. 1d). Two other ways in which helicoids depart from the ideal model are distortions and defects (Mazur et al. 1982 and references therein).

This description emphasizes the analogies of helicoidal fibrous biological structures with true cholesteric liquid crystals (Friedel 1922); the term "cholesteric" was introduced for the simple historical reason that it was first observed in some esters of cholesterol.

The close analogy between the helicoidal structures of (usually extracellular) biological materials and the structure of cholesteric liquid crystals probably suggests that several tissues and organelles are self-assembled according to a mechanism that is very similar to the process allowing materials to form liquid crystals. Apparently, helicoids should pass through a liquid crystalline phase before solidifying. Therefore, it is important to determine in such cases the molecular mechanisms of self-assembly.

Self-assembling systems are important in biology as they are economical in energy terms, requiring neither enzymatic control nor the expenditure of energy-rich bonds. They are particularly appropriate for building extracellular skeletal structures outside the cells which make them (Bouligand 1978b; Neville 1986).

Natural helicoidal composites occur in several combinations: polysaccharide fibers in a polysaccharide matrix (plant cell walls), polysaccharide fibers in a protein matrix, (arthropod cuticle), and protein fibers in a protein matrix (insect and fish eggshells), to mention just a few examples. In all cases principles of molecular recognition should govern the self-assembly mechanisms (Neville 1986).

1.2 Scope of This Chapter

Over the last decade we have been using the silkmoth chorion, the major component of the eggshell, a proteinaceous protective and functional layer surrounding the oocyte, as a model system to study how proteins fold and self-assemble to form complex, physiologically important structures. Silkmoth chorion is an example of a helicoidal composite of protein fibers in a protein matrix. In a previous communication (Hamodrakas 1984) by briefly surveying data collected from the silkmoth *Antheraea polyphemus* (Saturnidae) chorion, we proposed the twisted β-pleated sheet as the molecular conformation which dictates the formation of the helicoidal architecture in proteinaceous eggshells.

Since then several other helicoidal proteinaceous eggshells (both insect and fish) have also been studied, providing both novel information and interesting comparisons. These include eggshells from the domesticated silkmoth *Bombyx mori* (Bombycidae), *Antheraea pernyi* (Saturnidae), the lepidopteran *Manduca sexta* (tobacco hornworm moth, Sphingoidae), and *Sesamia nonagrioides* (an insect harmful to various crops, especially maize, Noctuidae) and the teleost fish *Salmo gairdneri* (trout).

This review will briefly cover the large amount of new information accumulated in the last 7 years which verifies, in the systems studied so far, the existence of a common molecular denominator, the β-pleated sheet, which apparently governs the self-assembly process in helicoidal proteinaceous eggshells, biological analogs of cholesteric liquid crystals.

1.3 Insect and Fish Eggshell: Terminology and Brief Description of the Systems

Hereafter, the terms insect eggshell and fish eggshell will simply denote the Lepidopteran eggshell and the *S. gairdneri* eggshell, respectively, to avoid possible misinterpretations.

1.3.1 Insect Eggshell

The insect eggshell (chorion and vitelline membrane) and the associated follicular epithelium, which secretes its constituent layers, have been the subject of numerous investigations at the cellular and molecular level, providing a model system in several areas of current biological research: physiology of the eggshell layers and morphogenesis of supramolecular structure, control of gene expression in differentiating cells, evolution of multigene families, and structural protein folding and organization.

Work in these areas has been reviewed recently (Kafatos et al. 1977; Hamodrakas 1984; Margaritis 1985; Regier and Kafatos 1985), and, together with the book by Hinton (1981), should be consulted for background information on the biology, biochemistry, physiology, and morphology of insect follicles and eggshells.

The ovary of an adult female lepidopteran usually consists of ovarioles, i.e., strings of follicles attached to each other in linear arrays from least mature to ovulated (Fig. 2a). Hereafter, we shall refer to a certain follicle by declaring its position within the ovariole with respect to the first follicle. With this terminology, the notation 5/22 means the 5th out of 22 choriogenic follicles in an ovariole.

The follicular cells, which surround the oocyte, synthesize and secrete, according to a precise spatial and temporal program, a set of structural proteins onto the surface of the oocyte, which self-assemble to form the multilayered eggshell (Fig. 2b). Protein synthesis occurs over a 2-day period in the silkmoths (Kafatos et al., 1977); in contrast, in *Drosophila melanogaster*, it lasts only 5 h (Margaritis 1985).

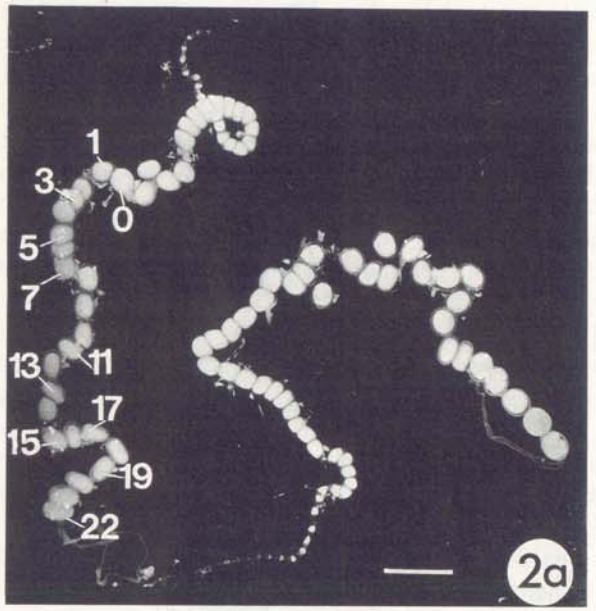
The eggshell performs certain functions, permitting sperm entry-fertilization, exchange of the respiratory gases, mechanical and thermal insulation, waterproofing, resistance to external high pressures, exclusion of microorganisms, and hatching (Kafatos et al. 1977; Margaritis 1985).

Chorion, which is largely proteinaceous (over 96% of its dry weight is protein in the silkmoth; Kafatos et al., 1977), accounts for at least 95% of the total eggshell mass in the species studied and it is the part of the eggshell which exhibits a helicoidal architecture.

Hereafter, for the insect eggshell, we shall be using both terms eggshell and chorion to refer to chorion.

1.3.2 Fish Eggshell

The tough eggshells of fish eggs play an essential role in controlling the relations between the external and internal egg environments allowing gaseous diffusion, providing physical protection and provision for sperm entry.



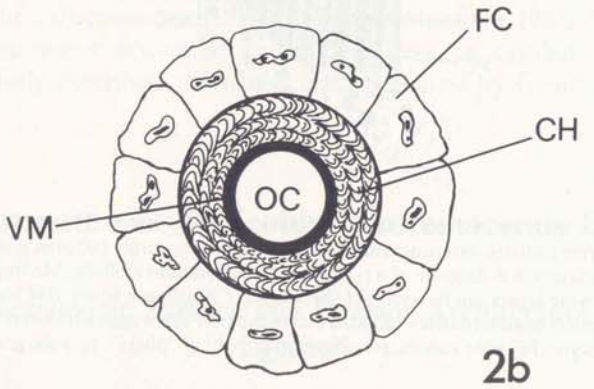


Fig. 2. a Two ovarioles dissected from a developing adult female *B. mori*. The paired ovaries in *B. mori* are each composed of four ovarioles. Follicles in progressively advanced stages of development (*0, 1, 2 . . . , 22*) are interconnected in a linear array within each ovariole; number 0 follicle indicates the beginning of choriogenesis. (Papanicolaou et al. 1986). **b** Schematic diagram of a follicle. *FC* follicle cells; *OC* oocyte. The eggshell consists of chorion (*CH*) and vitelline membrane (*VM*)

There is considerable variation in the nomenclature used to describe the eggshell of fishes (Grierson and Neville 1981; Groot and Alderdice 1985). Commonly used terms for this outer covering of the egg include: zona radiate, zona pellucida, radiate membrane, cortical membrane, vitelline envelope, egg envelope, egg capsule, cortex radiatus, chorion, eggshell.

Hereafter, for the fish eggshell, we have chosen to use the terms eggshell and/or alternatively, "cortex radiatus externus" when referring to its thin protein-polysac-

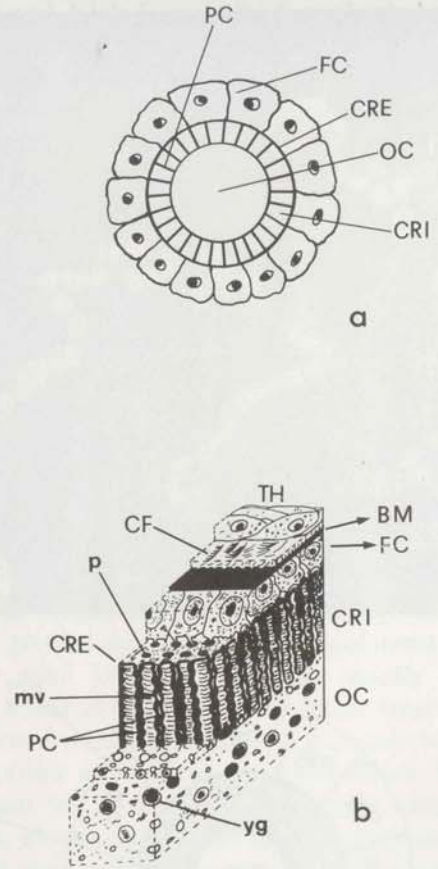


Fig. 3. a Schematic diagram of a fish follicle. *FC* follicle cells; *OC* oocyte. The eggshell consists of *CRE* (outer eggshell; cortex radiatus externus) and *CRI* (inner eggshell; cortex radiatus internus). Pore canals (*PC*) traverse these layers. **b** A diagram of a typical mature *S. gairdneri* follicle. Moving inwards from the periphery the following layers can be seen: *TH* theca layer; *CF* collagen fibers; *BM* basal membrane; *FC* follicle cells; *CRE* outer eggshell (cortex radiatus externus); *CRI* inner eggshell (cortex radiatus internus); *OC* oocyte (*PC* pore canals; *mv* oocyte microvilli; *p* "plugs"; *yg* yolk granules)

charide outermost layer and "cortex radiatus internus" when referring to its largely proteinaceous thick inner layer which has a helicoidal architecture. (Fig. 3a, b).

A wealth of interpretable information has been gathered from numerous studies on the fish eggshell, concerning, its fine structure and development, the structural changes after fertilization and water activation, its chemical composition as well as its degradation by hatching enzyme, its characteristics in relation to the prevention of polyspermy during fertilization and to the ecological significance of variation in eggshell structure. These studies are fully referenced by Groot and Alderdice (1985).

Initially, a few follicle cells are present close to the very young oocytes of fish ovaries (Anderson 1967; Flugel 1967). As the oocyte grows, the follicle cells increase in number, probably by mitosis to constitue a continuous follicular epithelium

which retains its single-layered structure throughout oocyte growth in teleosts. The theca layer is composed of collagenous fibers, fibroplasts, blood vessels and large thecal cells that surround the follicular epithelium outside the basal membrane (Hurley and Fischer 1966; Fig. 3a, b).

During oogenesis the oocyte develops microvilli on its surface. Synthesis of the eggshell begins on the oocyte surface at the base of these microvilli and proceeds inwards, towards the oocyte surface. The first layer to be deposited is the thin (0.15 μm) polysaccharide-protein cortex radiatus externus (CRE) followed by the thick (30 μm) proteinaceous, helicoidal cortex radiatus internus (CRI) layer. As the eggshell increases in thickness, the oocyte microvilli increase in length and become enclosed in the presumptive microvillar "pore canals". Follicular cell processes are also found in these pore canals, often in close association with the oocyte microvilli (Fig. 3b) throughout eggshell formation. (Papadopoulou P and Hamodrakas SJ, in prep.; Tesoriero 1977).

Our observations support the view that both oocyte and, perhaps to a lesser extent, the follicular cells contribute to the formation of the eggshell (see also Groot and Alderdice 1985 and references therein). Before ovulation the oocyte microvilli and folicular cell processes withdraw from the pore canals and the outermost (closest to the follicle cells) pore canal openings are blocked with "plugs" of material. However, remnants of the oocyte microvilli and follicular cells processes are, occasionally, found in "plugged" pore canals even after ovulation (our unpubl. observations; for a discussion see also Groot and Alderdice 1985). The important events following water activation and their effects on eggshell structure and chemistry are fully described, discussed, and referenced by Groot and Alderdice (1985).

2 Ultrastructure of Helicoidal Proteinaceous Eggshells

2.1 Macroscopic Surface and Interior Architecture of the Eggshell

Lepidopteran and fish eggs are usually large and extremely hard. Their shape is that of a laterally flattened ellipsoid, with a major diameter ranging from 0.9 mm for *S. nonagrioides*, 1.5 mm for *B. mori*, 2.1 mm for *M. sexta*, 3 mm for *A. polyphemus* and *A. pernyi* to 5.5 mm in *S. gairdneri* and exhibit anterior-posterior polarity (Fig. 4). The dry weight of the eggshell varies accordingly, being approximately 40 μg in *B. mori* and 460 μg in *A. polyphemus* (Kafatos et al. 1977).

In the lepidopteran species, a prominent feature of its surface is a polygonal network of ridges (Figs. 5a, b, 6a). The ridges correspond to the edges of the follicular cells which secrete chorion. They are formed by overproduction of chorionic proteins in the intercellular spaces. Each polygon corresponds to the overlying secretory cell — it is a follicular cell "imprint" — and each ridge to a two-cell junction.

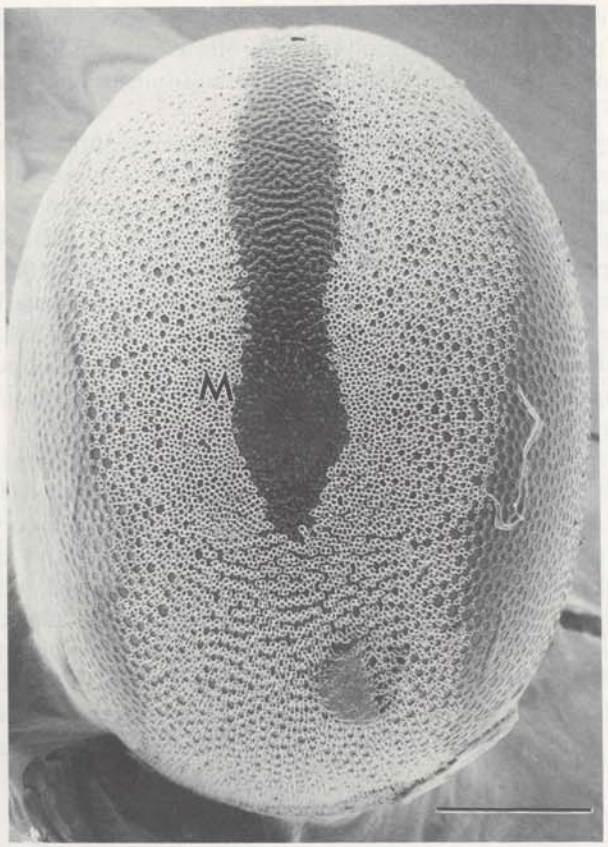


Fig. 4. Scanning electron micrograph of a mature (ovulated) follicle of *A. polyphemus*. At the anterior pole, the micropyle region (*M*) is seen. *Bar* 0.4 mm. (Courtesy of Dr. LH Margaritis)

Each polygonal imprint includes a small number of knobs which varies not only among species, but also among different strains of *Bombyx* (Kafatos et al. 1977 and references therein; Papanicolaou et al. 1986; our unpubl. data).

At ridge corners, aeropyles are found (Figs. 5b, c, 6a). These are round openings leading to internal radial air-channels; they are found normally at the borders of

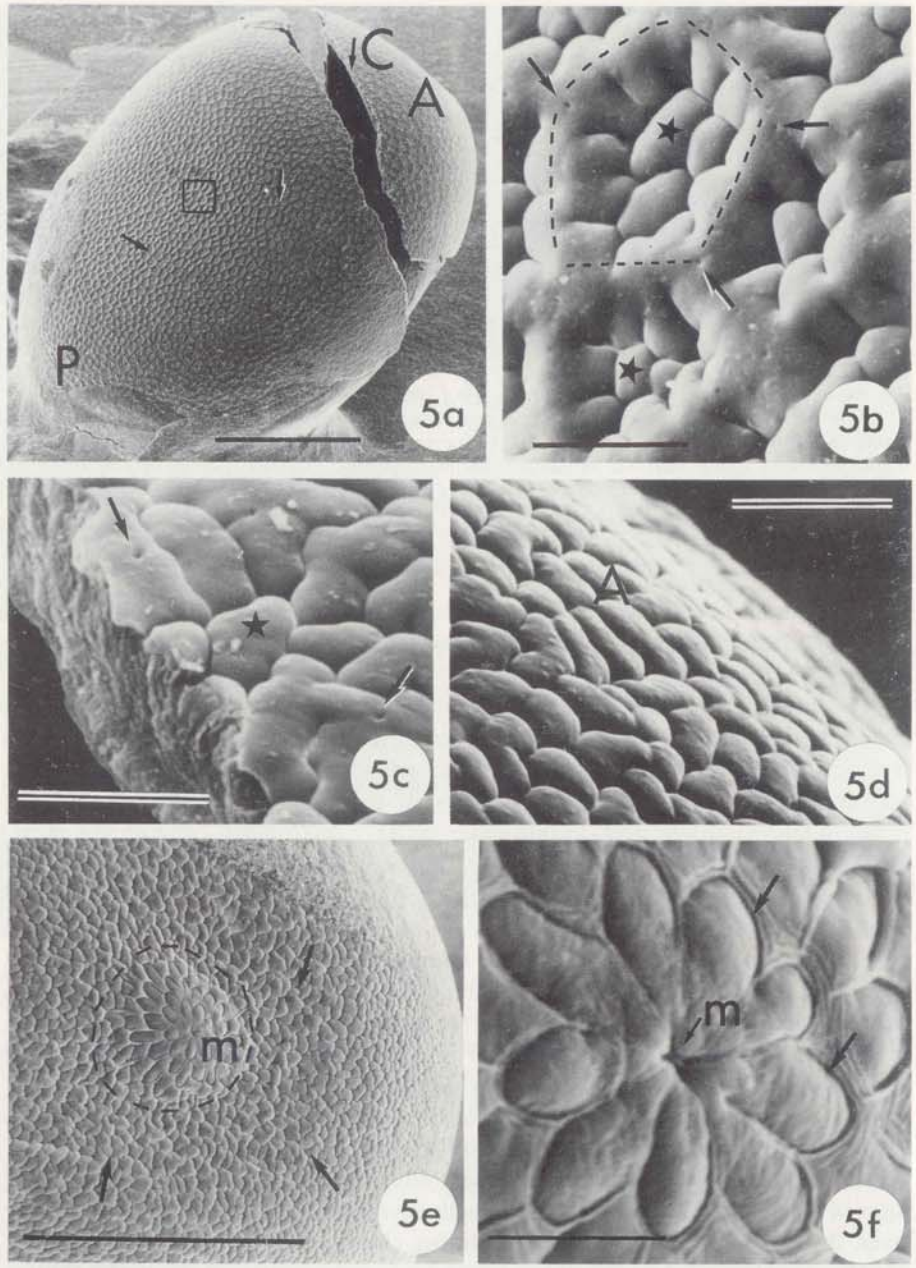


Fig. 5a–f. Scanning electron micrographs of a purified chorion from a mature *B. mori* follicle, 22/22. **a** There are two planes of symmetry perpendicular to each other, both passing through the longest meridian of the eggshell (one is parallel to the plane of the micrograph). The surface of the chorion shows polygonal imprints (*arrows*), each one created by one follicular (epithelial) cell. The anterior pole (*A*) is slightly flattened. *P* posterior pole. *C* (*arrow*) indicates a ripped region of the chorion. *Bar* 0.4 mm. **b** Closer view of the squared region of the chorion surface shown in **a**. The follicle cell imprints are marked by wide polygonal ridges (*dotted lines*) which correspond to intercellular regions of the follicle cells. Several knobs rest within each imprint. Small pores (aeropyles) are seen at the corners of polygons (*arrows*). *Bar* 20 μm. **c** Side view of the ripped region C of the chorion shown in **a**. The ridges of the imprints become flat and polygons disappear. Aeropyles (*arrows*) are seen on the knobs. *Bar* 20 μm. **d** Detail of the anterior pole (*A*) of the chorion. Polygonal imprints completely disappear (compare with **b**). The unfocused region is the chorion dome where the micropyle lies. *Bar* 20 μm. **e** Front view of the micropyle dome (*dotted lines*). The micropyle (*m*) is surrounded by four concentric "circles" of cell imprints. The remainder of the surface is covered by knobs (*arrows*). *Bar* 20 μm. **f** Detail of the micropyle (*m*, *arrow*) shown in **e** where the surrounding cell imprints (*arrows*) take the shape of a rosette. The ridges of the micropyle region are very thin (compare with **b**) and are devoid of aeropyles. The micropyle appears to contain four channels. *Bar* 20 μm. (Papanicolaou et al. 1986)

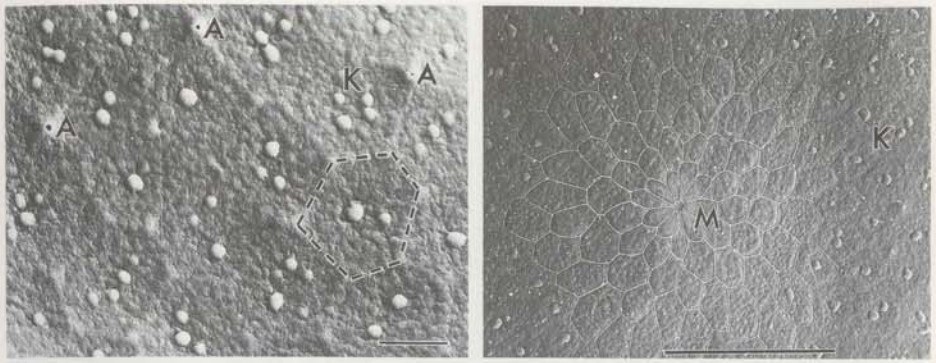


Fig. 6a, b. Scanning electron micrographs showing surface regions of a mature *M. sexta* chorion. **a** Follicle cell imprints are marked by wide polygonal ridges (*dotted lines*). Several "knobs" (*K*) are seen within each imprint. Aeropyles (*A*) are found at the corners of polygons. *Bar* 20 μm. **b** The micropyle (*M*) is surrounded by four concentric "circles" of cell imprints. It is discerned by the fine ridges and the elongated shape of the follicle cell imprints. *Bar* 100 μm

three-cell junctions. The aeropyles may be simple small holes or elaborate chimney-like structures, reflecting the respiratory physiology of the different species (Kafatos et al. 1977).

The anterior pole of chorion contains the micropyle, with the micropylar channels (their number is also species specific), through which sperm entry occurs (Figs. 5d, e, f, 6b). Its external opening is surrounded by a rosette of petal-shaped cell imprints. The ridges are lower than in the bulk of the chorion surface, substantially narrower, and devoid of aeropyles.

In the fish eggshell, the outer surface of the major layer CRI is decorated by numerous "plugs" (Fig. 7a) which block the outer openings of the radial "pore canals" traversing the eggshell (Fig. 7c). An almost hexagonal arrangement of the pore canals openings is revealed, after complete removal of the outer thin CRE layer with a treatment involving a combination of a denaturant (6 M urea) and a reducing agent (1% β-mercaptoethanol) (Fig. 7b).

The pore canal openings retain their regular arrangement in the inner (closest to the oocyte) surface of the eggshell, which shows a fibrous texture (Fig. 7d).

The outer opening of the micropyle, through which sperm entry occurs, is surrounded by a zone devoid of plugs and an outer zone with large plugs (Fig. 7f).

In vertically (to the surface) ripped sections of a lepidopteran eggshell, the first signs of a helicoidal architecture are lamellae lying above a usually thin trabecular layer, the first chorionic layer formed during choriogenesis (Fig. 8a–d).

The number of lamellae varies among species, ranging from 10–20 for *S. nonagrioides* to more than 100 in the moth *Hyalophora cecropia* (Smith et al. 1971). Their orientation relative to the oocyte surface is also variable and, similarly, their thickness and density in different regions of the eggshell. These variations presumably reflect inter-species-specific physiological needs, and/or perhaps different morphogenetic modes and local chemistry.

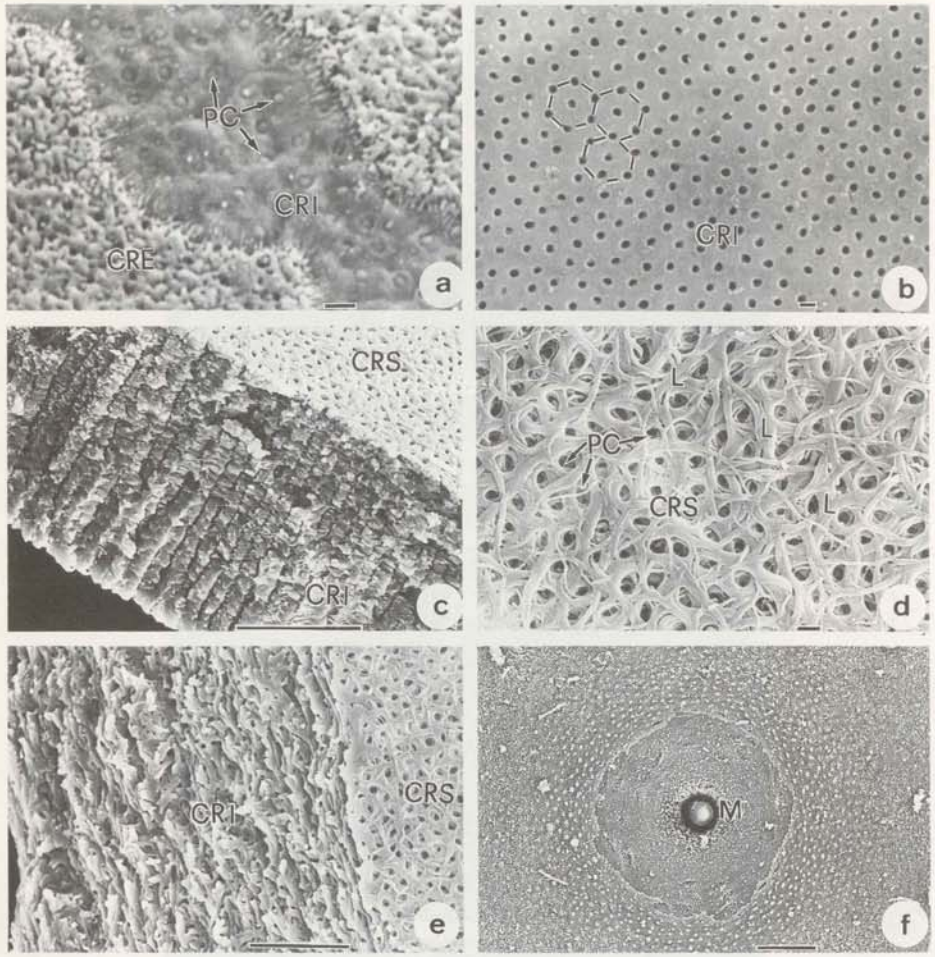
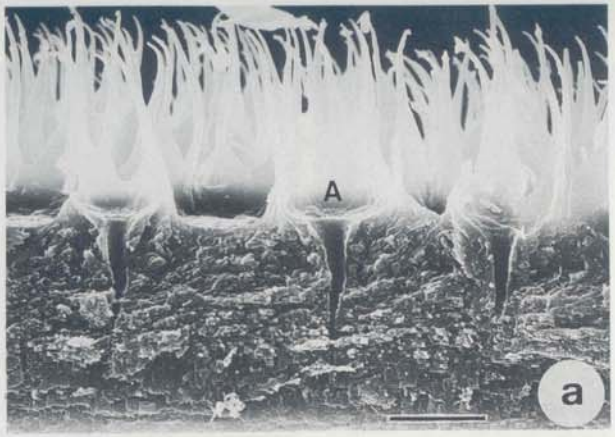
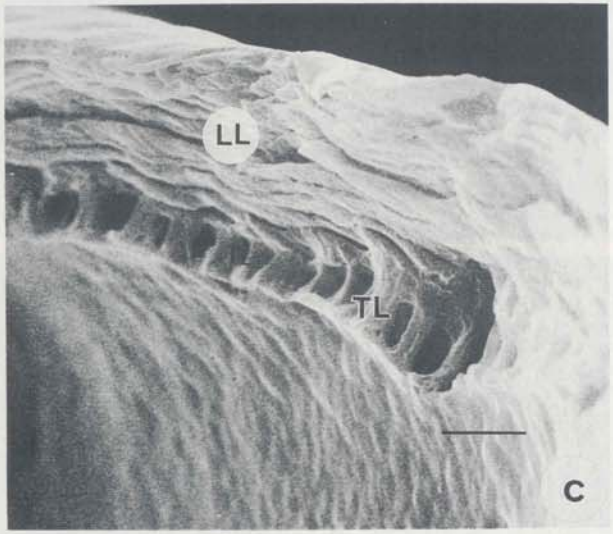
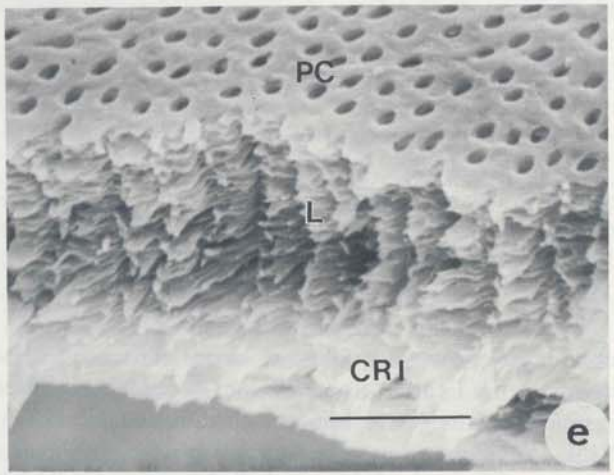


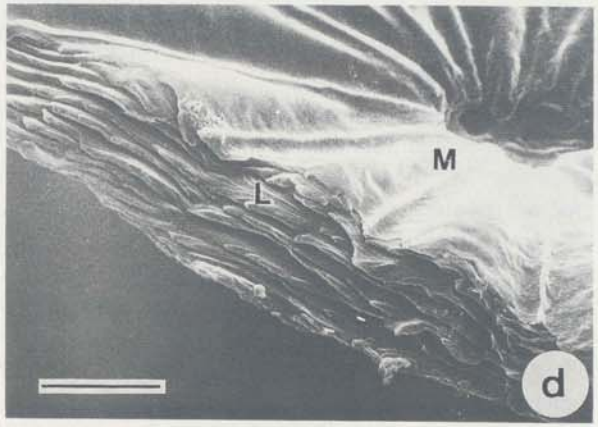
Fig. 7a–f. Scanning electron micrographs showing surface regions and interior architecture of the *S. gairdneri* eggshell. **a** CRE and the outer CRI with "plugs" (*P*) blocking the pore canals are seen (2 days after fertilization). *Bar* 1 μm. **b** Treatment with 6 M urea removes CRE and the plugs, revealing an almost hexagonal arrangement of the pore canals (2 days after fertilization). *Bar* 1 μm. **c** A vertical cross-sectional rip through the eggshell reveals the lamellar organization of CRI. The pore canals are easily discerned traversing the eggshell (18 days after fertilization). *Bar* 10 μm. **d** The inner surface (*CRS*) of the eggshell is seen. Note the network of lamellae, their interconnections and twist around the pore canals (*PC*) (18 days after fertilization). *Bar* 1 μM. **e** An oblique cross-sectional rip through an eggshell showing a different view of the lamellar organization (18 days after fertilization). *Bar* 10 μm. **f** The micropyle, 5.4 μm in diameter, is seen (2 days after fertilization). *Bar* 10 μm

The trabecular layer consists of "columns" oriented perpendicular to the oocyte surface connecting a "roof" and a "floor" (Fig. 8). It contains cavities acting as air-tanks, which are important for respiration (Kafatos et al. 1977; Hinton 1981; Margaritis 1985). Its thickness is also species-specific. It is usually very thin compared to the lamellar part of the eggshell; however, in extreme cases, as in *S. nonagrioides*, it represents almost one third of mature chorion overall thickness









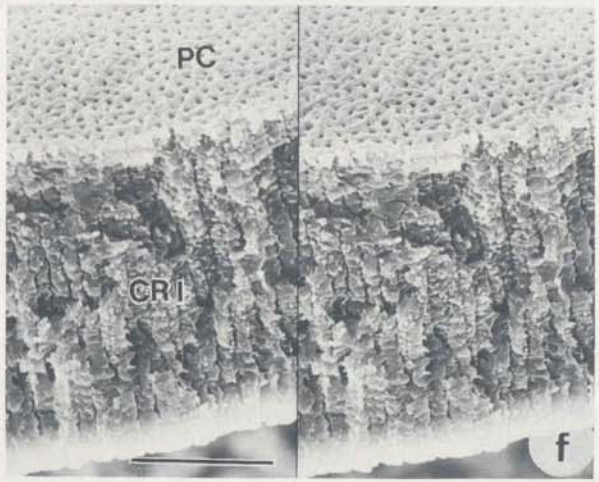


Fig. 8a–f. Scanning electron micrographs of cross-sectional rips through the mature eggshell showing the interior lamellar architecture of: **a** *A. polyphemus*; chimney-like tall aeropyles (*A*) are seen. *Bar* 20 μm. (Courtesy of Dr. LH Margaritis). **b** *M. sexta*; nearest to the oocyte is the trabecular layer (*TL*) which consists of pillars surrounding air-filled spaces. In the inner lamellar layer (*IL*) lamellae are lying parallel to the oocyte surface, whereas in the outer lamellar layer (*OL*) at an oblique angle to the rest. *Bar* 2.5 μm. **c**, **d** *S. nonagrioides*; the trabecular layer (*TL*) is almost 1/3 of total chorion thickness. The micropyle region is devoid of trabecular layer. In **c** *bar* 2.5 μm and in **d** 5 μm, **e**, **f** *S. gairdneri*; **e** shows an oblique cross-sectional rip, whereas **f** is a stereo-scanning electron micrograph showing the pore canals (*PC*) traversing regularly arranged lamellae. In **e** *bar* 5 μm and in **f** 10 μm

(Fig. 8c). The lamellar structure of fish eggshell can also be seen in vertically to its surface ripped sections (Fig. 8e, f).

2.2 Fine Structure and Morphogenesis of the Eggshell

2.2.1 Eggshell Fine Structure

The eggshells of Lepidoptera are clearly distinguished from those in other orders of insects by their predominantly lamellar ultrastructure (e.g., Fig. 9a, b, c; Smith et al. 1971; Furneaux and Mackay 1972; Kafatos et al. 1977; Hinton 1981; Mazur et al. 1982; Papanicolaou et al. 1986; Fehrenbach et al. 1987; Regier and Vlahos 1988).

A lamellar supramolecular structure also predominates in several fish eggshells (e.g., Fig. 9d; Grierson and Neville 1981 and references therein). Distortions and defects disrupt the regular lamellar arrangement (Mazur et al. 1982) both in lepidopteran and fish eggshells.

In Lepidoptera, under higher magnification, each lamella is seen to consist of fibers, 70–200 Å in diameter, embedded in an amorphous matrix (Fig. 10). The paraboloidal appearance of these fibers in oblique sections clearly suggests a helicoidal architecture (Bouligand 1972). Initially, these fibers are distributed sparsely, but as choriogenesis progresses they thicken and/or increase in number, until they coalesce (Smith et al. 1971; Kafatos et al. 1977; Regier and Vlahos 1988).

In perpendicular sections the paraboloidal "arcs" disappear, leaving alternating bands (zones) of lines and dots. The zones where the fibers are oriented nearly parallel to the plane of section appear darker, giving rise to the lamellar appearance. However, a closer examination of electron micrographs taken from eggshells in the late choriogenetic stages reveals parabolic arrays of thin, 30–40 Å, fibrils confirming the helicoidal model of architecture even at these developmental stages (Hamodrakas 1984; Papanicolaou et al. 1986). Our recent unpublished transmission electron microscopy and optical diffraction data (Hamodrakas SJ and Ottensmeyer FP, in prep.) suggest the existence of 30–40 Å fibrils as constituents of the 70–200 Å thicker fibers throughout choriogenesis.

Applying carefully the technique described by Neville (1975, p. 228), we were able to determine the sense of rotation of the helicoidal structure in lepidopteran eggshells (data not shown): clearly, proteinaceous eggshells in Lepidoptera are left-handed (anti-clockwise) helicoids, which means that successive planes of fibrils twist anticlockwise in a direction further away from the observer. As mentioned earlier, most helicoidal biostructures are left-handed (Neville 1975; Grierson and Neville 1981). Knowledge of the absolute sense of rotation of these helicoids might be crucial in determining the molecular mechanisms of self-assembly (packing of protein fibrils, see below).

The two-phase system (fibers embedded in a matrix) of lepidopteran proteinaceous eggshells appears to be an ideal solution to the problem of constructing eggshells providing strong and elastic mechanical protection and support (Kafatos et al. 1977 and references therein). Our data, however, pose the question whether

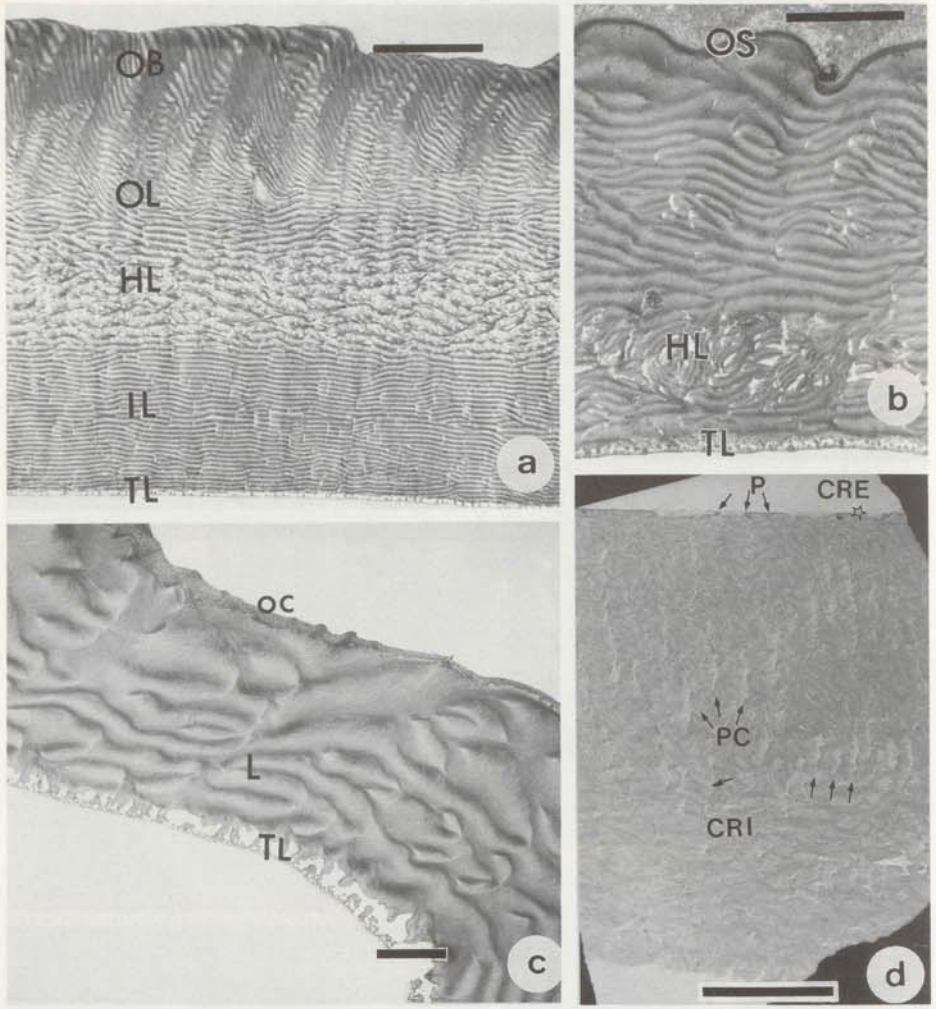


Fig. 9a–d. Transmission electron micrographs of thin sections cut through chorions of: **a** *A. polyphemus*; within the bulk of the lamellar chorion, four types of lamellae can be distinguished: thin lamellae of the inner lamellar layer (*IL*), thick, distorted, spongy lamellae of the holey layer (*HL*), lamellae of the outer layer (*OL*) and, lying at an angle to the rest, the thick lamellae of the oblique layer (*OB*). Nearest to the oocyte is the trabecular (*TL*) layer, which consists of pillars surrounding air-filled spaces. *Bar* 10 μm. (Hamodrakas et al. 1986). **b** *B. mori*; formation of the outer osmiophilic layer (*OS*) containing proteins with a high-cysteine content has just begun; the trabecular layer (*TL*) and the remainder of the chorion have already been laid down. *Bar* 5 μm. (Papanicolaou et al. 1986). **c** *S. nonagrioides*; a "sticky" outer covering (*oc*) has been deposited on the outer surface of a laid egg. It serves as a "glue" sticking together laid eggs. The main bulk of the chorion consists of lamellae (*L*) deposited on a thick trabecular layer (*TL*). *Bar* 2 μm. **d** *S. gairdneri*; the radial pore canals (*PC*) can be discerned (*arrows*) in the lamellar CRI, with a characteristic twisted ribbon structure (Grierson and Neville 1981), implying helicoidal architecture for the eggshell. On the outer surface of the eggshell, the thin osmiophilic *CRE* is the eggshell layer deposited first. "Plugs" can be seen (*P*, *arrows*) blocking pore canals. *Bar* 7 μm. (Hamodrakas et al. 1987)

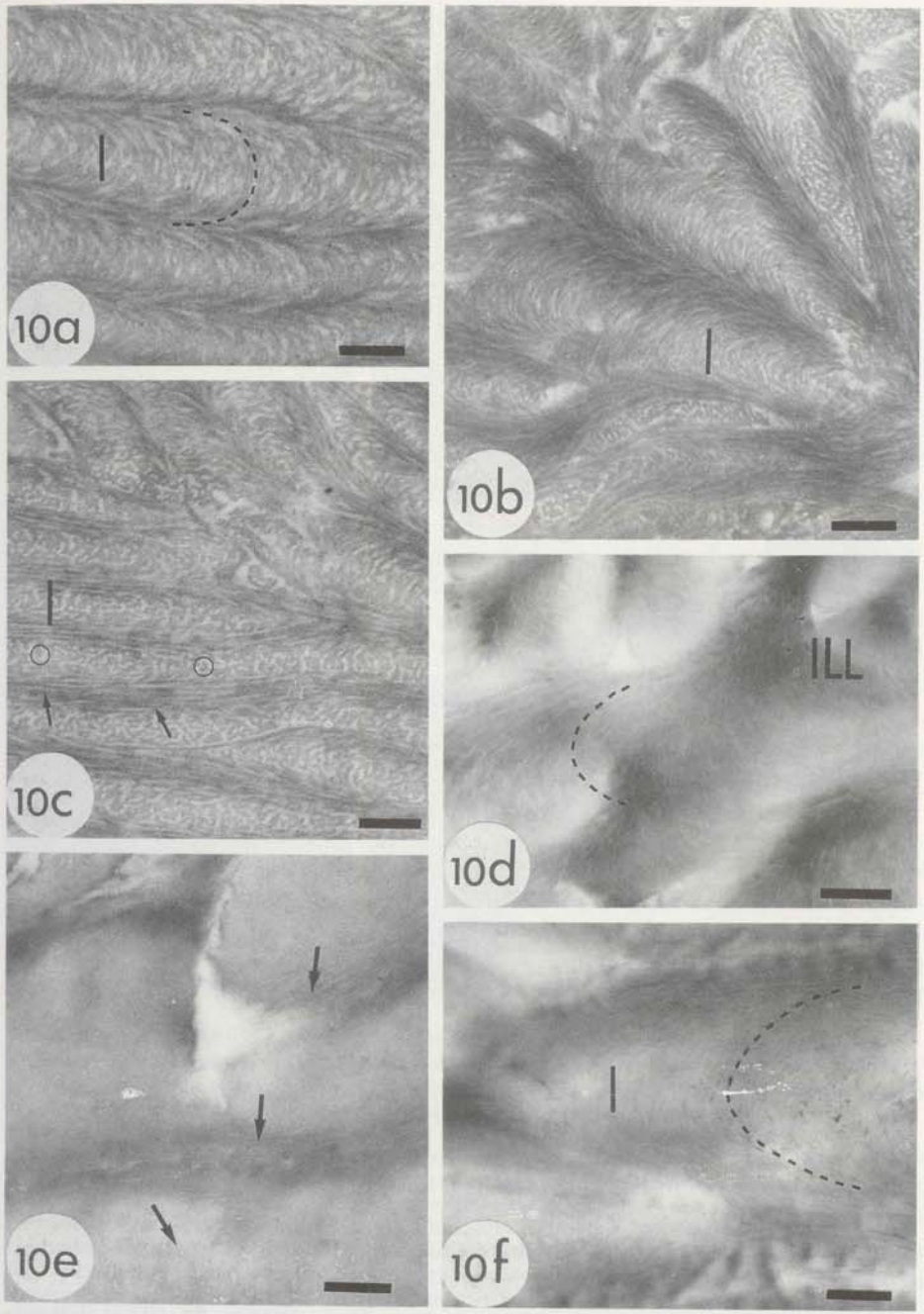


Fig. 10a–f. Transmission electron micrographs of thin sections cut through the chorion of *B. mori*, showing the fibrous ultrastructure of its lamellae. **a** The parabolic pattern of fibers (ca. 110 Å; *dotted lines*) within each lamella (*l*) in an oblique section of follicle 11/22 is seen. *Bar* 0.2 μm. **b**, **c** Fibrous lamellae (*l*) in oblique (**b**) and vertical (**c**) (with respect to the surface of chorion) sections of the 13/22 chorion. In **c** parabolic arcs disappear; in the lamellar boundaries, fibers cut longitudinally are seen as *parallel lines* (*arrows*) in contrast to those in the center, cut transversely, which are seen as *dots* (*small circles*). *Bar* 0.2 μm. **d** A micrograph showing the inner lamellar layer (*ILL*) of the chorion of follicle 15/22. Fibers ca. 90 Å in diameter form parabolic arcs within individual lamellae (*dotted lines*). *Bar* 0.45 μm. **e** Micrograph of the inner part of chorion (follicle 19/22) where the fibrous ultrastructure is obscured. *Arrows* show fibrillar remnants of the chorion lamellae. *Bar* 0.2 μm. **f** In the mature chorion of the follicle 33/33, fibers of approximately 30–40 Å in diameter can be seen. The *dotted lines* show the arrangement of the parabolic arcs of fibers within one lamella (*l*). *Bar* 0.2 μm. (Papanicolaou et al. 1986)

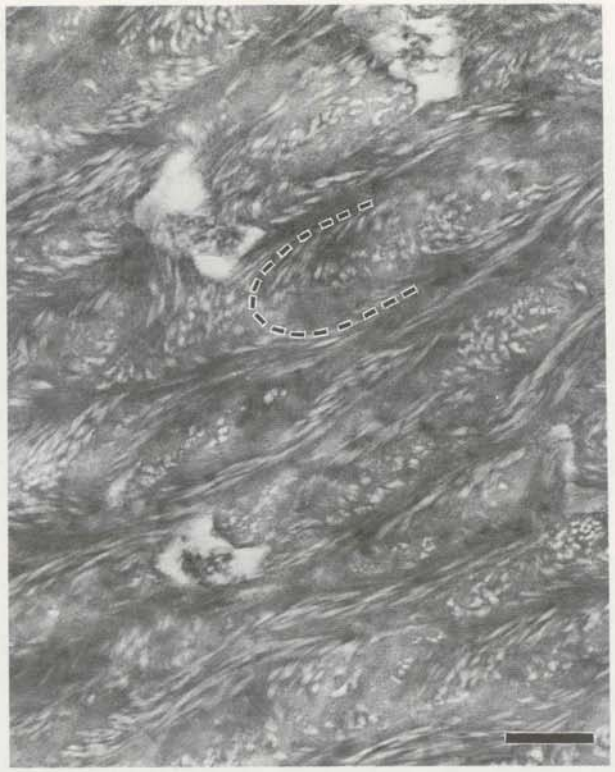


Fig. 11. Thioglycollic acid treatment (after Grierson and Neville 1981) reveals parabolic patterns of fibrils (*dotted lines*) in electron micrographs of oblique, thin sections cut through the eggshell of *S. gairdneri*. *Bar* 0.4 μm

true two-phase systems actually exist in the lepidopteran eggshell: different fixation and/or staining techniques give entirely different images of the fiber-matrix systems (Hamodrakas SJ, unpubl.). Furthermore, the technique of freeze-fracturing (cf. below) does not give any indication of a separate matrix phase. An attractive alternative, might be that the light (not heavily stained) areas of electron micrographs correspond to hydrophobic (not stain-absorbing) parts of the protein molecules constituting the "core" of eggshell fibrils, whereas the dark areas represent hydrophilic, stain-absorbing, portions of the molecules. It might, therefore, be that different portions of the same protein molecule(s) correspond to fiber

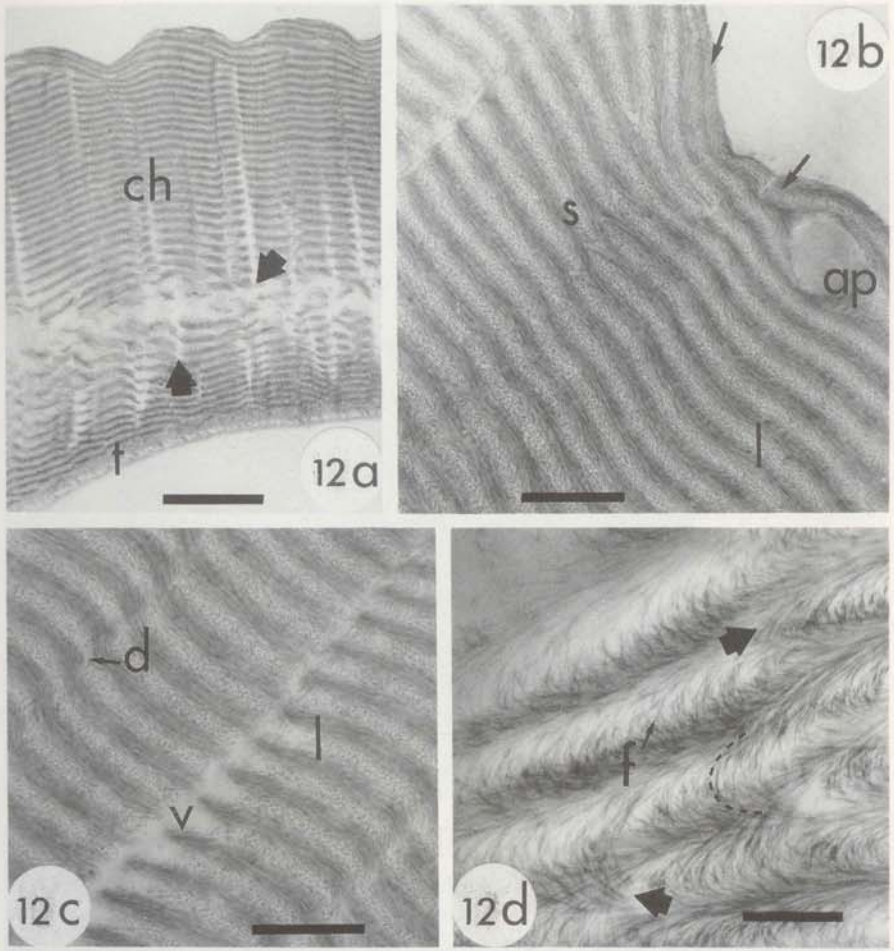


Fig. 12a–d. Transmission electron micrographs of thin sections cut through the chorion of *B. mori* from a number of follicles treated with a denaturing agent before fixation (6 M urea; see Papanicolaou et al. 1986). **a** Low power survey field of the chorion (*ch*) of follicle 19/38. Note the parallel arrangement of the lamellae and the vacancies (*arrows*) between them. *Bar* 5 μm. **b** The outer part of the same chorion at higher magnification reveals an empty aeropyle (*ap*) and a screw dislocation (*s*) (see also Mazur et al. 1982) within the lamellar organization (*l*) Note also the 180° change in the orientation of the parabolic pattern of fibers within individual lamellae, in certain chorion regions. *Bar* 1 μm. **c** Image of the inner part of chorion shown in **a**, illustrating in detail the lamellar organization. The parallel arrangement of the lamellae (*l*) and their fibrous ultrastructure is evident; *d* arrow, and *v* indicate a distortion and a vacancy in the lamellar organization, respectively. *Bar* 1 μm. **d** Fibrous ultrastructure of chorion from follicle 26/38. Thin individual fibers, ca. 30 Å in diameter (*f*, *arrow*) are clearly resolved, forming parabolic arcs (*dotted lines*). In several cases fibers of adjacent lamellae seem to overlap (*heavy arrows*). *Bar* 0.5 μm. (Papanicolaou et al. 1986)

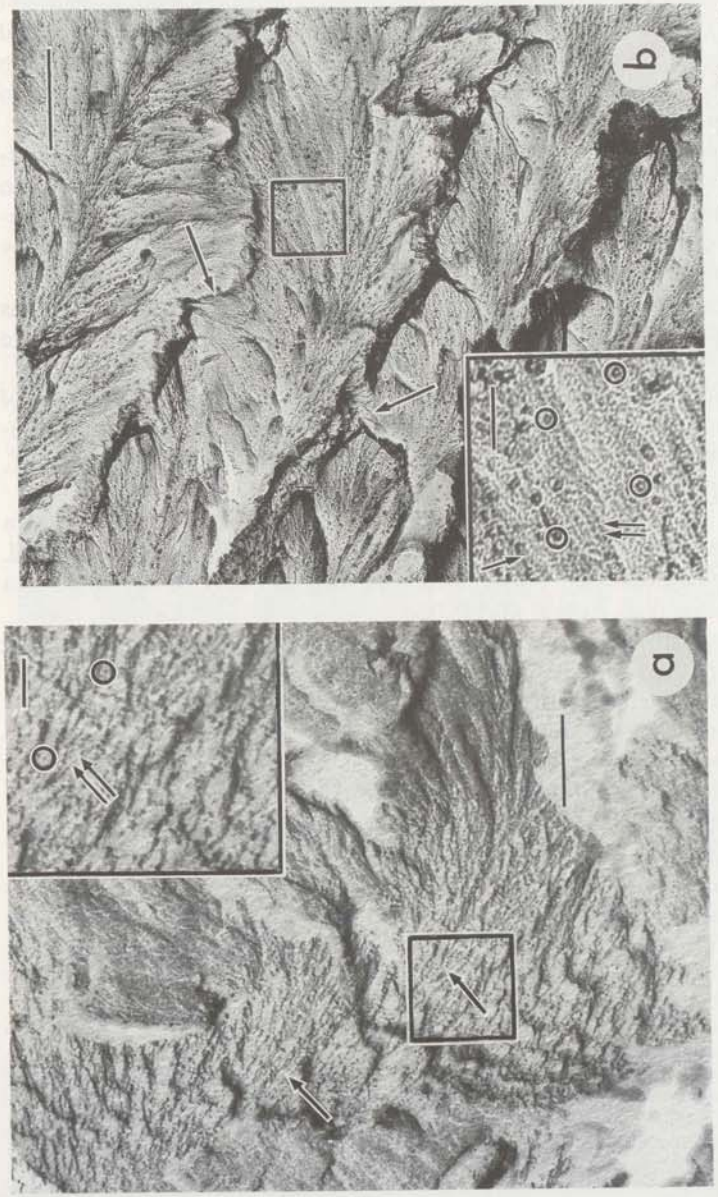


Fig. 13 a. Platinum/carbon (Pt_1C) replica of a freeze-fracture plane within the chorion of *A. polyphemus* (*unidirectional shadowing*). The replica shows clearly the fibrillar nature of chorion (*arrows*). At higher magnification (*inset*), filaments of diameter approximately 3–4 nm are longitudinally (*arrow*) or transversely (*circles*) seen. There are indications (*beading*; *multiple arrows*) that these filaments have helical structure. *Bar* 0.2 μm. *Inset*: *Bar* 60 nm. **b** Pt_1C replica of a freeze-fracture plane within the chorion of *A. polyphemus* (*rotary shadowing*). The fracture has advanced across successive lamellae producing a series of steps. Therefore the lamellar structure of silkmoth chorion is clearly seen. The helicoidal arrangement of chorion filaments is also obvious (*arrows*). In the inset (*boxed area*), at higher magnification, the helical (?) (*multiple arrows*) chorion filaments are seen either longitudinally (*arrows*) or transversely (*circles*). *Bar* 0.3 μm. Inset: *Bar* 50 nm. (Hamodrakas et al. 1986)

and matrix respectively. In the early stages of choriogenesis (Fig. 10) large stained areas may represent spaces, empty of protein, occupied by stain.

To discern microfibrillar patterns showing characteristic rows of parabolic "arcs", indicative of a helicoidal architecture, in sections of the fish eggshells cut obliquely (Fig. 11), we were forced to pretreat fish eggshells with thioglycollic acid (an agent disrupting disulfide bonds) following a method suggested by Grierson and Neville (1981). The method was first introduced by Filshie and Rogers (1962) to resolve fine structure in feather keratin.

Pretreatment of lepidopteran eggshells with a urea buffer (6 M urea, 0.4 M Tris-HCl, pH 8.5, a denaturing agent) substantially modifies the lamellar organization (Fig. 12), causing extensive rearrangements and texture disruption. Possible reasons for the significant changes observed and their meaning are discussed by Papanicolaou et al. (1986).

Transmission electron microscopy (lamellar organization — parabolic patterns of fibrils) provides only indirect evidence for the existence of helicoidal architecture in lepidopteran and fish eggshells.

Direct visualization of the helicoidal arrangement of fibrils for the formation of eggshell architecture was made possible by freeze-fracturing, utilizing both unidirectional and rotary shadowing (Hamodrakas et al. 1986; Orfanidou C, Hamodrakas SJ, Margaritis LH, and Gulik-Krzywicki T, in prep. Figs. 13, 14).

Apparently, in Lepidoptera, fibrils, hereafter called also "filaments", with a diameter of approximately 30–40 Å are the basic structural units of the eggshell. The beaded appearance of the fibrils along their long axis may be attributed to a helical substructure (e.g., Aebi et al. 1983).

It is perhaps instructive to remember at this point that Rudall, some 35 years ago (1955), had put forward the proposal that helical molecular structures are ideal candidates to serve as basic structural elements of a helicoidal architecture.

The existence of 30–40 Å fibrils as basic structural units of helicoidal proteinaceous eggshells, unraveled by freeze-fracturing, is in good agreement with X-ray diffraction (cf. below; Hamodrakas et al. 1984, 1986) and recent transmission electron microscopy and optical diffraction data (Hamodrakas SJ and Ottensmeyer FP, in prep.).

Freeze-fracturing has been used to reveal the helicoidal organization of microfibrils in the cuticle of a crayfish (Filshie and Smith 1980) and also in studies of the cholesteric liquid crystalline phases of polymer solutions (Livolant and Bouligand 1989) and cholesteric liquid crystalline DNA (Rill et al. 1989).

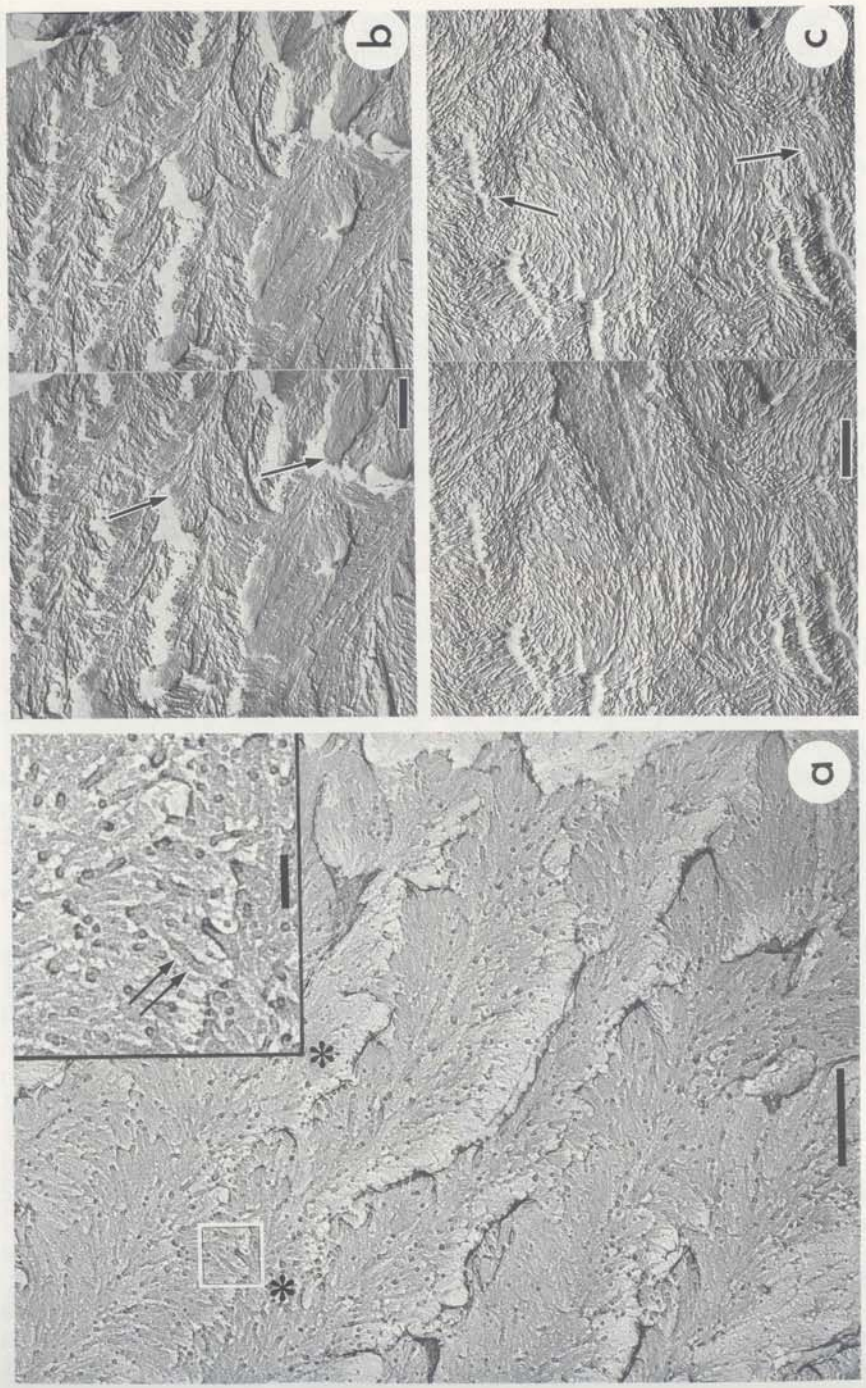


Fig. 14a. Pt/C replica of a freeze-fracture plane within the chorion of *M. sexta* (*rotary shadowing*). The fracture has advanced across successive lamellae producing a series of steps, revealing the lamellar structure of chorion and the helicoidal arrangement of its constituent fibrils (compare with **b**). In the inset (*boxed area*; *), at higher magnification, chorion fibrils are seen either tansversely (*circles*; 3–4 nm in diameter) or longitudinally (*double arrows*). The apparent beading of the fibrils suggests, most probably, a helical fibrillar structure. *Bar* 0.2 μm. Inset: *bar* 50 nm. **b** Stereoscopic view of a Pt/C replica of a freeze-fracture plane within the chorion of *M. sexta* (*unidirectional shadowing*) showing clearly the helicoidal packing modes of its constituent beaded (helical?) fibrils (*arrows*), 3–4 nm in diameter. *Bar* 0.2 μm. **c** Stereoscopic view of a Pt/C replica of a freeze-fracture plane within the chorion of *S. gairdneri* (*unidirectional shadowing*) showing the beaded structure of its constituent fibrils (*arrows*). *Bar* 0.3 μm.

Scanning electron microscopy also provides direct visualization of the helicoid in coelacanth scales (Giraud et al. 1978), the cuticle in *Carcinus maenas* and the test of *Halocynthia papillosa* (Gubb 1975).

Biological structures, in general, are formed by a hierarchy of assembled subunits, each forming the construction unit for the next size level of organization (Crick 1953); this has the advantage that errors are rectified en route by the rejection of faulty components. Helicoidal proteinaceous eggshells are not an exception to this rule; they are typical examples of protein engineering in vivo (Hamodrakas et al. 1988).

The basic structural unit in Lepidopteran helicoidal eggshells appears to be a fibril 30–40 Å in diameter. The next levels of hierarchy are: packing of fibrils into sheets, stacking of sheets one on top of another, helicoid formation.

Questions of interest which remain unanswered to date are: what is the structure of a fibril? Is it formed by one or more protein molecules? How are these protein molecules folded and how do they interact? We shall try to provide partial answers to these questions below. However, it is clear that any simple or heuristic model of self-assembly for helicoidal proteinaceous eggshells should be based on the determination of the folding patterns of individual protein molecules and of protein intermolecular interactions.

2.2.2 Morphogenesis of the Eggshell

Comparison of chorion morphogenesis and protein biochemistry among species offer considerable promise for answering fundamental questions about the relationship between development and evolution (Regier and Vlahos 1988). Unfortunately, to date, a detailed description of Lepidopteran chorion morphogenesis is restricted to only four species: *H. cecropia*, *A. polyphemus*, *B. mori* and *M. sexta* (King and Aggarwal 1965; Telfer and Smith 1970; Smith et al. 1971; Mazur et al. 1980; Regier et al. 1982; Papanicolaou et al. 1986; Regier and Vlahos 1988).

A superb review on silkmoth chorion morphogenesis has been presented by Mazur et al. (1982).

Four different modes of growth appear to occur during formation of the lamellar chorion following the construction of the trabecular layer (Mazur et al. 1982): (1) lamellogenesis or framework formation; most lamellae are deposited early in choriogenesis, (2) permeation or intercalation of newly secreted protein components that leads to expansion of this framework, (3) densification by addition of protein components, and (4) regionalization that leads to apposition of additional lamellae and surface sculpturing.

Species-specific variations are evident (refer to original publications for detailed descriptions) and an evolutionary model has been presented to account for the differences observed (Regier and Vlahos 1988).

Views of sequential events occurring during *B. mori* chorion formation and morphogenesis of chorion components are given as examples in Figs. 15, 16, 17.

Little is actually known about the mechanisms operating during fish eggshell morphogenesis, although the literature concerning fish eggshell formation is relatively rich (cf. Groot and Alderdice 1985). A detailed review of this subject is

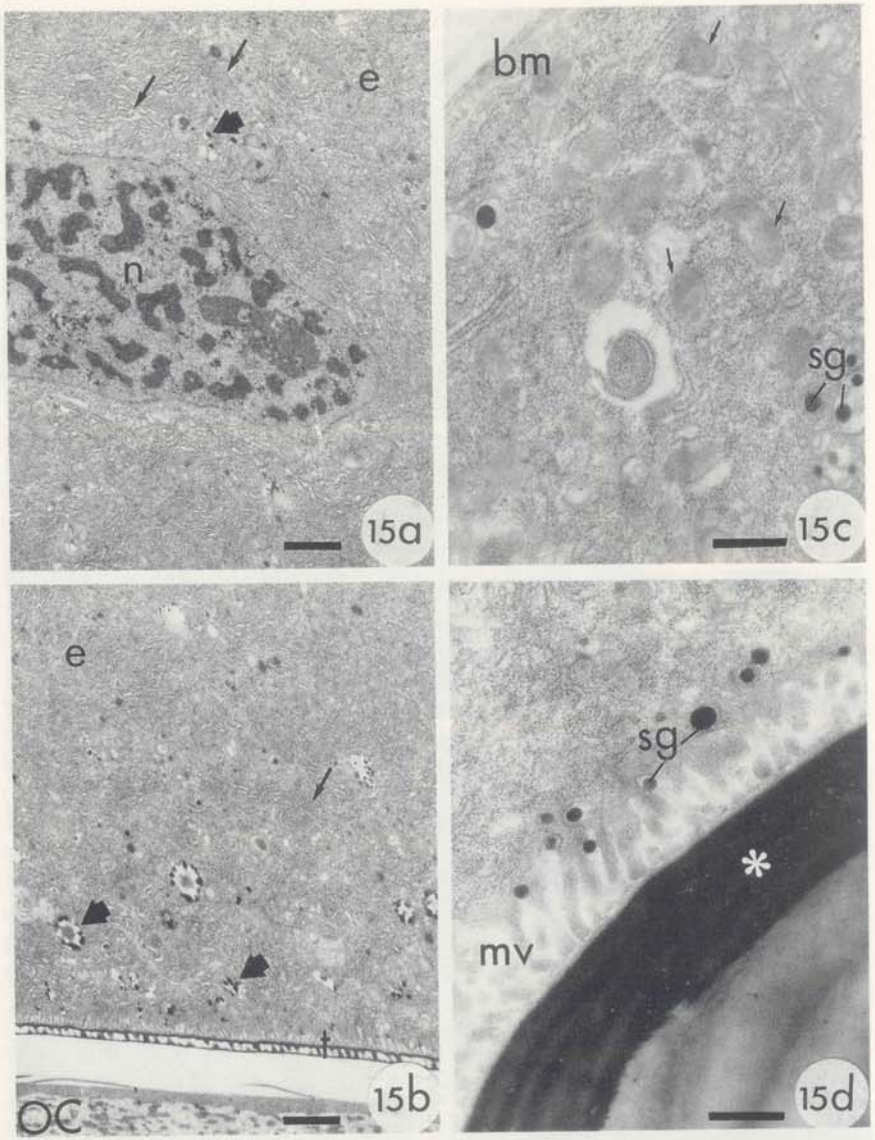


Fig. 15. a, b Transmission electron micrographs of thin sections cut through a follicle cell of *B. mori* (*e* follicle 1/22) showing: **a** its nucleus (*n*), and **b** the cytoplasm close to the secretory region. The high synthetic activity of the cell is apparent by the abundance of RER and Golgi complexes (*arrows*) and also by the osmiophilic secretory granules (*heavy arrows*) in its cytoplasm. *t* Trabecular layer; *OC* oocyte. *Bar* 1 μm. **c, d** Transmission electron micrographs of thin sections cut through a follicle cell (follicle 19/22) showing: **c** part of its cytoplasm close to the basal membrane (*bm*) and **d** part of its cytoplasm near the outer surface of the chorion. Secretory granules (*sg*) full of osmiophilic material are seen between, near, or within the microvilli (*mv*) and also close to the basal membrane. A large number of mitochondria (*thin arrows*), RER membranes, and ribosomes are also seen in the cytoplasm; * osmiophilic layer; *Bar* 0.5 μm. (Papanicolaou et al. 1986)

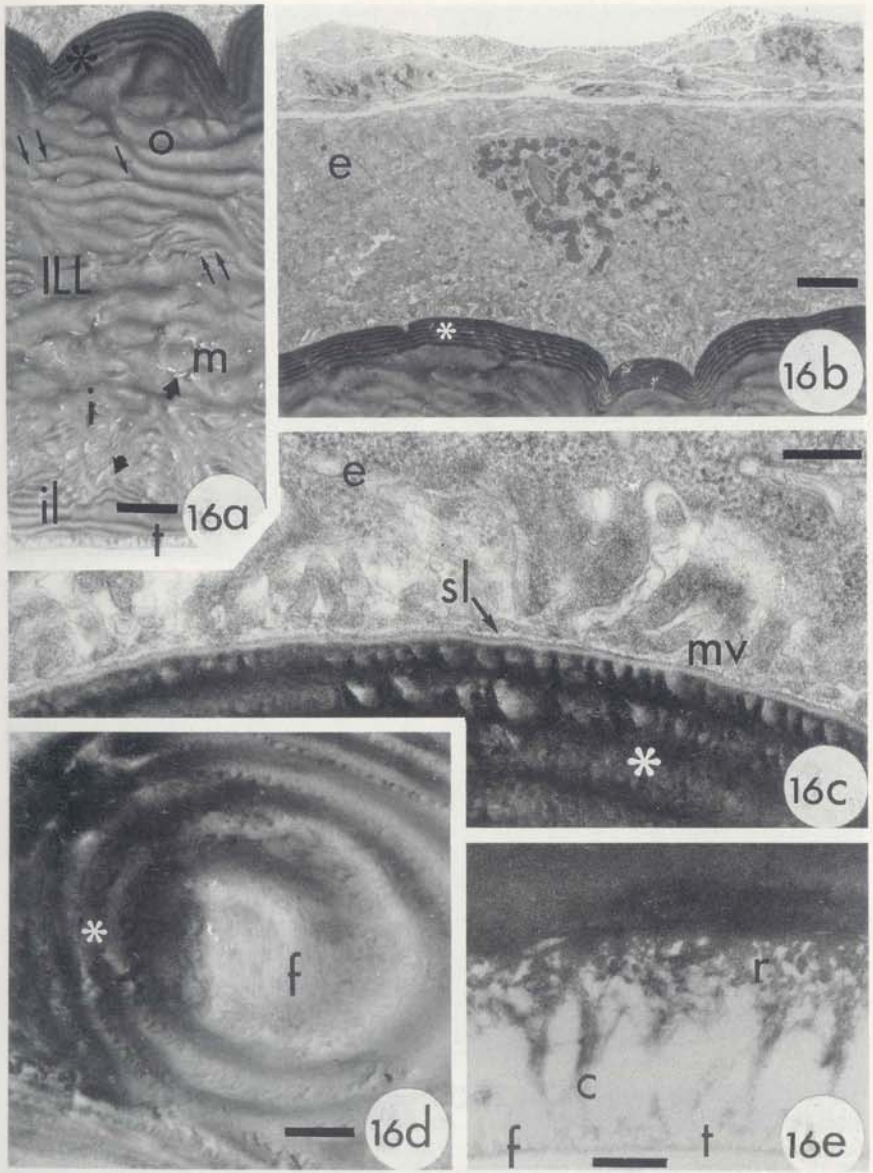


Fig. 16a–e. Transmission electron micrographs of a nearly mature *B. mori* follicle, 22/22. **a** Cross-section of the chorion. Lying closest to the oocyte is the trabecular layer (*t*) and then a thick inner lamellar layer (*ILL*) followed by an osmiophilic outer layer (*). The ILL can be divided into four sublayers, on the basis of their differences in lamellar thickness and orientation; the innermost sublayer (*il*), consisting of two thin, parallel lamellae close to the trabecular layer; the inner sublayer (*i*), consisting of very thin randomly oriented lamellae; the middle sublayer (*m*) with four to five thick lamellae; and the outermost sublayer (*o*) made up of lamellae of moderate thickness. The bulk of the inner lamellar layer

beyond the scope of this chapter, and the reader should consult for details original publications by Anderson (1967), Flugel (1967), Wourms (1976) and our forthcoming article (Papadopoulou P and Hamodrakas SJ, in prep.) as examples.

3 The Biochemistry of Eggshell Proteins

Detailed reviews of silkmoth chorion protein biochemistry have been made by Kafatos et al. (1977) and Regier and Kafatos (1985).

Biochemically, silkmoth chorion is surprisingly complex: as many as 186 protein components have been resolved by two-dimensional gel electrophoresis, from the chorions produced by an individual *A. polyphemus* moth (Regier et al. 1980) and more than 150 different polypeptides in *B. mori* (Regier and Kafatos 1985).

Silkmoth chorion proteins have been classified into four classes called A to D, based on observed clustering in SDS gels. The most abundant are the A's, B's, and C's which together account for approximately 97% of total chorion mass. In *A. polyphemus* the A's, B's, and C's represent about 38, 50, and 9% respectively of the chorion's dry mass, whereas the D's are quantitatively minor. (Kafatos et al. 1977). Molecular weight ranges for A, B, and C proteins are approximately 9000–12 000, 12 000–14 000, and 16 000–20 000 Da, respectively.

A fifth class of minor proteins, the E's (divided into two subclasses, E1 and E2, with molecular weights approximately 15 000 and 85 000 Da, respectively), uniquely assemble to form the "filler" substructure, the bulk of which is localized as little bundles resting on *A. polyphemus* chorion surface, in the so-called aeropyle crown region (Mazur et al. 1980; Regier 1986).

In *B. mori*, another class of species-specific proteins, the Hc's (high-cysteine), are responsible for the formation of the outer osmiophilic eggshell layer (Figs. 15, 16, 17), presumably a very tough weather-resistant covering; *B. mori* embryos are the only silkmoth embryos that undergo prolonged periods of diapause, thus requiring protection. This class is further subdivided into HcA and HcB which are homologous to A's and B's respectively (Rodakis and Kafatos 1982).

shows many small scattered discontinuities (*small arrows*) along with small holes (*heavy arrows*) especially in its inner parts. The outer osmiophilic layer consists of five to six parallel lamellae which are uniform in thickness and show very few or no discontinuities. The osmiophilic layer forms the knobs visible in scanning electron micrographs, in the outer surface of the chorion. *Bar* 2 μm. **b** View of the 11-μm-thick follicular epithelium (*e*) which is attached to the osmiophilic layer. *Bar* 3 μm. **c** Details of the osmiophilic layer (*) and the microvilli (*mv*) of the follicle cells (*e*). In between lies the thin sieve layer (300 Å), consisting of three minor layers (*sl, arrow*). Its central layer appears as an electron-dense line ca. 150 Å thick. *Bar* 0.2 μm. **d** Spiral arrangement of the osmiophilic layer as seen in a tangential section of a knob (*). Such spirals were also documented by Bouligand (1972) in the crub cuticle. In the center of the spiral the fibrous (*f*) ultrastructure of the osmiophilic layer is revealed. *Bar* 0.4 μm. **e** Closer view of the trabecular layer (*t*) formed by fibrous columns (*c*) separating two, also fibrous, spongy layers: the roof (*r*) and the floor (*f*). *Bar* 0.3 μm. (Papanicolaou et al. 1986)

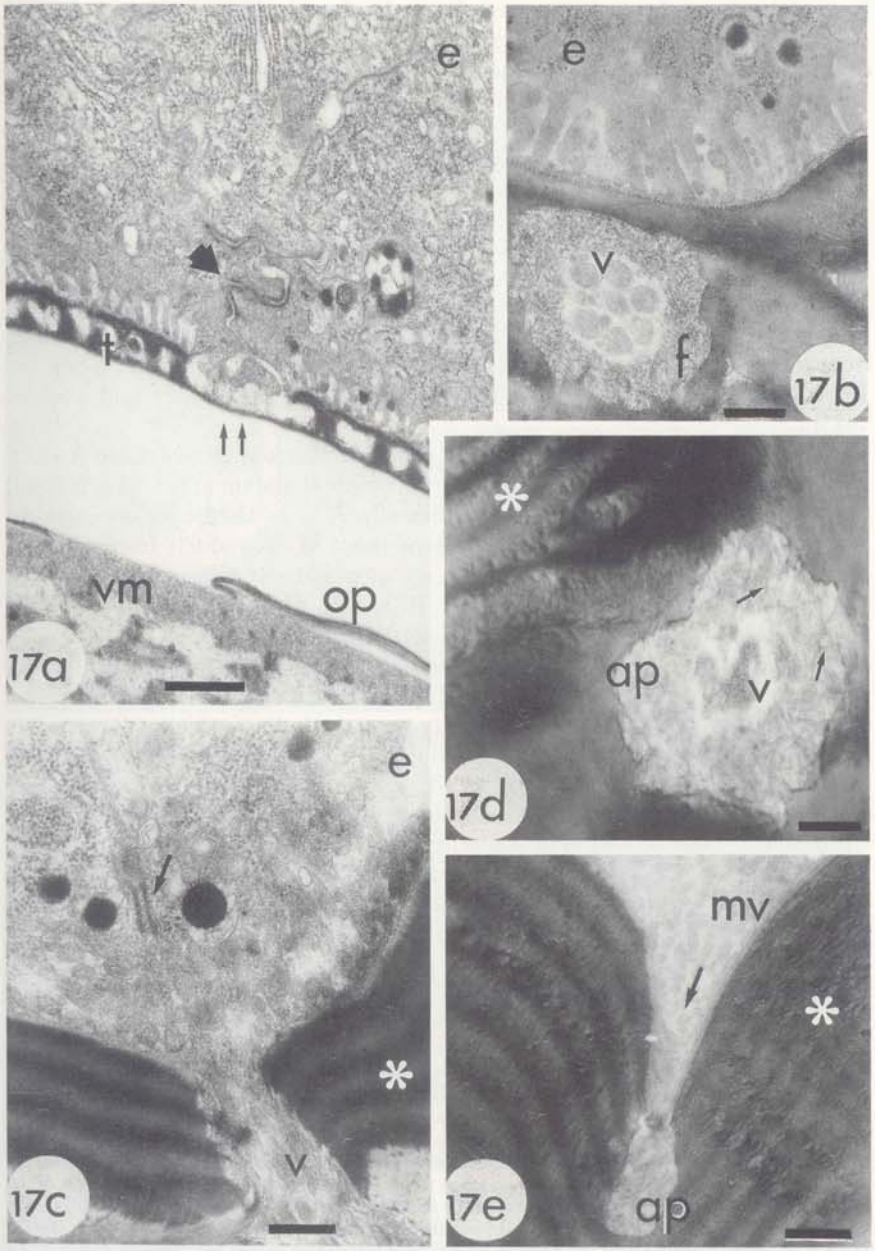


Fig. 17a–e. Transmission electron micrograph of an early stage of choriogenesis (follicle 1/22) showing the formation of an aeropyle in *B. mori*. **a** The aeropyle channel (*arrows*) forms around a bundle of extended microvilli, under a cell junction (*heavy arrow*), during the formation of the trabecular layer (*t*): *e* follicle cell; *vm* vitelline membrane; *op* overlapping plates of the vitelline membrane. *Bar* 0.4 μm.

The C class of proteins, which constitutes the bulk of the "early" proteins, responsible for the formation of the initial chorion framework and perhaps required for the organization of the quantitatively dominant components which are secreted later, can be subdivided further into two subfamilies CA and CB, sharing homologies with A's and B's respectively (Regier et al. 1983; Lecanidou et al. 1986)

It appears that the number of silkmoth genes is at least as high (Eickbush and Kafatos 1982). However, most chorion genes are related: they belong to a small number of gene families (A, B, CA, CB, HcA, HcB...), each encompassing multiple genes that arose during evolution by reduplication followed by sequence divergence (see, for example, Regier et al. 1978; Jones and Kafatos 1980, 1982). These multigene families are themselves related, and constitute a superfamily, with one branch encompassing the A, CA, and HcA families, and the other the B, CB, and HcB families (Rodakis and Kafatos 1982; Regier et al. 1983; Goldsmith and Kafatos 1984; Iatrou 1984; Rodakis et al. 1984; Eickbush et al. 1985; Burke and Eickbush 1986; Spoerel et al. 1989; Lecanidou et al. 1986). Clear diagrams of sequence relationships and nomenclature in the moth chorion gene superfamily were given by Lecanidou et al. (1986), who called the two branches α and β.

Interestingly, in both *M. sexta* (Regier and Vlahos 1988). and *S. nonagrioides* (Orfanidou C and Hamodrakas SJ unpubl. data), three major protein classes are also resolved on SDS-polyacrylamide gels, the A's, B's, and C's, with similar molecular weights to the corresponding classes of silkmoth chorion. Unfortunately, sequence data are not available to date for members of these classes; however, they are predicted to share sequence homologies with silkmoth chorion proteins (Regier and Vlahos 1988).

Primary sequences have been determined for all six silkmoth gene families and their products (references listed above), comprising a substantial amount of fascinating and, we hope, interpretable information.

Protein sequence comparisons (Fig. 18) and predictions of secondary structure (cf. below) have revealed that chorion proteins have a tripartite structure (Hamodrakas et al. 1982a, b, 1988; Fig. 19). A central domain is highly conserved within each family and can be recognized as highly homologous between families of the same branch (α or β).

A and B central domains show, however, distant similarities, suggesting that the chorion genes constitute a superfamily derived from a single ancestral gene (Lecanidou et al. 1986 and references therein).

b Electron micrograph of a thin cross-section cut through an aeropyle channel in the outer chorion of follicle 15/22. The extended villi (*v*), which are used as a mold for the formation of the channel, are closely packed and surrounded by a loose fibrous matrix (*f* filler); *e* follicle cell. *Bar* 0.3 μm. **c** Thin section of the top of an aeropyle of 19/22 follicle. A bundle of microvilli (*v*) traverses both the distupted sieve layer and the osmiophilic layer (*) forming the basis for a future air channel. The aeropyle is formed under a cell junction (*arrow*) of the follicular epithelium *e*. *Bar* 0.3 μm. **d** Thin section of an aeropyle (*ap*) right under the osmiophilic layer (*) of follicle 22/22. Inside the villi (*v*) of the channel, numerous cytoplasmic ribosomes can be resolved. Note also the loose fibers (*arrows*) of the channel filler. *Bar* 0.3 μm. **e** A micrograph indicating that the furrow (*arrow*) between the osmiophilic knobs (*) of the outer surface of the chorion is a site where an aeropyle (*ap*) is formed. *mv* microvilli (follicle 21/22). *Bar* 0.4 μm. (Papanicolaou et al. 1986)

Fig. 18. a Sequence of typical A and HcA proteins determined from complementary DNA clones: pc18, pc609, and pc202 are members of the A family from *A. polyphemus* (Jones and Kafatos 1982), m2774 is an A component from *B. mori* (Rodakis et al. 1982) and m1911 is an HcA component from *B. mori* (Rodakis and Kafatos 1982). *Unbroken vertical lines* flank the portions of the sequences considered in this chapter, and a *broken line* indicates a slightly different border for the previously defined central domain (Jones and Kafatos 1982). *Dots* indicate that the m2774 sequence is incomplete. (Hamodrakas et al. 1985). **b** Sequences of typical B, HcB, and CB proteins. *Unbroken lines* mark the borders of the central domain and *broken lines* flank the portions of the sequences considered in this work. Proteins Be 2G12, Bm2807, and Bm1768 are representative proteins of the B family of *B. mori*, B Hc-B12 of the HcB family of *B. mori*, B pc401 of the B family of *A. polyphemus*, and C pc404-H12 of the CB family of *A. polyphemus*. (Hamodrakas et al. 1988)

Fig. 19. Schematically, the tripartite structure of the silkmoth eggshell proteins. A central highly conservative and regularly structured domain and two, more variable, flanking "arms", particularly enriched in cysteine (*C*), constitute each protein. (Hamodrakas 1984)

The flanking amino- and carboxyl-terminal domains or "arms" are more variable and are marked by the presence of characteristic tandemly repetitive peptides that do not appear in the central domain. Nevertheless, "arms" also show similarities both within and among chorion families.

Ultimately, we want to relate the primary sequences and tripartite composition of chorion proteins to the fibrous structures they assume, and the regular assemblies of these fibers. Although chorion is very complex, this undertaking is facilitated by the evolutionary relatedness of the components. Limited evolutionary variations amount to information that can help specify further the fundamental molecular properties of these proteins.

Lepidopteran chorion becomes insoluble by disulfide bond formation between its constituent proteins (Kawasaki et al. 1971). This process has been called cementing (Blau and Kafatos 1979), and has been demonstrated to occur in a relatively short period (8 h) around the time of ovulation (Hamodrakas et al. 1982b, Regier and Wong 1988).

In contrast, developing chorions isolated from all stages of choriogenesis prior to ovulation can be almost completely dissolved with protein denaturing agents only (e.g. guanidinium chloride, SDS, 6 M urea) without a sulfhydryl reducing agent (e.g., dithiothreitol or β-mercaptoethanol) (Kafatos et al. 1977; Regier and Wong 1988; Orfanidou C and Hamodrakas SJ, unpubl.).

Several important findings for silkmoth chorion assembly emerged from the work of Regier and Wong (1988), who have tried to find in vivo patterns of disulfide bond formation. These can be summarized as follows:

During choriogenesis, low molecular weight disulfide-bonded chorion protein multimers form that are soluble in the absence of a reducing agent.

Multimers generally contain two to six protein subunits, with dimers and trimers the most abundant. Within multimers, only a small portion of cysteine residues per protein are involved in disulfide bond formation. These multimers, that assemble inside follicle cells, within minutes after protein synthesis, are, apparently, secreted normally (secretion rate is the same as that of monomers), and remain intact (their sizes and concentration do not increase) throughout choriogenesis.

Only a relative minor fraction of the total chorion (5–10% of total) proteins participate in multimer formation. Even so, all classes of chorion proteins are present in multimers. Furthermore, certain linkage patterns are preferred. In particular, class B proteins cross-link abundantly with A's and with themselves, but A's do not abundantly cross-link with themselves.

This evidence suggests that multimers represent normal intermediates in the pathway of chorion protein assembly.

Our opinion is that they may serve as "nuclei" in the "crystallization" process of self-assembly, "guiding", perhaps, the modes of packing of monomers by non-covalent weak interactions, and presumably playing an important role in the "cementing" process as well.

However, multimer formation during choriogenesis and "cementing", which occurs around the time of ovulation, appear to represent quite distinct phenomena, despite the fact that they both depend on formation of disulphide bonds.

Biochemical data on fish eggshell proteins are unexpectedly poor. Apart from papers determining the amino acid composition of fish eggshell proteins, which indicate the relative abundance of Glu, Asp, Pro, and of certain hydrophobic and polar residues (e.g., Leu, Ala, Thr) in these proteins (Ohzu and Kusa 1981; Kobayashi 1982), to our knowledge, no other significant work on fish eggshell protein biochemistry has been published.

Apparently, attempts to study fish eggshells are hindered by the fact that isopeptide bonds between glutamic and/or aspartic acid and lysine/arginine residues (peptide bonds formed between the side chains of these residues, e.g., Glu-Lys) render fish eggshells insoluble to normal denaturing and reducing agents and/or their combinations. It has been shown conclusively that these strong covalent bonds are primarily responsible for the process of "hardening" in the fish eggshell, a result of the so-called cortical reaction thought to occur after water activation (Hagenmaier et al. 1976; Lonning et al. 1984; Davenport et al. 1986).

We, ourselves, have experienced great difficulties in solubilizing eggshells isolated from *S. gairdneri* eggs and characterizing their constituent proteins on SDS-polyacrylamide gels. Recently, we have been able to obtain electrophoretic patterns, which suggest the presence of a complex set of proteins in the fish eggshell, with molecular weights in the range 12–100 kDa, (with prominent bands at approximately 12, 18, 32, 35, 43, 48, 54, 70, 85 and 96 kDa some of which might be arising from multimers), but, these data should only be considered as preliminary (Papadopoulou P and Hamodrakas SJ, unpubl.).

4 Elucidation of Eggshell Protein Structure

Between the detail revealed by low resolution (20–30 Å) electron microscopy and the wealth of information available in the determined amino acid sequences of silkmoth chorion proteins there is an information gap. This is a well known problem in fibrous proteins (e.g., keratins, collagen, etc.) and their supramolecular assemblies; it has been proved to be an extremely difficult task to bridge this gap

with the use of existing experimental methods (Fraser and McRae 1973; Squire and Vibert 1987).

Our work on silkmoth chorion and similar-in-structure proteinaceous eggshells has been focused in attempts to partially "fill" this gap.

Ideally, this can be achieved by isolating, crystallizing, and solving the three dimensional crystal structure of individual protein components by X-ray crystallographic methods near atomic resolution, and, in the next step, determining the modes of protein-protein interactions for the formation of fibrils and of higher order helicoidal structure. Unfortunately, this approach seems to be unrealistic at present, since it requires the use of suitable single crystals of fibrous chorion proteins which cannot be produced easily (if at all!) and for various other obvious reasons.

Alternatively, a promising approach would be to isolate individual protein components, from suitable fibers and study these fibers by X-ray diffraction to determine the major structural features of the molecules. However, this is a difficult task because of the substantial complexity of chorion proteins. Nevertheless, it remains one of our future aims although our preliminary attempts towards achieving this goal have failed.

Therefore, we decided to perform an "anatomy" of silkmoth chorion protein sequences, hoping to determine folding patterns and modes of protein-protein interactions, utilizing information "hidden" in the primary sequences. We considered it likely that this undertaking might be facilitated by the possible existence of periodicities (repetitions or tandemly repeating motifs) in the primary sequences of silkmoth chorion proteins, a characteristic and common feature of fibrous protein structure (Fraser and McRae 1973).

In the last two decades, the amino acid sequences of >3000 proteins have been determined (Barker et al. 1986), and also the three-dimensional structure of >300 proteins has been elucidated mainly by single crystal X-ray crystallographic methods and more recently by NMR methods as well, at atomic or near atomic resolution (Bernstein et al. 1977). Since experimental findings indicate that all the necessary information for a protein to fold into its native structure is coded into its amino acid sequence (Anfinsen 1973), several attempts have been made to predict three-dimensional structure from sequence (Taylor 1987 and references therein). In these attempts the correct prediction of elements of secondary structure, frequently, plays a key role, since these elements may represent the initial nuclei in the process of protein folding.

For various other reasons, not mentioned here but described extensively in recent reviews (see Hamodrakas 1988 and references therein) it is also important to predict correctly the secondary structure of proteins from their sequence alone.

In the case of chorion proteins, there is also another reason which makes valuable the correct prediction of their secondary structure elements: the packing of protein molecules, which dictates self-assembly, is known to depend on the interactions of their secondary structure elements (α-helices and β-sheets). Therefore, the first step towards unraveling the modes of interaction of protein molecules for the formation of helicoidal proteinaceous eggshells is the determination of eggshell protein secondary structure.

Our approach was both theoretical (secondary structure prediction) in cases where amino acid sequences were available and experimental (X-ray diffraction and laser-Raman, infrared and circular dichroism spectroscopy) to: (1) verify theoretical results and (2) determine secondary structure when protein primary sequence information was lacking.

In the course of our investigation an attractive idea emerged: the possibility of the existence of a universal folding pattern which dictates the formation of the helicoidal architecture in proteinaceous eggshells, biological analogs of cholesteric liquid crystals. All evidence, both theoretical and experimental, in the systems studied so far, favors this notion.

4.1 Secondary Structure Prediction

Although amino acid sequence dictates native conformation, secondary structure predictions based on primary sequences should be undertaken with full awareness of their limitations. Even in the case of globular proteins, for which they were initially developed and applied, the accuracy of prediction methods is limited (Chou and Fasman 1978). The methods fail altogether when short precise repeats are present, and thus have been applied only rarely to structural proteins in which internal repeats are widespread. However, comparisons of evolutionarily related sequences or imprecise internal repeats are invaluable in overcoming these limitations: limited variation should reduce the "noise" and help identify consistent structural features (e.g., Hamodrakas and Kafatos 1984; Hamodrakas et al. 1989; Aggeli et al. 1991).

Several prediction algorithms have been published and can be classified into two categories — statistical or stereochemical — but their success has been rather limited. It has been claimed that combined prediction schemes provide a higher degree of accuracy than individual prediction methods and on this basis we have constructed and published an algorithm which combines six of the most popular and successful prediction methods. The prediction package runs on microcomputers and requires as input only the amino acid sequence of the protein(s) under study (Hamodrakas 1988).

The algorithm was applied to the determined amino acid sequences of the A, B, and C families of silkmoth chorion proteins. Detailed results on representative, predominant members of these families are shown in Figs. 20, 21, 22 (Hamodrakas et al. 1982a; Regier et al. 1983).

For each protein individual predictions for α-helix (α), β-pleated sheet (β) and β-turns (T) were made according to six prediction methods. Joint prediction histograms were then constructed by tallying individual predictions since they are more dependable than individual prediction schemes. Structures predicted by three or more methods (out of six, or out of five in the case of turns) were considered as most probable and are shaded. To make secondary structure prediction plots most comparable sequences within each family or sequences from different homologous families were aligned.

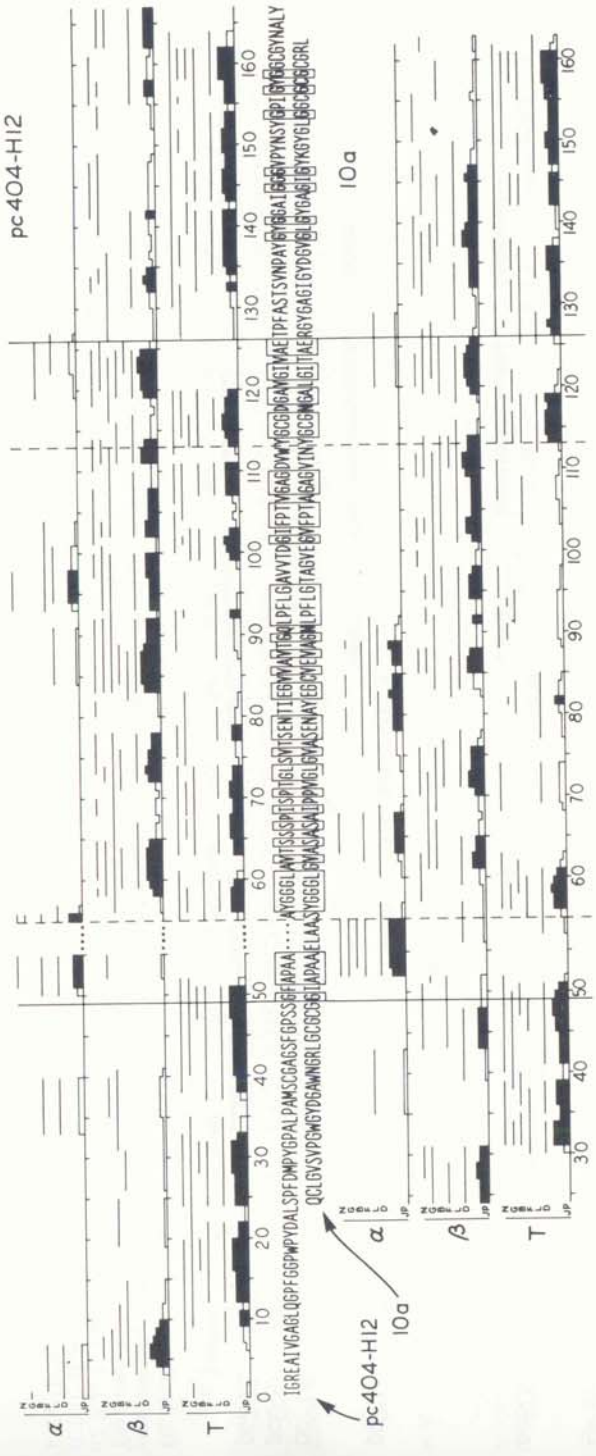


Fig. 20. Secondary structure predictions for protein sequences of the pc404-H12 (CB family) and 10a (B family) chorion components. Individual predictions for α-helix (α), β-sheet (β), or β-turn (*T*) are shown by *horizontal lines*, as derived according to Nagano (*N*), Garnier et al. (*G*), Burgess et al. (*B*), Chou and Fasman (*F*), Lim (*L*), and Dufton and Hider (*D*). See Hamodrakas et al. (1982a) for a complete listing of references. Joint prediction histograms (*JP*), constructed by tallying the individual predictions, are also shown. The most probable structures, predicted by three or more methods, are *shaded*. Sequences are numbered from the amino- to carboxyl-terminus of pc404-H12. Identical residues are enclosed in *boxes*. A gap between positions 55 and 56 in pc404-H12 is indicated by *dots*. The *solid vertical lines* indicate the borders of the clear interfamily homologies; the *dashed vertical lines*, the conservative central region of B protein sequences. (Regier et al. 1983)

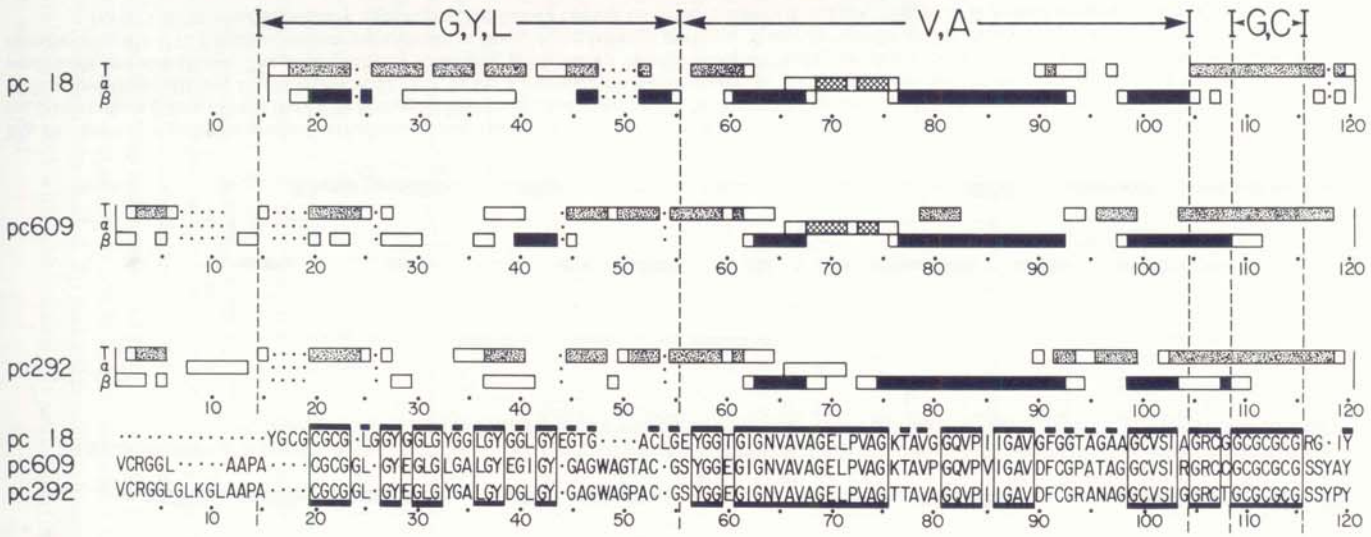


Fig. 21. Summary of joint prediction histograms for secondary structure in proteins of the A family of silkmoth chorion proteins. Structures predicted by two methods are shown as *open rectangles*, while the most probable structures predicted by three or more methods are shown *solid* (β-sheet), *hatched* (α-helix), or *stippled* (β-turn, *T*). The amino- and carboxyl-termini are indicated by *short vertical lines*, and regions enriched in certain amino acids are outlined by *dashed lines* and named accordingly (*G, Y, L*; *V, A*; *G, C*). Sequences are aligned and numbered, and necessary gaps are indicated by *dots*. The actual sequences are presented at the bottom, according to the IUPAC-IUB one-letter code, as follows: *A* Ala; *C* Cys; *D* Asp; *E* Glu; *F* Phe; *G* Gly; *H* His; I Ile; *K* Lys; *L* Leu; *M* Met; *N* Asn; *P* Pro; *Q* Gln; *R* Arg; *S* Ser; *T* Thr; *Y* Tyr; *V* Val; *W* Trp. *Blocks* outline two or more residues which are invariant in these and four additional related sequences (18b, 18c, 292a, 292b); individual invariant residues are *overlined*. (Hamodrakas et al. 1982a)

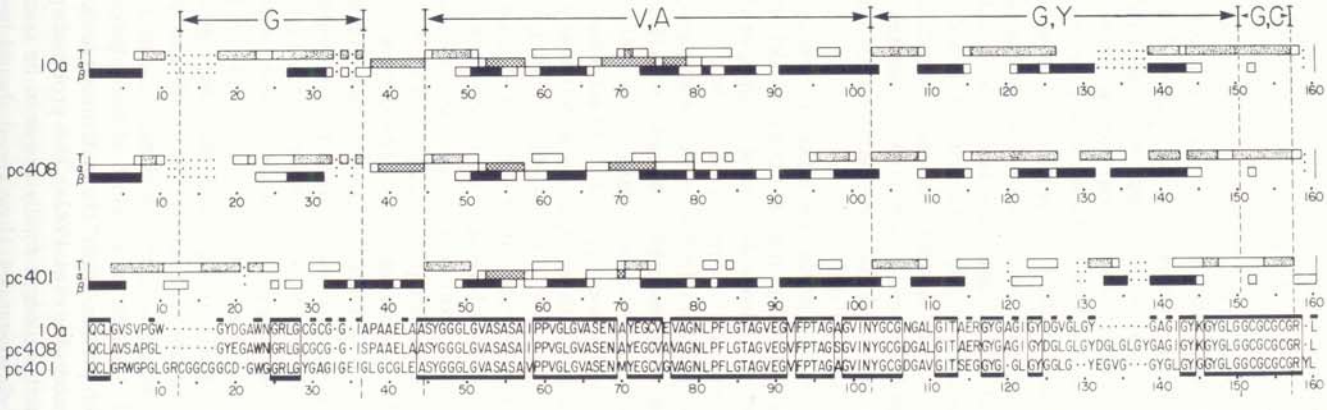


Fig. 22. Summary of joint prediction histograms for secondary structure in proteins of the B family. Details as in Fig. 21. Blocks and overlinings refer to four additional sequences (pc10, 10b, 401a, 401b). (Hamodrakas et al. 1982a)

The diagrams are accompanied by the actual protein sequences presented in the one letter code and displayed in a manner emphasizing regions rich in certain amino acids and also sequence conservation (Jones and Kafatos 1982).

The prediction results clearly indicate that in members of the A, B, and C families β-sheet structure predominates, whereas α-helix is almost totally absent. β-Turns are predicted frequently, and may often connect short, presumably antiparallel β-sheet strands. The evidence that the β-sheets are antiparallel has been provided from laser-Raman and infrared spectroscopy data (see Sect. 4.3) and for purely topological reasons (short connections between β-strands).

The proteins can be divided into a number of distinct regions or "domains" according to the degree of evolutionary constancy in sequence, the amino acid composition and secondary structure features. They appear to have a tripartite structure (Figs. 19, 20, 21, 22): a central domain ("core") highly structured into β-sheet structure, evolutionarily highly conserved, rich in Val, Ala, and two more variable flanking "arms" (or "tails"), varying in length, particularly enriched in Cys in the most abundant families A and B (and in their homologous Hc's in *B. mori*), which appear to be less structured and presumably reflect protein and/or family specific functions.

The most intriguing feature of prediction, which, however, was the clue that led us to the proposal of the detailed protein structural models presented below (cf. Sect. 5) was an apparent periodicity of β-sheet strands in the central domain evidenced by the nearly periodic occurrence of maxima for β-sheet predictions (Fig. 20; Hamodrakas et al. 1982a). In between these maxima, β-turns were also predicted. This feature, which is obscured in the summaries (Figs. 21, 22), indicated a regular model of alternating β-strands-β-turns, in the central domain.

It is interesting to note that insect eggshells which do not exhibit a helicoidal architecture contain proteins with a tripartite structure (e.g., *D. melanogaster*; *Ceratitis capitata*; Hamodrakas et al. 1989; Aggeli et al. 1991). A tripartite structure is also a characteristic feature of several other fibrous proteins, e.g., avian — feather and scale — keratins (Fraser and McRae 1973, 1976; Gregg et al. 1984).

On the basis of partial sequencing data, Regier et al. (1978b) pointed out that, as in avian keratin, in silkmoth chorion proteins, cysteines are preferentially localized near the two ends of the polypeptide chain. The complete chorion sequences confirmed this feature and the tripartite structure which emerged after secondary structure prediction further emphasizes the analogy with avian keratins.

In both feather and scale keratin, a fiber-matrix texture can be detected under the electron microscope when appropriate stains are used. Amino- and carboxyl-terminal ends, which are rich in cysteine, are thought to provide the stain-absorbing matrix, whereas it was suggested that the regular β-sheet central, conservative portion of keratin constitutes the fibrils (Fraser and McRae 1976).

In analogy with avian keratins, we proposed that the central, highly structured, evolutionarily conserved domains of silkmoth chorion proteins constitute the fibrils and the cysteine-rich protein "arms" the matrix seen in electron micrographs of silkmoth chorion (Hamodrakas et al. 1982a). This proposal appeared to be valid for a number of years and yet might finally prove to be correct during several developmental stages of choriogenesis. However, it should be slightly revised to take into account novel information, provided by X-ray diffraction, freeze-fracturing (Hamodrakas et al. 1986) and protein model-building studies (Hamodrakas et al. 1988) (see also Sects. 2.2.1 and 5).

The structure of silkmoth chorion protein "arms" deserves special attention. Protein arms were for a long time thought to be less structured than the central domain, as prediction results indicate (Hamodrakas et al. 1982a). However, since they contain tandem repeats and prediction methods fail in such cases, this interpretation may be entirely erroneous! On the contrary, the periodical structure of the peptides constituting the arms (see, e.g., Lecanidou et al. 1986) strongly favors a regular type of structure. The question, of course, is: what type of structure? To provide a plausible answer, let us examine briefly the primary structure features of the "arms".

Thus, most members of the abundant A and B families contain reduplications of Gly-Cys (G-C) repeats which are thought to serve for cross-linking chorion proteins via disulfide bonds. These are predicted mostly as β-strands. The alternation of "small" (G) and "bulky" (C) residues is reminiscent of the β-sheet structure in silk fibroin (Marsh et al. 1955), with small and bulky residues pointing to opposite sides of the β-sheet. In these and other parts of the molecules (even in the central domain), especially in sequences predicted as β-sheet strands, relatively small residues (G, A, or T) tend to alternate with bulky residues (V, L, I, Y). This might be important in chorion for the packing of successive β-sheets (Hamodrakas et al. 1982a).

In both the left and right arms, "late" *B. mori* HcA and HcB sequences are highly similar, consist almost entirely of glycine and cysteine, and contain almost exclusively G-G-C and G-C subrepeats which form a $(GGC)_2$-$(GC)_2$ major repeat (Burke and Eickbush 1986).

The "early" CA and CB protein "arms" are proline-rich and cysteine-poor (Lecanidou et al. 1986).

However, the most striking feature in the arms of the abundant A and B proteins is the existence of tandem repeats of the pentapeptide GYGGL (or LGYGG) or its variants (Regier and Kafatos 1985). The repeat array is frequently predicted as β-sheet, although an alternative prediction of a series of β-turns GYGG separated by a single residue (L) is possible (Hamodrakas et al. 1982a). In scale keratin, a repetitive sequence of the type $(GGX)_n$ (where X = F, L or Y) exists (Walker and Bridgen 1976) and it is thought to form β-sheet structure, possibly facilitating the characteristic hexagonal packing of scale keratin (Stewart 1977). Therefore, it appears that, the fiber-matrix systems of feather and scale keratins are analogous to silkmoth chorion at all structural levels: primary and secondary structure of protein components and also ultrastructure (Hamodrakas et al. 1986). For historical reasons, we should perhaps mention at this point that the term "ichthulokeratin" was introduced by Young and Inman (1938) to describe the eggshell proteins of Salmon eggs.

The protein sequence data banks NBRF and SWISSPROT were searched for the possible occurrence in other proteins of the tandem motifs (e.g., LGYGG) appearing in silkmoth chorion protein arms, utilizing the commands SCAN and MATCH of the Protein Sequence Query (PSQ) program. This is the main access

and retrieval program of the Protein Identification Resource (PIR; Orcutt et al. 1983). The findings of this search were unexpected and require further investigation: the pentapeptide LGYGG and its variants are constituents of proteins (or portions of proteins) exhibiting a fiber-matrix texture of ultrastructure. Thus, it is found repeatedly at the carboxyl-terminal end of scale keratin, which consists of 4×13 amino acid residue repeats, most probably adopting a characteristic antiparallel β-sheet structure comprising the matrix (Gregg et al. 1984), and also in the amino-terminal sequences of several cuticular matrix proteins from the locust *Locusta migratoria* (Hojrup et al. 1986). Insect cuticle is an extracellular layer, exhibiting a helicoidal architecture with a fiber-matrix texture, surrounding the whole animal. The cuticle consists of fibers of chitin embedded in a protein matrix (Neville 1975).

Furthermore, a study of the literature shows that the protein oothecin, secreted by the left colleterial gland of the cockroach *Periplaneta americana* to form the eggcase, consists, almost entirely, of tandem repeats of this pentapeptide (Pau 1984). Similar peptides appear also in the sequence of *Schistosoma mansoni* eggshell proteins and are predicted to form a β-sheet type of structure (Vanderlei et al. 1989) and also in the β-sheet structure of spider silk as well (Xu and Lewis 1990). Homologies of small portions of these proteins are summarized in Table 1.

In all the proteins (or protein segments) of Table 1, the peptide LGYGG and its variants, apparently, adopt a β-sheet type of structure. Its appearance in the structure of proteins which exhibit a fiber-matrix texture most probably indicates that this peptide is well tailored to play its role in the matrix of fiber-matrix systems: it may be a common architectural feature of matrix, or, if the alternative interpretation of fiber and matrix is taken (cf. Sects. 2.2.1 and 4.2), its presence perhaps ensures a better packing of the system components. To gain further insight, we performed a comparison of silkmoth chorion proteins with scale keratins. A sample of such a comparison performed with the FASTP program of Lipman and Pearson (1985) is shown in Fig. 23. To our surprise, it was found that the abundant silkmoth chorion protein Bpc401 (JBA041) shares a 41% similarity with chicken scale keratin (KRCHS) over a segment of 87 residues. This relatively high homology awaits further interpretation, in terms of both structure and evolution.

Summarizing the evidence presented above, we propose that the predominant secondary structure of chorion protein "arms" is that of β-sheet.

Table 1. Homologies between segments of proteins containing the pentapeptide GYGGL. Gaps necessary for alignment are represented with dashes. References are given in the text

Protein	Sequence segment
Silkmoth chorion Bpc401	. . . G L G Y G G – Y G L G G . . . .
Silkmoth chorion Bpc401	. . . G L G Y G G L – G Y G G . . . .
Silkmoth chorion Apc18	. . . G L G Y G G L – G Y G G . . . .
Scale keratin	. . . S L G Y G G L Y G Y G G . . . .
Locust cuticle Pr37	. . . G L G Y G G – Y G Y G . . . . .
Spider silk	. . . Q G G Y G G L G G Q G . . . . .
Oothecin	. . . G L G Y G G L – G Y G G . . . .


```
          80        90        100       110       120       130
JBA041    GLGVASENMYEGCVGVAGNLPFLGTAGVEGYFPTAGAGVINYGCGDSAVGITSEGG
                   ...:  ... :.:.....  .. : : . :::.....: ..  :
KRCHS     FPGPILSSFPQDSVVGSSGAPIFGGSSLGYGGSSLGYGGL-YGYGGSSLGYGGLYG
            50        60        70        80        90        100

               140       150       160       170
JBA041    YGG--LGYGGLGYEGVGGYGLGYGG-YGLG-GCGCG-SGRYL
          :::  :::::: : : :: .::::: :: : ..:.: :.:.
KRCHS     YGGSSLGYGGL-Y-GYGGSSLGYGGLYGYGRSYGSGYSSPYSYRYNRYRRGSCGPC
                  110       120       130       140       150
```

Fig. 23. Sequence similarities between silkmoth chorion protein Bpc401 (*JBA*041) and chicken scale keratin (*KRCHS*). The alignment optimized by the program FASTP (Lipman and Pearson 1985) is denoted by a *colon* for an identity and a *dot* for a conservative replacement. Insertions are marked with a *dash*. A z value of 3.69 standard deviations above mean was calculated for the initial score and a z value of 4.88 standard deviations above mean was calculated for the aligned score with the program RDF (Lipman and Pearson 1985). Fifty random sequences were generated. z values of initial and optimized scores >3 are considered as possibly significant, whereas z values >6 are probably significant according to Lipman and Pearson (1985)

Having predicted the secondary structure of silkmoth chorion proteins, experimental evidence was necessary to verify predictions. Experimental studies of eggshell protein structure were not confined, however, to silkmoths. Since they do not require previous knowledge of protein primary structure, they were extended to include helicoidal eggshells of other Lepidoptera (*M. sexta* and *S. nonagrioides*) species and of fish (*S. gairdneri*) as well. We shall briefly describe below the results of these studies.

4.2 X-Ray Diffraction Studies of Eggshells

X-ray diffraction studies of fibrous proteins usually provide information about their secondary structure and the modes of orientation and packing of protein molecules; frequently, these studies lead to the proposal of a protein structural model (Fraser and McRae 1973). However, samples should fulfill certain requirements: they should have intrinsic high order and diffract X-rays strongly to give interpretable X-ray diffraction patterns. Our experience showed that, unfortunately, both requirements are not easily met by the proteinaceous eggshells; it proved to be difficult to obtain useful X-ray diffraction patterns (Hamodrakas et al. 1983, 1986). Samples used for X-ray diffraction experiments were either hemispherical half-chorions, or stacked arrays of almost flat chorion fragments. High and low angle diffraction patterns were taken with the geometries shown in Fig. 24. Full details of the experimental procedures are given in the original publications.

In the "in-plane geometry" (Fig. 24a) the beam was parallel to the chorion surface and the planes of the stacked fibrils. In this case, the vertical axis of the diffraction pattern (meridian) corresponded to the radial axis of the chorion, i.e., revealed order along the axis from the inner to the outer surface; the horizontal axis (equator) of the diffraction pattern corresponded to the lateral axis of the chorion, i.e., revealed order within planes parallel to the surface. In the perpendicu-

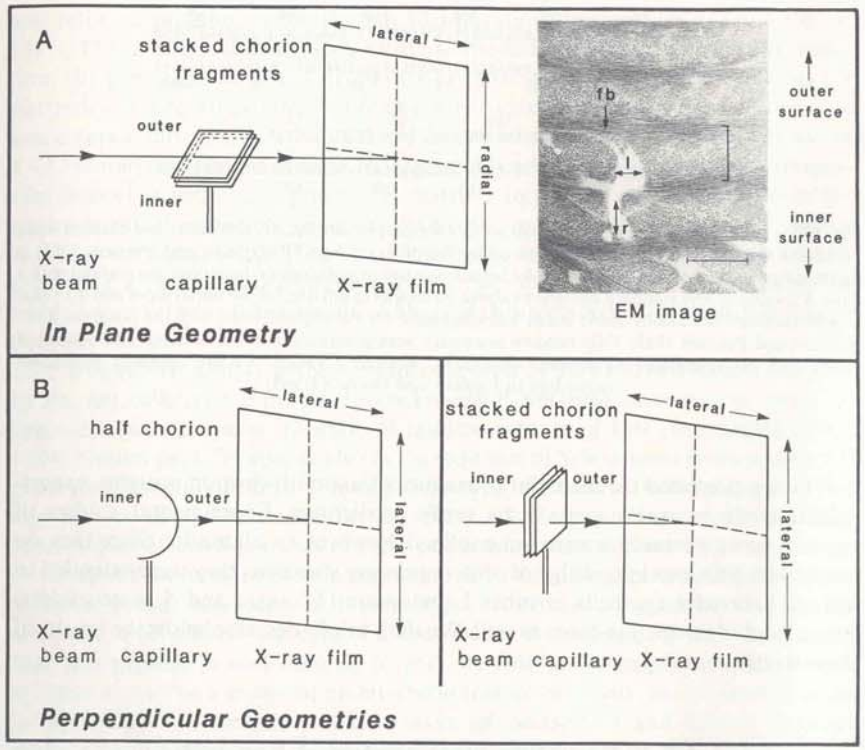


Fig. 24A, B. Geometries employed in the high angle X-ray diffraction experiment. In **A**, a stacked array of almost flat fragments, obtained from the hollow chorion sphere, is irradiated with the beam in the plane of the fragments, i.e., parallel to the outer and inner surfaces of the chorion. Note that the horizontal axis of the film is parallel to the fragments and thus reveals lateral periodicities, whereas the vertical axis of the film reveals radial periodicities (along the axis between inner and outer chorion surfaces). The electron micrograph is from an immature chorion and corresponds to the cut face that would be encountered by the beam. Fibrils (*fb*) are seen in orientations that vary with the plane; the *bracket* outlines one lamella (180°) rotation in fibril orientation), and the lateral (*l*) and radial (*r*) axes are indicated, as are the outer and inner chorion surfaces. *Bar* 500 nm. In **B**, the beam is oriented along the radial axis of the chorion hemisphere, or of stacked chorion fragments; consequently, both the horizontal and the vertical axes of the film reveal lateral periodicities. (Hamodrakas et al. 1983)

lar geometry (Fig. 24b) the beam was perpendicular to both the chorion surfaces and the planes of the stacked fibrils. In this case, both the vertical and the horizontal axes of the diffraction pattern revealed lateral order, within these planes.

In both orientations the high angle diffraction patterns had certain features in common (Fig. 25a,b). However, asymmetrical texture (scattering in an ellipse) was only observed in the in-plane geometry (Figs. 25a, 26, 27), indicating preferential orientation of the molecular chains. According to Kakudo and Kasai (1972, p. 404), scattering in an ellipse denotes scattering by stacked lamellar crystals with

cylindrically symmetrical orientation, which is (locally) the case of helicoidal structure. Therefore, X-ray diffraction provides a direct proof of the existence of a helicoidal architecture.

Predictions of secondary structure based on known sequences of major silkmoth chorion proteins led us to the suggestion that β-sheet structure is predominant in the silkmoth chorion. The preponderance of β-sheet structure was also supported by the results of laser-Raman and infrared studies (Hamodrakas et al. 1982b, 1984, 1987; cf. Sect. 4.3). The X-ray diffraction studies further support this notion.

Both perpendicular and in-plane geometries yielded broad, nearly uniform rings, centred at 0.464 nm. This spacing is typical of antiparallel β-sheet structures

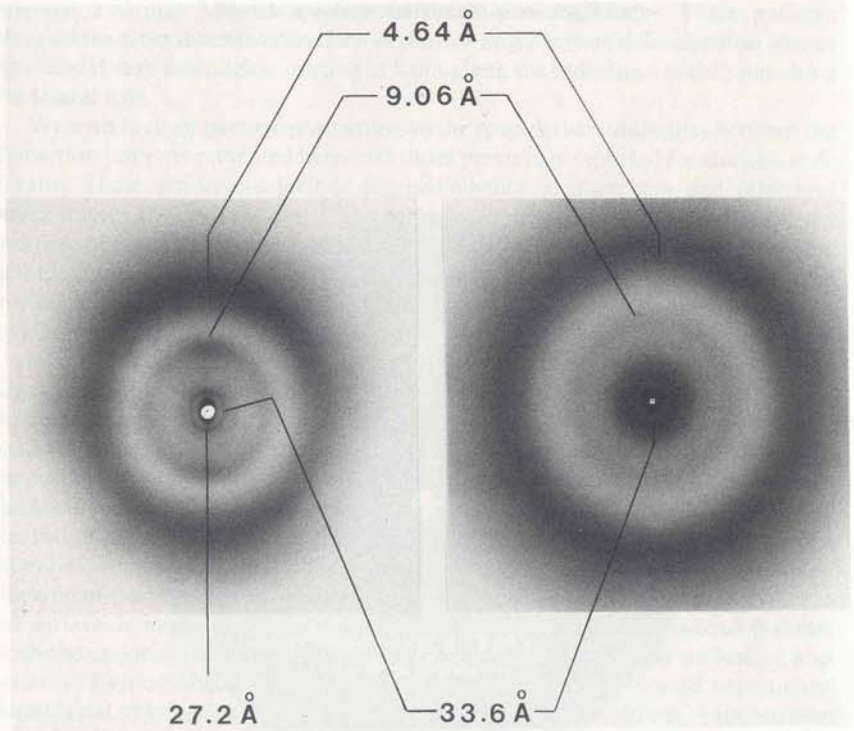


Fig. 25. High-angle X-ray diffraction patterns from mature silkmoth chorions of *A. polyphemus*. Pattern (*left*) is from the in-plane geometry (cf. Fig. 24a) and (*right*) from the perpendicular geometry (cf. Fig. 24b; half chorion). Note the presence of 0.464 nm, 0.906 nm, and 3.36 nm reflections which form rings in the perpendicular geometry. In the in-plane geometry, the same reflections occur but the 0.906 nm repeat is stronger and predominantly oriented along the radial axis. A 2.72 nm radial repeat is also seen, whereas the 3.36 nm repeat is confined to the lateral axis. (Hamodrakas et al. 1983)

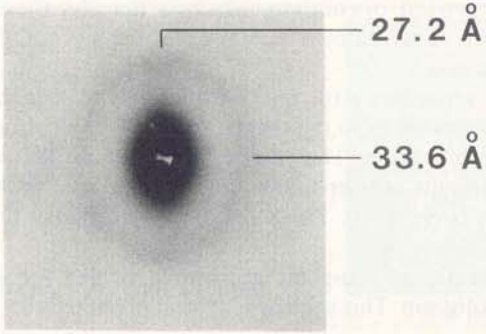


Fig. 26. Low-angle X-ray diffraction pattern from mature silkmoth chorion of *A. polyphemus*. Elliptical scattering, obtained using the in-plane geometry, suggests a helicoidal organisation (see text). This pattern confirms the results seen in the low-angle region of Fig. 25a; there is a 2.72 nm radial repeat and a 3.36 nm lateral repeat. (Hamodrakas et al. 1983)

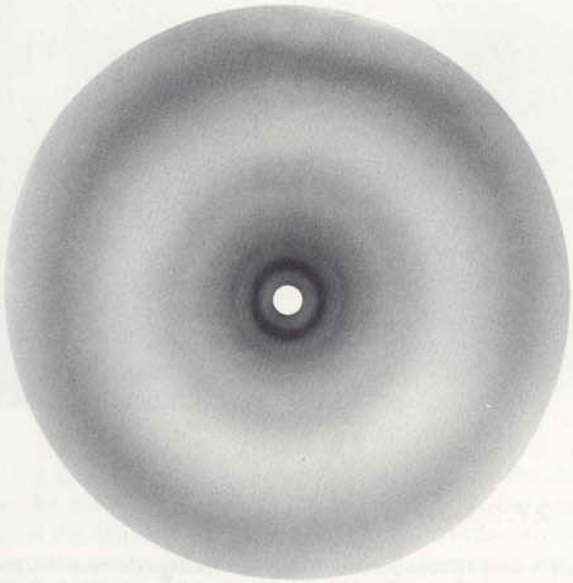


Fig. 27. X-ray diffraction pattern from mature silkmoth chorion of *A. polyphemus*. Incident beam parallel to the chorion surface which is horizontal (in-plane geometry). The plane of the film is vertical. The chorion sample was treated with 2% uranyl acetate for 60 min before irradiation. The concomitant significant increase in the intensity of the ca. 3 nm reflection can be seen. A toroidal camera was employed. Specimen to film distance was approximately 70 mm. (Hamodrakas et al. 1986)

corresponding to the interchain distance, between hydrogen bonded peptide chains of each sheet. A second ring was observed, at a 0.906-nm spacing, which may be attributed to the intersheet packing distance between regularly packed adjacent β-sheets. In the perpendicular geometry this ring was uniform and relatively weak (Fig. 25b). By contrast, in the in-plane geometry it included strong and broad reflections along the radial axis (Fig. 25a).

The X-ray diffraction patterns do not contain a ca. 0.54 nm reflection which could be representative of α-helix, confirming the suggested paucity of α-helical structure in the chorion (cf. 4.1). A third ring corresponded to spacings of ca. 3 nm. In the perpendicular geometry it was circular and corresponded to 3.36 nm. In the in-plane geometry it was oblong (elliptical), and in some patterns it was clearly resolved into a strong 2.72 nm reflection along the radial axis and a weaker 3.36 nm reflection along the lateral axis (Fig. 25a).

Low-angle X-ray diffraction patterns (Fig. 26) confirmed the existence of oriented 2.72 and 3.36 nm spacings in the in-plane geometry. These patterns revealed no other discrete reflections in the low angle region; diffuse central scatter was seen at very low angles, starting at 8 nm along the radial axis and 11 nm along the lateral axis.

We wish to draw particular attention to the remarkable similarities between the diffraction patterns presented here, and those previously reported for chicken scale keratin. These similarities include the prominence of interchain and intersheet reflections (0.47 and 0.94 nm^{-1} respectively, in the case of scale keratin); the presence of oriented reflections in the in-plane but not the perpendicular geometry; and the presence of oriented reflections in the vicinity of 3 nm^{-1}, in both high and low angle X-ray diffraction patterns obtained using the in-plane geometry (2, 2.3 and 3.5 nm^{-1} in the case of scale keratin).

When considered together with the electron microscopic evidence (Fig. 24a), the X-ray diffraction patterns suggest substantial orientation of the β-sheets relative to the fibrils. Since the 0.906 nm^{-1} reflections are most prominent along the radial axis (in-plane geometry), and weak along the lateral axis (both in-plane and perpendicular geometries), it would appear that β-sheets tend to be stacked across the fibril, occupying planes parallel to the chorion surface (rather than say, being stacked along the long axis of the fibril). The weak 0.906 nm^{-1} reflections in the lateral axes could be ascribed either to the disrupted and oblique lamellae, or to stacking of β-sheets in nonradial orientations even in fibrils which are parallel to the surface. It might also arise if the β-sheets are not "flat", but twisted β-sheets (with the majority of their strands more or less parallel to chorion surface; cf. also Sect. 5). Twisted sheets and disrupted or oblique lamellae would explain the paradoxical observation that in the in-plane geometry the 0.464 nm^{-1} (interchain) reflection is not preferentially observed at a 90° angle relative to the 0.906 nm^{-1} (intersheet reflection) and/or nonstacked sheets or sheets stacked in various orientations. Oriented 0.94 nm^{-1} and unoriented 0.47 nm^{-1} reflections are also observed in the in-plane patterns derived from scale keratin [see Fig. 2b in Stewart (1977)].

Prominent reflections at ca. 3 nm^{-1} are also observed in the chorion patterns. In the perpendicular geometry they are unoriented and correspond to 3.36 nm,

whereas in the in-plane geometry oriented 3.36 nm and stronger, 2.72 nm reflections are observed. One possible explanation for these reflections is that the 8–20 nm fibers seen by electron microscopy are aggregates of oriented ca. 3 nm filaments. The shape and stacking of these filaments would be asymmetrical in cross-section, resulting in the 2.72 nm radial and 3.36 nm lateral periodicities. Similar filaments, measuring ca. 3 nm in diameter, and spaced at distances of 3 to 3.5 nm, have been seen in both feather rachis and scale keratin (Stewart 1977).

To examine whether the ca. $1/3\,nm^{-1}$ reflections appearing in the X-ray diagrams correspond to periodicities between fibrous elements of chorion, a simple experiment was carried out, similar to that of Fraser and McRae (1959) in studies of feather keratin: the silkmoth chorion was treated with 2% osmium tetroxide or 2% uranyl acetate, which results in considerable deposition of osmium or uranium within the structure, as evidenced by the brown and yellow coloration, produced respectively. The high angle diffraction pattern is unaffected by these treatments and it may be concluded as in feather keratin, that deposition of osmium or uranium occurs mainly between, rather than within, the fibrous elements of the structure. At low angles, the principal effect is a dramatic intensification of the $1/3\,nm^{-1}$ reflection (Fig. 27) which indicates periodic fluctuations of considerable amplitude in the density of osmium and uranium in the structure, of the same order of magnitude.

If the $1/3\,nm^{-1}$ reflection is considered as indicating packing distances between chorion filaments, then the dramatic increase in the intensity of this reflection may be easily explained, since the contrast between the filaments and the matrix is considerably increased, if osmium tetroxide or uranyl acetate bind to the matrix intervening between the filaments. The concept that osmium tetroxide or uranyl acetate bind preferentially to the matrix is supported by the proposal put forward by Hamodrakas et al. (1982a), that the "less structured", high in cysteine content, variable arms of chorion proteins constitute the matrix. The high affinity of osmium tetroxide and uranyl acetate to cysteine is well known. However, an alternative explanation is that the filaments constituting chorion are formed by the hydrophobic cores of the folded proteins and that the matrix corresponds to the hydrophilic exterior of the proteins, with osmium tetroxide and uranyl acetate binding to the polar groups of the protein surface; groups of the protein surface include the side chains of glutamate, aspartate, and cysteinyl residues (Hamodrakas et al. 1985) to which uranyl acetate and osmium tetroxide preferentially attach (Blundell and Johnson 1976).

Supporting and conclusive evidence for the existence of ca. 3–4 nm fibrils in silkmoth chorion was provided by freeze fracturing studies (Hamodrakas et al. 1986; cf. Sect. 2.2.1) and recent transmission electron microscopy and optical diffraction data (Hamodrakas SJ and Ottensmeyer FP, in prep.).

Similar, but very weak X-ray diffraction patterns were obtained also from eggshells of the fish *S. gairdneri* (Papadopoulou P and Hamodrakas SJ, unpubl.) and *M. sexta* (Orfanidou C and Hamodrakas SJ, unpubl.).

In conclusion, X-ray diffraction studies confirm the prevalence of β-sheet structure in the chorion, reveal a degree of regular orientation (preferred packing) of β-sheets relative to its surface, provide a direct proof of its helicoidal architecture and suggest the existence of fibrils ca. 3–4 nm in diameter as its basic structural elements.

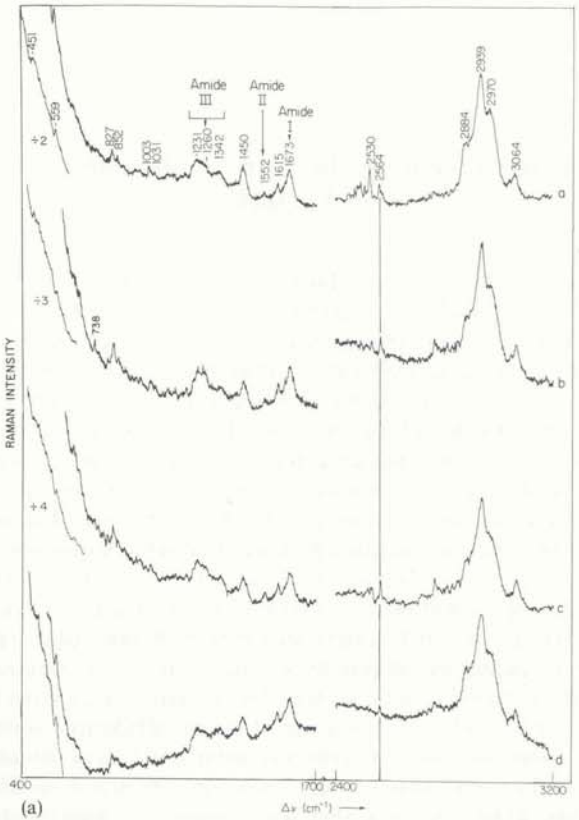
4.3 Laser-Raman and Infrared Spectroscopic Studies of Eggshells

Laser-Raman spectroscopy is a relatively new technique for examining changes in the frequency of emitted radiation due to molecular vibrations. It is related to, but distinct from, infrared (IR) spectroscopy, which examines absorption of radiation due to the same molecular vibrations. Both techniques have been empirically demonstrated to be quite sensitive to protein conformation; they are very powerful and reliable techniques for the determination of protein secondary structure. Laser-Raman spectroscopy can also provide a wealth of useful information about the state of certain amino acid residues in proteins and protein systems, thus offering distinct advantages compared to IR spectroscopy. The weak emission as opposed to absorption of radiation by water is another important advantage of Raman spectroscopy for studies of biological materials. (Parker 1971; Frushour and Koenig 1975; Spiro and Gaber 1977; Yu 1977; Carey 1982).

The eggshell is a relatively favorable structure for Raman spectroscopy studies. It consists almost exclusively of protein and functions as an essentially dry shell; thus it meets the criterion of high protein density which is required for adequate sensitivity in Raman studies. The absence of major admixtures, such as chitin or other carbohydrates, minimizes interference from other than protein vibrations. Furthermore, we have encountered only limited interference from fluorescence in our studies. These features have permitted analysis of the intact eggshell structure, as opposed to protein extracts, ensuring that the structural features observed reflect a physiological state.

Serious difficulties were encountered in our initial attempts to obtain eggshell samples [in the form of KBr pellets, containing about 2% (w/w) material, which was thoroughly ground in a vibrating mill before mixing with KBr] suitable for IR spectroscopy. This was due to the toughness, relative opacity and elasticity of the eggshells. Nowadays, however, useful IR spectra from eggshells are obtained in an almost routine manner.

To date, Raman spectra have been obtained from eggshells of *A. polyphemus* (Hamodrakas et al. 1982b), *B. mori* (Hamodrakas et al. 1984), *A. pernyi* (Hamodrakas SJ and Petrou A, unpubl.), and *S. gairdneri* (Hamodrakas et al. 1987; Papadopoulou P, Kamitsos EI, and Hamodrakas SJ, in prep.) and IR spectra from *A. polyphemus*, *B. mori*, *A. peryi*, *S. gairdneri*, *M. sexta*, and *S. nonagrioides* (Orfanidou C and Hamodrakas SJ, in prep.). Representative laser-Raman spectra from *A. polyphemus*, *B. mori*, and *S. gairdneri* eggshells are shown in Fig. 28a, b, and c, whereas Fig. 29 shows a representative IR spectrum from *S. gairdneri* (Hamodrakas et al. 1982b, 1984, 1987).



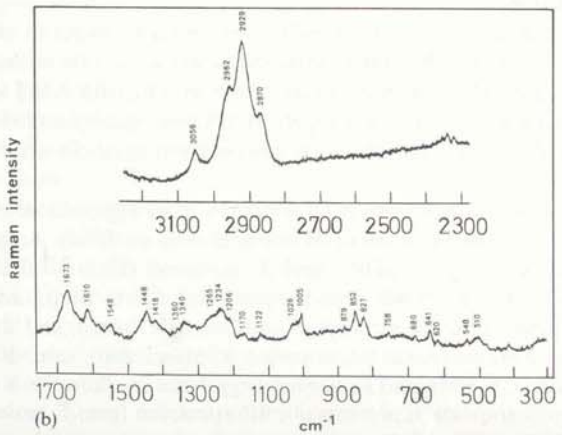


Fig. 28a, b

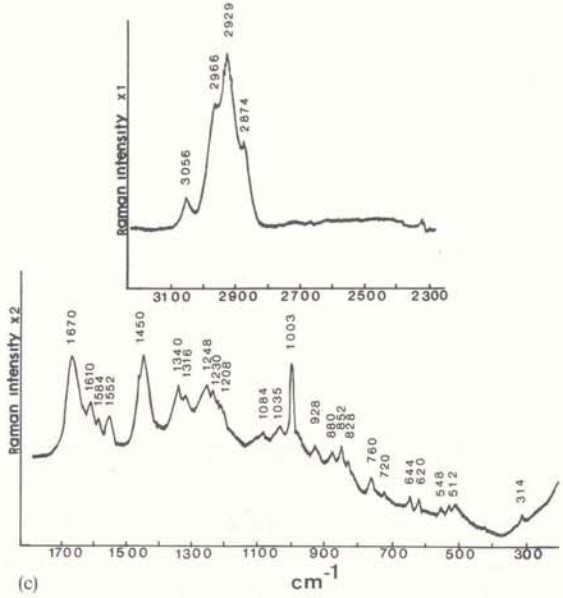


Fig. 28 a. Laser Raman spectra of *A. polyphemus* chorion. Early developing chorion (*a*) or ovulated chorion (*b–d*) samples were used. *a–c* were examined tangentially to the surface, and *d* in the radial orientation along a cut edge. In *c*, the inner structure was examined tangentially, after scraping off the outer layers of the chorion. Several Raman bands are identified, including those discussed in the text. A change in scale was necessary for plotting the low end of the spectrum in *a*, *b*, and *c*. Instrumental conditions: excitation wavelength = 5145 Å; average slit width = 5.2 cm^{-1}; integration time = 1 s; laser power at the sample = 100 mW. (Hamodrakas et al. 1982b). **b** Laser Raman spectrum of the eggshell of the silkmoth *B. mori*. A 90° scattering geometry was employed with the laser beam hitting the eggshell surface tangentially (outer lamellar layer). Instrumental conditions: excitation wavelength 514.5 nm; scanning speed = 10 cm^{-1} min^{-1}; time constant = 2 s; spectral resolution 5 cm^{-1}; laser power at the sample = 100 mW. (Hamodrakas et al. 1984). **c** Laser Raman spectrum of the eggshell of the fish *S. gairdneri*. A 90° scattering geometry was employed, with the laser beam hitting the eggshell surface tangentially. Instrumental conditions: excitation wavelength, 514.5 nm; scanning speed, 10 cm^{-1} min^{-1}; time constant, 2 s; spectral resolution, 5 cm^{-1}; laser power at the sample, 100 mW (Hamodrakas et al. 1987)

4.3.1 Raman and Infrared Spectra: Secondary Structure of Eggshell Proteins

According to extensive theoretical and experimental studies, the frequencies of amide I, II, and III bands in Raman spectra are useful indicators of protein and polypeptide secondary structure (Carey 1982). The amide I bands (which occur in the region 1630–1690 cm^{-1}) appear to have contributions from C–O stretching (approximately 70%) and C–N stretching (approximately 16%). The generally weak amide II Raman bands (1510–1570 cm^{-1}), and the amide III bands (1220–1330 cm^{-1}) have significant contributions from N–H in plane bending and

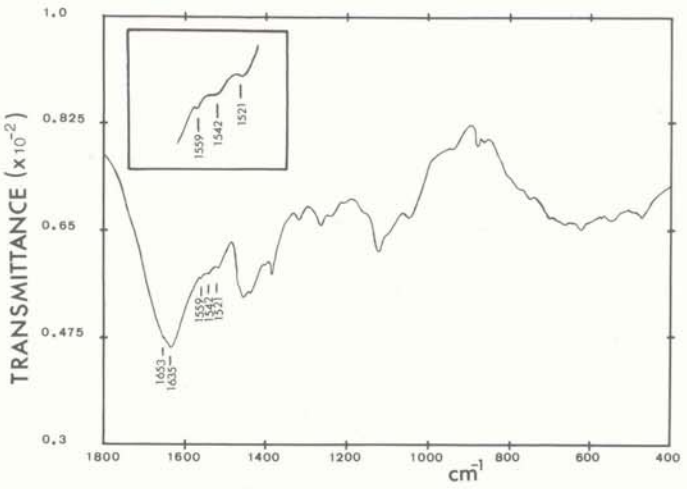


Fig. 29. Fourier transform infrared spectrum of the eggshell of the fish *S. gairdneri*. The spectrum is the result of signal averaging of 100 scans, at 2 cm^{-1} resolution. Samples were in the form of KBr pellets, containing about 2% wt. material, thoroughly ground in a vibrating mill, before mixing with KBr. *Inset* shows, on an expanded scale, the 1500–1600 cm^{-1} spectral region used for identifying the bands at 1521, 1542, and 1559 cm^{-1}. (Hamodrakas et al. 1987)

C–N stretching. Table 2 summarizes the diagnostic locations of these bands for α-helical, β-sheet, and β-turn structures, and lists the corresponding frequencies observed in laser Raman spectra of *A. polyphemus*, *B. mori*, and *S. gairdneri* chorion samples (Fig. 28a, b, c).

Table 3 gives the frequencies and our tentative assignments of the bands appearing in the laser-Raman (Fig. 28b) spectrum of the *B. mori* eggshell. Additional bands are resolved but not tabulated, because insufficient data are available for unambiguous assignments. The IR spectra exhibit a wealth of information. However, in such complicated proteinaceous systems it is difficult to assign all absorption bands to certain vibrations of defined chemical groups. Therefore we limit our attention to identifying amide I, II, and III bands suggesting a certain type of secondary structure.

The Raman spectra clearly indicate that the β-sheet conformation is predominant in proteinaceous eggshells (Fig. 28; Tables 2, 3): The bands at ca. 1673 cm^{-1} (*A. polyphemus*), 1673 cm^{-1} (*B. mori*), 1670 cm^{-1} (*S. gairdneri*) [amide I] and 1231 cm^{-1} (*A. polyphemus*), 1234 cm^{-1} (*B. mori*), 1230 and 1248 cm^{-1} (*S. gairdneri*) [amide III], can best be interpreted as resulting from abundant antiparallel β-pleated sheet structure in eggshell proteins.

The amide I band of the spectra was analyzed following the method of Williams and Dunker (1981) to estimate the percentage (%) of secondary structure of eggshell proteins. The method is described fully by Hamodrakas et al. (1984). The

Table 2. Summary of diagnostic laser-Raman amide bands and their observation in proteinaceous eggshells (Fig. 28)

Band nature	Bands characteristic of			Observed in chorion
	α-helix	β-sheet	β-turn	
Amide I	1650 to 1660	1665 to 1680	1665, 1690	1673 *A. polyphemus* 1673 *B. mori* 1670 *S. gairdneri*
Amide II	1516, 1545	1535, 1560	(I) 1550 to 1555, 1567 (II) 1545, 1555, 1560	
Amide III	1260 to 1290	1230 to 1240	1290 to 1330	1231 *A. polyphemus* 1234 *B. mori* 1230, 1248 *S. gairdneri*

Table 3. Wavenumbers and tentative assignments of bands in the laser-Raman spectrum of the eggshell of *B. mori* (Fig. 28b)

Wavenumber (cm^{-1})	Tentative assignment
510	S–S stretch
540	S–S stretch
620	Phe
641	Tyr
680	C–S stretch? Trp?
758	C–S stretch? Trp?
827[a](+)	Tyr
852[a](+)	Tyr
879	Trp
1005[a](+)	Phe or C–C stretch (β-sheet)
1016	Phe-Trp
1026	Phe
1122	C–N stretch
1170	Tyr
1206	Tyr, Phe
1234[a](+)	Amide III (antiparallel β-sheet)
1265[b](+)	Amide III (β-turns? cross-β? α-helix? coil?)
1340	Amide III (β-turns) or Trp
1360	Trp
1418	Trp
1448[a](+)	CH_2 deformation
1548	Amide II (β-turns) or Trp
1610	Tyr, Phe, Trp
1673[a](+)	Amide I (antiparallel β-sheet)
2800–3100	C–H stretch

[a] (+) A strong peak.
[b] (sh) A shoulder.

analysis suggests that the proteins of silkmoth chorion consist of 60–70% antiparallel β-pleated sheet and the remainder 30–40% of β-turns (a 2:1 ratio). For *S. gairdneri*, it is estimated that 50–60% is antiparallel β-pleated sheet, 30–40% β-turns and 10% α-helix (Papadopoulou P, Kamitsos EI, and Hamodrakas SJ, in prep.).

The distribution of the phi and psi angles in the β-sheets appears to be rather narrow (in other words the β-sheets exhibit a uniform structure), since the amide I band at ca. 1670 cm^{-1} is sharp: its half-width is approximately 40–45 cm^{-1} compared to the 27 cm^{-1} observed in the very uniform silk fibroin and the 76 cm^{-1} in the less uniform β-keratin (Frushour and Koenig 1975).

Supporting evidence for the prevalence of antiparallel β-pleated sheet in the proteins of helicoidal eggshells was supplied by Fourier transform IR spectroscopy: For the *S. gairdneri* eggshell proteins, the observation of a very intense absorption band at 1635 cm^{-1} (amide I) and of a weak band at 1521 cm^{-1} (amide II) in the IR spectrum (Fig. 29) strongly indicate an antiparallel β-pleated sheet conformation. Similar results were obtained from IR spectra of *A. polyphemus*, *B. mori*, *A. pernyi*, *M. sexta* and *S. nonagioides* (Orfanidou C and Hamodrakas SJ, in prep.). In at least two cases, *A. polyphemus* (Fig. 28a) and *S. gairdneri* (Papadopoulou P and Hamodrakas SJ, in prep.) a study of eggshell protein structure has been performed at different developmental stages, utilizing laser-Raman spectroscopy. The relative invariance, during development, of the bands indicative of antiparallel β-sheet suggests the preponderance of this structure in eggshell proteins throughout choriogenesis and is significant in terms of eggshell morphogenesis.

4.3.2 Raman Spectra: Side Chain Environments

Raman spectra yield useful information on amino acid residues of Cys, Tyr, Phe, and Trp (Carey 1982). We shall focus our attention here only on Cys and Tyr.

Bands in the 500–550 cm^{-1} region are typically associated with the S–S stretching mode of the C–C–S–S–C–C structural unit of disulfide bonds. Following Sugeta et al. (1972), bands at 510, 525 and 540 cm^{-1} may be assigned to S–S bridges in *g-g-g*, *g-g-t* and *t-g-t* (*g* and *t* denote *gauche* and *trans*) conformations, respectively. Bands in the 2530–2580 cm^{-1} spectral region are typically associated with the –S–H stretching mode (Yu 1977). Our studies on *A. polyphemus* eggshells have shown that the highly localized cysteines of chorion proteins are apparently found in diverse environments during chorion development, as can be judged by the presence of multiple bands, assigned to free sulfhydryls, in the 2530–2580 cm^{-1} region in developing eggshells and are cross-linked by disulfide bonds at/or near ovulation. This is strongly supported by the evident suppression and even disappearance of these bands in the ovulated samples (Fig. 28a; Hamodrakas et al. 1982b).

In the outer osmiophilic layer of a mature *B. mori* eggshell formed by the Hc proteins, the great majority of the cysteines are cross-linked by disulfide bonds in *g-g-g* and *t-g-t* conformations, as can be judged by the presence of bands at 510 and 540 cm^{-1} and the absence of significant features in the 2530–2580 cm^{-1} spectral region (Fig. 28b; Hamodrakas et al. 1984). These bonds harden and waterproof the eggshell during the unusually long diapause periods, which is essential for the oocyte survival (Kafatos et al. 1977).

Determination of the exact time of formation of disulfide bonds in *S. gairdneri* eggshells does not seem to be necessary; apparently, its cysteines are cross-linked via disulfide bridges throughout development (Papadopoulou P, Hamodrakas SJ, and Kamitsos EI, in prep.). All possible types of disulfide conformation exist in mature eggshells, as can be judged by the presence of bands at ca. 510, 525, 548 cm^{-1} (Fig 28c; Hamodrakas et al. 1987).

The intensity ratio of the tyrosine doublet at 850 and 830 cm^{-1}, $R = I_{850}/I_{830}$, is sensitive to the nature of hydrogen bonding, or to the state of ionization of the phenolic hydroxyl group: basically, if tyrosine functions as a strong hydrogen-bond donor to a negative acceptor in a hydrophobic environment, the ratio R is low, perhaps 0.3–0.5. Hydrogen bonding in which the phenolic oxygen serves as a weaker donor or as an acceptor yields a higher ratio (Siamwiza et al. 1975). In the *A. polyphemus* spectra (Fig. 28a), $R = 0.3 \pm 0.1$, suggesting that the tyrosines are buried in a hydrophobic environment and strongly hydrogen bonded. This is significant, since the tyrosine residues are highly localized within the "arms" of chorion proteins. In *B. mori* spectra (Fig. 28b), $R = 1.2 \pm 0.1$, which probably suggests that most tyrosines act as much weaker hydrogen bond donors or as acceptors. In *S. gairdneri* (Fig. 28c), $R = 2.2 \pm 0.1$, an unusually high value (Yu 1977).

Summarizing, all evidence to date, both theoretical and experimental, strongly suggests that uniform antiparallel β-pleated sheet is the predominant molecular conformation of helicoidal eggshell protein structure during development. Apparently, this conformation dictates protein self-assembly for the formation of higher order helicoidal architecture. The follicle cells (or the oocyte) do not play a direct role in the process, since self-assembly takes place extracellularly; (1) the proteins interact to form the three dimensional structure of the eggshell at some distance from the points of their secretion (Smith et al. 1971); (2) secreted protein molecules pass through the porous sieve layer (Fig. 16) to reach their destination: this ensures minimum follicle cell involvement in the self-assembly process, apart from secretion of multimers, which may act as "nuclei" in the self-assembly process. The question now arising is: how is self-assembly done in molecular terms? To answer this question, models of protein structure are needed. Obviously, our description will be confined to silkmoth chorion proteins with known primary structure.

5 Structural Motifs in Proteins of Helicoidal Eggshells and Protein Structural Models

Tandemly repeating peptides have been found in the sequences of most fibrous proteins and play an important role in the formation of the fibrous structure (Fraser and McRae 1973; Parry 1979). Individual repeat units tend to be conformationally equivalent. If the equivalence is exact, a helical structure results, if not, the local conformations of the repeat units are likely to be similar (Parry et al. 1979). An important question which always arises in such cases is what type of structure is formed by these repetitive peptides.

Our analysis clearly established that silkmoth chorion proteins have an unusual tripartite structure: a central domain, highly conserved within each family

and recognizably homologous between families of the same branch, and two variable flanking domains or "arms", marked by the presence of tandemly repetitive peptides (Figs. 18–22) that are not apparent in the central domain. The evolutionary conservation and length invariance of the central domain, in each branch (α or β) of chorion protein families, suggest that this domain assumes a precise and functionally important three-dimensional structure.

5.1 Structure of the Central Domain

The existence of tandemly repeating peptide motifs as a characteristic feature of fibrous protein structure and the observation of periodicities of β-sheet maxima in the central domain, alternating with β-turns, in secondary structure prediction histograms (cf. Sect. 4.1), led us to search for "hidden" periodical patterns of residues in the central domain.

5.1.1 Definition of the Sequences Considered

Figure 18 presents typical chorion sequences of the α (A, HcA, CA) and β (B, HcB, CB) branches and shows the borders of the central domains analyzed (Hamodrakas et al. 1985, 1988). The analysis has been performed on the sequences shown; our conclusions apply to all other available sequences, published and unpublished.

5.1.2 Hexad Periodicities

The region considered is highly conserved among subfamilies of the α and β branches, and has not undergone deletions or insertions over more 50 million years (Rodakis et al. 1982). For preliminary analysis of periodicities a Fortran program was written (Hamodrakas SJ, unpubl.), which revealed a sixfold periodicity for Gly. Accordingly, sequences were written out in rows of six residues (Figs. 30a, 31a) and the significance of the nonrandom distribution of residues in the six columns thus generated, was analyzed with the method of McLachlan (1977), calculating the pattern strength, P. This measure is the difference between observed and randomly expected unevenness in the distribution of amino acid residues over the columns, divided by the standard deviation. Therefore, values of P greater than 3.0 are highly significant. The analysis established clear sixfold periodicities for various types of residues (Table 4) which may be summarized by declaring that the periodically repeating hexapeptides have the general form:

α branch: Gly-X-large hydrophobic-Y-large hydrophobic-hydrophobic

β branch: Gly-X-large hydrophobic-Y-large hydrophobic-Z,

with X usually a β-turn former residue and Y, Z of a more general type.

Similar hexapeptide periodicities were detected by Fourier transform analysis of the sequences. Fourier transforms were obtained essentially as outlined by

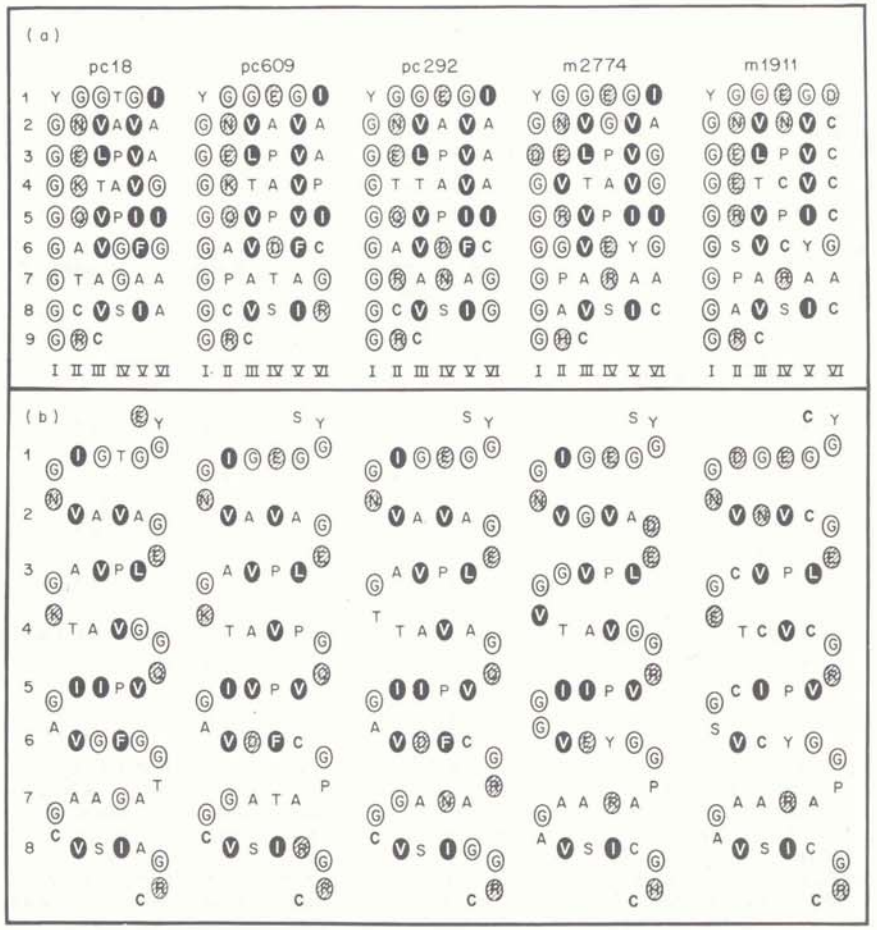


Fig. 30. a Regular amino acid distribution within the central domain of the A/HcA chorion proteins (cf. Fig. 18). To reveal the sixfold periodicities, sequences have been written in rows (numbered *1* to *9*) of six residues each; they should be read *left to right, top to bottom. Vertical columns* (numbered *I* to *VI*) can thus be seen to have nonrandom prevalence of certain types of residues, which are distinctively marked (see text). **b** Anti-parallel β-sheet model for the central domain. Sequences should be read continuously, beginning at the top. Each row of **a** corresponds to a vertical β-turn and a similarly numbered horizontal β-sheet strand. For further details, see text. (Hamodrakas et al. 1985)

McLachlan (1977) and McLachlan and Stewart (1976), using a Fortran 77 computer program: each sequence of N residues was represented as a linear array of N terms, with each term given a value of 1 or 0, according to whether the condition considered (e.g, presence of a Gly residue) was or was not satisfied. To increase resolution, this array was embedded in a larger array of zeros. A summary of the Fourier analysis results for the A/HcA proteins is shown in Table 4.

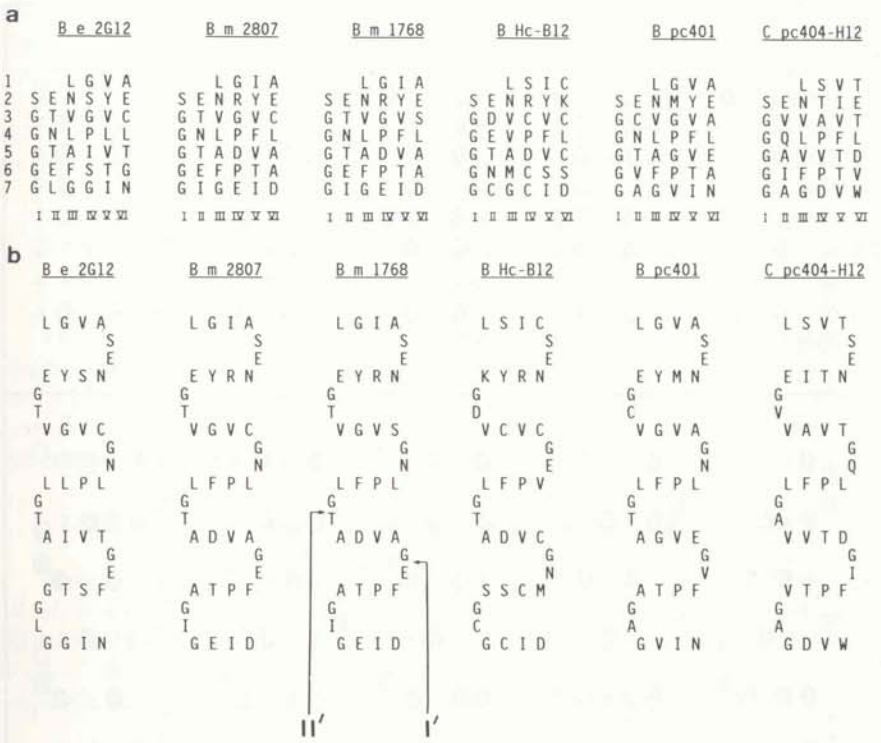


Fig. 31.a Regular amino acid distribution within the central domain of the B, HcB, CB silkmoth chorion proteins. To reveal the sixfold periodicities, sequences have been written in *rows* (numbered *1–7*) of six residues each; they should be read *left to right*, *top to bottom*. *Vertical columns* (numbered *I* to *VI*) can thus be seen to have nonrandom prevalence of certain types of residues. A distinct pattern of Gly-X-large hydrophobic-Y-large hydrophobic-Z emerges, where X is usually a β-turn former and Y, Z other types of residues. **b** Antiparallel β-sheet model for the portion of the central domain considered. Sequences should be read continuously, beginning at the top. Tentative I′ and II′ β-turns alternate with four-residue β strands. For further details, see text. (Hamodrakas et al. 1988)

5.1.3 Interpretation of Hexad Periodicities

The following arguments, described in detail by Hamodrakas et al. (1985), led us to interpret the hexad periodicities appearing in the central domains of the α and β branches of proteins, by the alternating β-turn/β-strand model of an antiparallel β-pleated sheet shown in Figs. 30b, 31b:

1. Secondary structure prediction indicates a regular alternation of β-sheet maxima with β-turns in the central domains of silkmoth chorion proteins (cf. Sect. 4.1).
2. Analysis of the amide I band of the laser-Raman spectra suggests a 2:1 ratio of antiparallel β-sheet/β-turns in chorion proteins (cf. Sect. 4.3).

Table 4. Residue periodicities in the central domain of silkmoth chorion A proteins

Pattern strength analysis

Type of residue	pc18	pc609	pc292	m2774	m1911
G	9.02	12.60	11.41	6.34	12.60
β-Turn[a]	6.62	6.70	6.17	4.86	9.46
V	4.07	4.82	4.07	2.97	4.07
β-Sheet[b]	7.19	6.11	6.48	5.97	4.94

Fourier transform analysis

Type of residue	pc18	pc609	pc292	m2774	m1911
G	3.48 (98°)	6.41 (113°)	7.61 (110°)	5.07 (114°)	6.60 (111°)
β-Turn[a]	3.11 (144°)	4.55 (131°)	4.49 (139°)	2.88 (115°)	3.31 (109°)
V	3.40 (−86°)	3.43 (−75°)	3.40 (−87°)	2.36 (−96°)	3.20 (−87°)
β-Sheet[b]	4.04 (−47°)	4.49 (−51°)	5.11 (−44°)	3.23 (−50°)	7.76 (−45°)

Top:
Pattern strength (P) values are shown for a periodicity of six residues; values are expressed in standard deviation units (σ).

Bottom:
Fourier transforms for a periodicity of 5.82 residues (a value giving consistent maxima at approximately six residues). Each entry includes a intensity (I) value and a phase angle in parentheses. Intensity values are scaled as recommended by McLachlan and Stewart (1976); the probability of observing by chance an intensity, I, at any particular periodicity is exp(−I).

[a] β-Turn formers, G, P, D, N, S, C, K, W, Y, Q, T, R, H, E; Chou and Fasman (1978).
[b] β-Sheet formers, V, L, I, F, W, Y, T, C; Chou and Fasman (1978).

3. Strong periodicities were observed for groups of β-sheet former and β-turn former residues with a period of approximately six (6) residues and a phase difference of approximately 180° (out-of-phase; Table 4).

In this model, the horizontal rows, each containing four residues, represent antiparallel β-sheet strands and consist of the residues shown in columns III to VI of Figs. 30a and 31a. The end residues of these short strands also participate in β-turns, together with the vertically displayed dipeptides (columns I and II of Figs. 30a and 31a; residues Gly-X in the hexapeptide motifs). The latter represent the central residues i + 1 and i + 2 of the β-turns, respectively.

The model is somewhat reminiscent of Silver Gull feather keratin (Fraser and McRae 1976), where an eight-residue periodicity in β-sheet propensities and a similar but out-of-phase periodicity of random coil propensities reflects the structure of a β-sheet consisting of short antiparallel strands. Similarly, the model is reminiscent of the cross-β-sheet structure in the shaft of the adenovirus fiber protein (Green et al. 1983); in that case, the antiparallel β-sheet strands are usually five or six residues long and together with two β-turns that punctuate them, result in an observed 15-residue periodicity.

5.1.4 β-Turn Type Determination and Modeling

To construct realistic models of the proposed antiparallel β-pleated structure for all chorion proteins, it was necessary to determine the types of the β-turns (Chou and Fasman 1977) which give the silkmoth chorion protein central conservative domains their characteristic fold (Figs. 30b, 31b).

Since the conformation of residues i, i + 3 is constrained to have values corresponding to those of an antiparallel β-sheet (these residues are parts of the four-residue β-strands), this means that we had to determine the conformation of residues i + 1, i + 2, of each β-turn.

A similar analysis has been performed by Geddes et al. (1968) to determine the types of β-turns which generate the cross-β conformation in proteins. These authors arrived at two types of turns to be the most favorable for the formation of the cross-β conformation, which they denoted as fold A and fold B. They essentially correspond to the II′ and II types of Chou and Fasman (1977), respectively. In these folds, a hydrogen bond is formed between the –NH group of residue i + 3 with the –C=O group of residue i.

Our analysis, which in general terms resembled the analysis of Geddes et al. (1968), but was carried out with our own Fortran program, was done as follows:

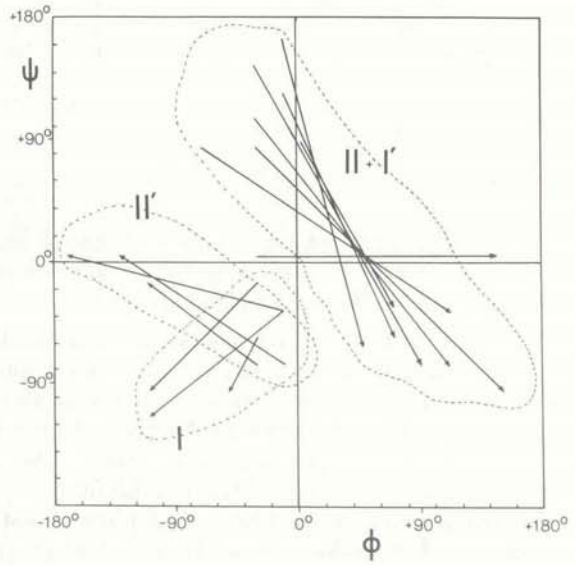


Fig. 32. Phi-Psi plot showing the conformation of the turn residues i + 1, i + 2 by *arrows*, as were found by our conformational analysis of the decapeptide VAVAGELPVA, as described in the text. The *beginning of each arrow* marks residue i + 1 and the end, residue i + 2. Three groups of possible β-turns were found: group 1, which corresponds to distorted types of II and I′ turns (Chou and Fasman 1977), group 2, corresponding to distorted types of II′ turns, and group 3, which corresponds to distorted types of III β-turns. Presumably, these distorted turns are necessary in the model because the phi, psi angles of the sheet residues were kept constant. (Hamodrakas et al. 1988)

1. Representative decapeptides of the type forming the characteristic structure of Figs. 30b, 31b were chosen, which contain two consecutive β-strands linked by a β-turn.
2. Backbone dihedral angles of the β-strand residues were set to phi = $-120°$ and psi = $+135°$. These values correspond to the centre of the allowed region for β-pleated sheets and they generate β-strands suitable for the formation of twisted β-pleated sheets (Schulz and Schirmer 1978). We have chosen these values since most known β-sheets, both in globular and structural proteins are twisted (Fraser and McRae 1976; Richardson 1981; Lotz et al. 1982), and since experimental findings (Hamodrakas et al. 1984, 1986) indicate that silkmoth chorion proteins contain twisted sheets.
3. The phi and psi angles of residues i + 1, i + 2 of the turns were varied in a systematic way, in order to generate allowed conformations without steric hindrance, and such that the –NH and =CO groups of the β-strands can create hydrogen bonds.

Several possible conformations were found from this analysis which are mostly distorted types of well known β-turns (Fig. 32). They can be classified into three major classes: Type I′ (or II), type II′, and type III β-turns. However, from model building, we selected β-turns I′ (or II), and II′, which, provided they alternate along the structure shown in Figs. 30b and 31b, create satisfactory models for the conservative domains of silkmoth chorion proteins. In these models, favorable hydrogen bonds are formed between the NH group of residue i + 3 with the CO group of residue i.

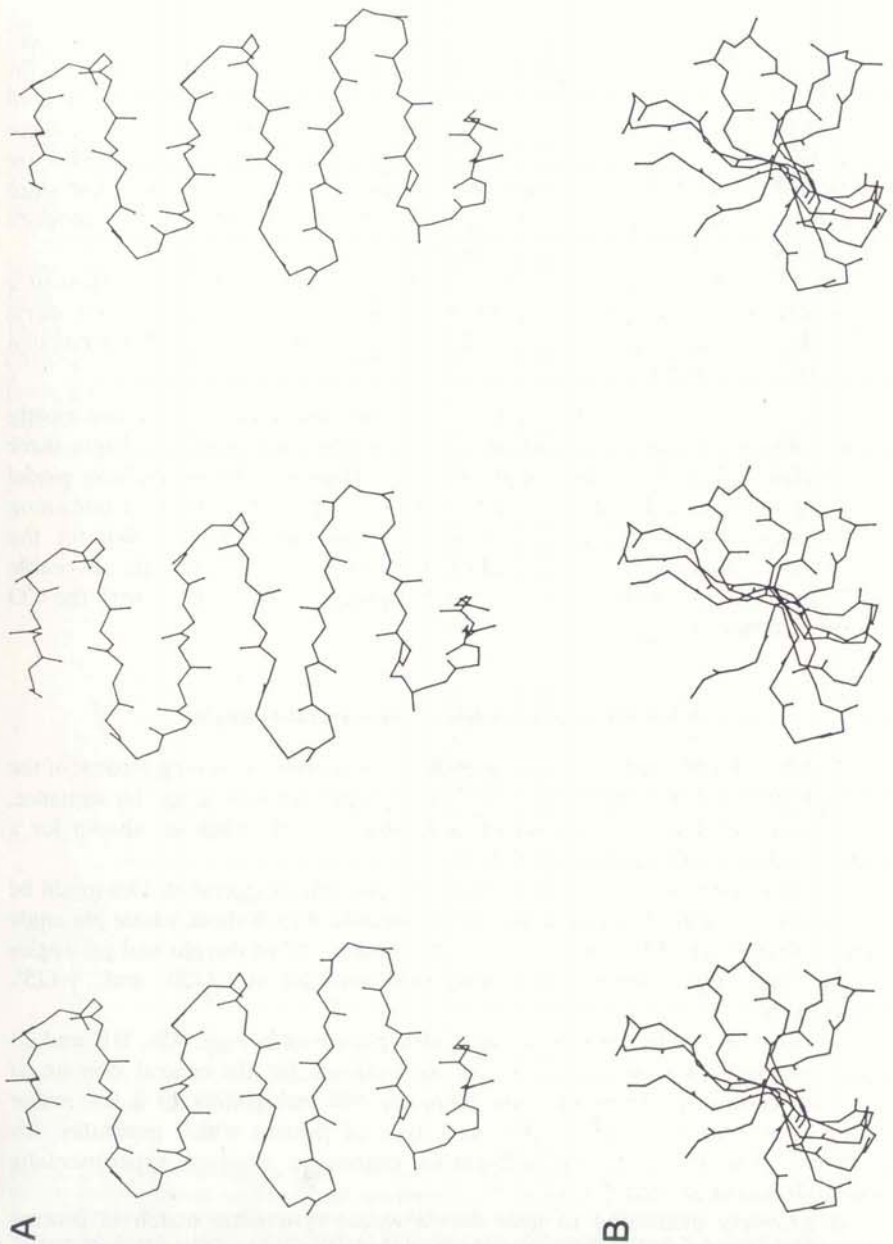
5.1.5 Structural Models of the Central Domain

The models obtained from this stereochemical analysis, assuming β-turns of the I′ (or II) and II′ type, alternating with four-residue β-strands along the sequence, were refined in detail on an interactive graphics system. They are shown for a representative A class protein in Fig. 33.

The β-pleated sheet twist of these models is possibly exaggerated. This might be a consequence of the proline residues in two strands of each sheet, whose phi angle has been constrained to about $-60°$ in the models and of the phi and psi angles of the remaining strand residues being held constant at $-120°$ and $+135°$, respectively.

A structure closely resembling the models presented in Figs. 30b, 31b, and 33 seems almost inevitable in view of all the evidence for the central domain of silkmoth chorion proteins. Perhaps, there are still ambiguities for a few minor points. For example, what is the exact type of β-turns which generates this structure? Unfortunately, the ambiguities cannot be resolved experimentally (Hamodrakas et al. 1988).

It was very interesting to note that a recent systematic search of β-turns connecting antiparallel β-sheet strands in globular proteins (Sibanda and Thornton 1985) has revealed that the types II′ and I′ turns clearly predominate. Out of the 29 β-turns observed experimentally, 15 were type I′, 10 type II′ and only 4 type



I. The type I′ turn differs from the type II turn in the phi and psi angles of only one residue (see Fig. 2 of Sibanda and Thornton 1985). Surprisingly, the type II turn is not observed to link adjacent antiparallel β-sheet strands in globular proteins.

A comparison of our theoretical search (Fig. 32) with Fig. 2 of Sibanda and Thornton (1985) clearly shows that our predicted structures do not differ substantially from those observed. Some of our type II turns can be considered as distorted I′ turns. Sibanda and Thornton (1985) suggest that the abundance of the type I′ turns is probably due to the fact that they have the correct twist to match the relative twist which is always observed between adjacent strands. Close study of our model verifies that this is indeed the case. Therefore, it appears that a further refinement of our proposed model is to adjust the phi and psi angles of only one residue, to belong to a type I′ turn rather than a type II turn. This modification leaves the remainder of the structure unaffected.

The model is further supported by the pattern of residues appearing in the β-turns. Sibanda and Thornton (1985) have found that the observed β-turns are strongly selective for amino acid type. For a type II′ turn they found that Gly predominates in the second position, whereas the third position is usually occupied by Ser, Thr, or a polar residue; for a type I′ turn, the second position is occupied mostly by Gly, Asp, or Asn and the third usually by Gly. Close study of our models (Figs. 30b, 31b) shows that for the type II′ turns the pattern is almost ideal: Gly is usually found in the second position, whereas in the third usually a Thr. For the type I′ (or II) turns the second position is occupied mostly by Gly, in agreement with the observed data, whereas, in the third position, charged (Glu) or polar (Asn) residues frequently appear.

5.1.6 Model Features

This model structure of the central domain, common to all chorion protein families and subfamilies, has the following characteristics:

1. It is highly conservative in each family and subfamily. Greater variability in sequence and perhaps in secondary structure (although this is not certain), is seen in the remainder of the molecules ("arms").
2. In the short four-residue β-strands relatively "small" residues (e.g., G) tend to alternate with "bulky" residues in several cases, e.g., VAVA, VSIG etc. This is reminiscent of the alternation of small (G) and bulky (A, S) residues to opposite sides of the β-sheet structure in silk fibroin (Marsh et al. 1955) and may be important in chorion for the packing of β-sheets to form higher order structure.
3. Both faces of the proposed β-sheets have a pronounced hydrophobic character, except for certain regions (e.g., in the strand VDFC), which might serve as sites

Fig. 33A, B. A skeletal model, obtained from an interactive graphics system, showing the characteristic β-pleated sheet fold of the central conservative domain of the A protein pc609 (main chain and carbonyl oxygens only). **A** View perpendicular to the "plane" of the β-sheet. **B** View perpendicular to the strands, parallel to the β-sheet "plane". The three-picture stereo system used in this figure enables readers with both normal and cross-over stereo vision to view the image. For normal vision select the *left* and *center images*, for cross-over vision use the *center* and *right images*. (Hamodrakas et al. 1988)

for specific recognition. The existence of charged or polar residues in these regions shows that they should be counterbalanced by complementary residues in neighboring β-sheets for the formation of hydrogen or salt bonds during morphogenesis of higher-order structure. It is well known (for a recent analysis, see Rashin and Honig 1984) that this is always the case whenever a charged or polar residue is found in the interior of water soluble globular proteins. A detailed analysis of protein-protein interactions which is well under way for silkmoth chorion proteins will provide further insights (Orfanidou C and Hamodrakas SJ, unpubl.). However, the hydrophobic character of both faces of the β-sheet structures clearly promotes favorable packing of protein molecules in the three-dimensional space along a radial direction of the eggshell.

4. Although both "edges" of the proposed β-sheet structure, i.e., the two central residues of the β-turns, consist mostly of Gly and polar residues, it is clearly seen from Figs. 30b and 31b that they show an uneven distribution of charges. Obviously, the right-hand side of the proposed sheets contains a number of charged residues (particularly Glu in the β branch and Glu, Arg in the α branch; Figs. 30b, 31b). The role of these charges has not yet been clarified, but they appear to be very important for the formation of higher order structure.
5. Certain residues occupy characteristic positions in this β-sheet structure: cysteines are often found in β-turns, in positions favorable to create disulfide bonds, which cross-link adjacent protein molecules during the late choriogenetic stages (Hamodrakas et al. 1984).

Two prolines in the A, B, and C proteins are always highly conserved and are always accompanied by large hydrophobic residues (I, V, L) in two β-strands which are distant by one strand. Their exact role is not yet fully understood. Prolines are not very common in β-strands and they do not favor β-sheet structure (Richardson 1981). Sometimes their presence is marked by the formation of β-bulges (Richardson 1981). Their appearance has seriously impaired our efforts to create satisfactory models of silkmoth chorion proteins. Perhaps they simply serve to increase the sheet twist, if the β-sheets are actually twisted.

Regarding the relation of the proposed structures (Fig. 33) with the helicoidal, higher-order structure of silkmoth chorion, a biological analog of a cholesteric liquid crystal, we can say that the (twisted? helical?) β-sheet ribbons of the central domains of chorion proteins are the basis for the morphogenesis of the helicoidal architecture (Hamodrakas 1984; Hamodrakas et al. 1986). We have already pointed out (cf. Sect. 2.2.1) that Rudall (1956) proposed some 35 years ago, and more recently Bouligand (1978a) as well, that a helicoidal structure can be formed from interactions of helical molecules.

5.2 Structure of the Flanking "Arms"

The quantitative measurement of secondary structure in silkmoth chorion proteins by analysis of the amide I band of the laser-Raman spectrum of chorion, suggests that chorion protein components consist of 60–70% antiparallel β-sheet and 30–40% β-turns, in a ratio 2:1 (cf. Sect. 4.3). These estimates, although in good agreement with the proposed model structure, cannot be attributed solely to the central conservative domains of chorion proteins since these domains account only for about 50% of total chorion mass. To fully account for the observed experimental percentages of secondary structure, the question remains, of course, what structure the variable chorion protein "arms" adopt. They also contain tandemly repetitive peptides evident from the sequences, different in nature from the peptides

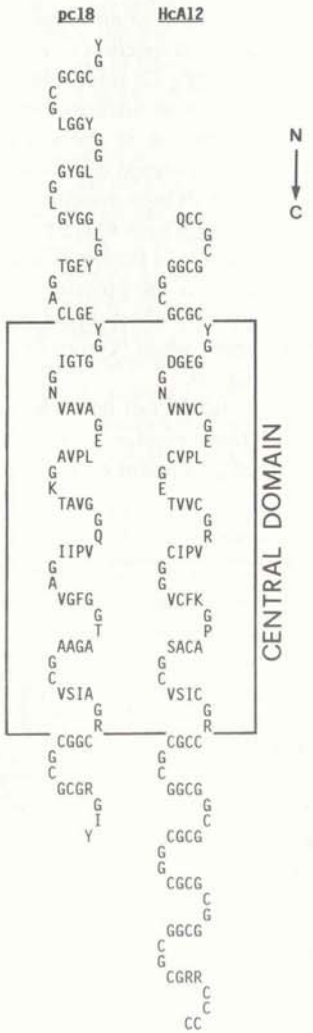


Fig. 34. Schematic, antiparallel β-sheet model of the silkmoth chorion A protein pc18 (*A. polyphemus*) and the HcA protein HcA12 (*B. mori*), assuming a uniform fold throughout their entire length.

of the central region and most probably adopting a β-sheet type of structure (cf. Sect. 4.1). The experimental percentages of protein secondary structure together with the uniform structure and packing of chorion fibrils evident from freeze-fracturing (Sect. 2.2.1) led us to attempt a folding of the protein "arms" similar to the folding of the central domain.

To our surprise, this folding can be done with no serious stereochemical constraints for all chorion proteins, with the possible exception, perhaps, of the "early" minor C proteins which are enriched in Pro in the "arms". Schematic sample folds for two proteins of the α branch are shown in Fig. 34 and a space-filling model of the amino-terminal arm of an A protein in Fig. 35. The resulting structures are attractive in several respects: (1) ensure uniform folding of the proteins throughout their entire length, (2) retain the characteristic features of the central domains: hydrophobicity of sheet surfaces, preferred patterns of residues in the β-turns, unequal charge distribution in the sheet "edges" and usually form "polar" β-sheets, having one "face" occupied mostly by Gly and the other by large hydrophobic residues (Tyr, Leu) which may promote efficient packing. The abundance of Gly in the "arms" implies high arm flexibility, which should be important during chorion formation. This aspect is further discussed by Papanicolaou et al. (1986). Our postulate is, however, that the protein "arms" may adopt a variety of conformations (with a prevalence of β-sheet structure) during choriogenesis and obtain their final uniform structure when "cementing" occurs, through disulfide bond formation at/or near ovulation.

Alternative model structures should not be excluded, however, for the protein arms, i.e., a model β-sheet of three-residue β-strands alternating with β-turns, taking into account the observed pentapeptide periodicity.

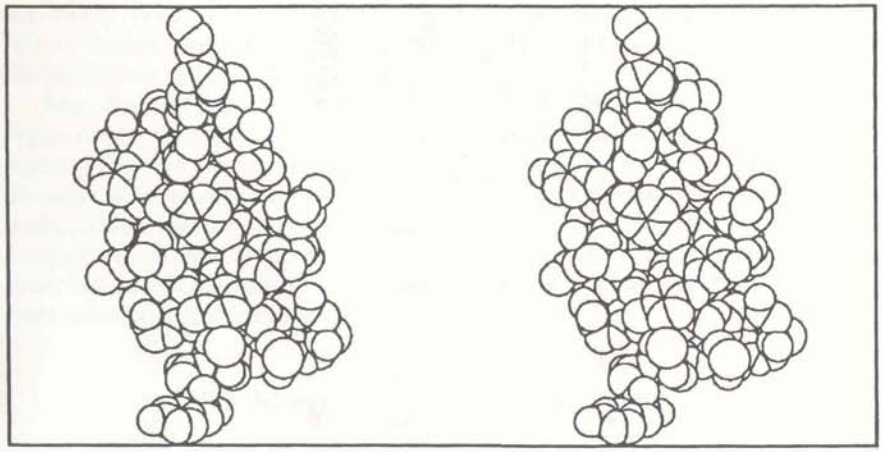


Fig. 35. A stereo (normal vision) space-filling model of the amino-terminal arm of the silkmoth chorion A protein pc18 (*A. polyphemus*). View perpendicular to the "plane" of the β-sheet

6 Models of Helicoidal Eggshell Protein Assembly

Formation of helicoidal eggshell architecture depends on the close packing of chorion proteins. The modes of packing of the detailed structural models presented above are currently under investigation to determine the rules of formation of the helicoidal structure (Orfanidou C and Hamodrakas SJ, in prep.); since β-sheet is the prevalent molecular conformation of individual proteins, these should be based on simple packing rules of β-sheets (Cohen et al. 1981; Chothia and Janin 1982). Twisted β-sheets in globular proteins usually pack with a small negative angle (~ 10–20°) between the sheet axes (or strands), or in an orthogonal fashion (angle close to 90°). Packing of the proposed β-sheets with a small negative angle would explain the observed anticlockwise (left-handed) sense of rotation of the helicoids.

Although it would clearly be premature to propose detailed models of protein-protein interactions due to the complexity of the system, some speculations may have heuristic value: to integrate available information and present a detailed model of three-dimensional chorion architecture, it is necessary to answer the fundamental question "how is a fibril formed?". The cross-sections of individual protein molecules (assuming they fold along their entire length as β-sheet ribbons) must be approximately 30 Å in diameter, taking into account the side chains of the turn residues. Measurements were made directly from molecular CPK space-filling and Kendrew skeletal models and on an interactive graphics screen. These values are in close agreement to the fibril diameter "seen" by freeze-fracturing and X-ray diffraction and are in favor of the attractive notion that one molecule corresponds to one fibril.

However, other key features drive us to assume that a basic structural unit in the bulk of the silkmoth chorion may be formed by a dimer of an A and a B protein instead of a single molecule (the fibril or a disulfide bonded dimer?): experimentally, it has been shown that (1) A and B protein pairs are produced in parallel and in equimolar amounts; their respective genes are clustered, divergently oriented and coordinately expressed (Jones and Kafatos 1980a), (2) disulfide bonded multimers (mostly dimers and trimers) may act as intermediates in the assembly process, showing preferred family linkage patterns (Regier and Wong 1988). Extensive analysis of the sequences and of models (Orfanidou C and Hamodrakas SJ, in prep.) suggests that all known, coordinately expressed A and B protein pairs (i.e., pc18-pc401, pc292-pc10a, AL11-BL11, AL12-BL12, etc; Regier and Kafatos 1985; Spoerel et al. 1989) may form dimers of the general form shown in Fig. 36.

These dimers which consist of an A and its complementary B molecule, each folded as a β-sheet ribbon throughout its entire length, and packed in an antiparallel fashion so that the "faces" of the β-sheet ribbons containing the side-chains of the invariant proline residues of their central domains (cf. Sect. 5.1) are "buried", combine the following significance features (Fig. 36):

1. Their cross-sections are of the same order of magnitude (30–40 Å) as those of individual molecules (~ 30 Å).

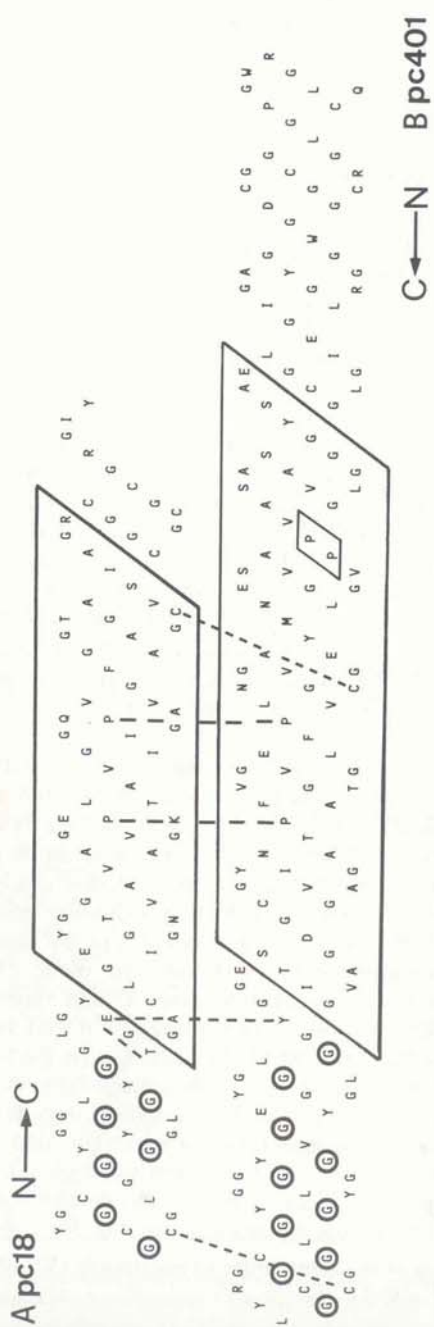


Fig. 36. Schematically, a packing model of an A–B silkmoth chorion protein dimer. The molecules pack in an antiparallel fashion and with the β-sheet faces containing the invariant proline residues of the central domain (*dotted lines*) "buried" in the hydrophobic core of the β-sheet sandwich. For details of the model features and specific interactions see text

2. Their central domains form a β-sheet sandwich, with the β-sheet surfaces exhibiting the most pronounced hydrophobic character buried in the interior.
3. The antiparallel fashion of packing ensures that the N-terminal arm of the A protein is packed against the distantly homologous (Regier and Kafatos 1985) carboxyl-terminal arm of the B, which has a similar length. Both contain tandem repeats of the pentapeptide LGYGG and its variants. The packing is such that the faces of the β-sheet ribbons containing mostly Gly are packed against each other, which ensures uniform packing. The beginning of the A central domain corresponds to the end of the B, whereas the end β-strand of the A central domain has as counterpart in the B, a β-strand containing two conserved, consecutive prolines in *A. polyphemus* sequences. This may be an "inflection point" in the conformation of the amino-terminal arm of the B's which has no "partner" from the A molecule. In *B. mori* sequences, the strain imposed by the two consecutive prolines is relieved since only one of them is conserved.
4. The four (two in each protein) invariant proline residues of the central domains pack with their side chains facing each other.
5. Buried polar or charged residues are fully counterbalanced by adjacent polar or of opposite charge residues, which is very important for the stability of the hydrophobic core. In the example shown (Fig. 36) a glutamyl residue just before the beginning of the central domain in Apc18 is counterbalanced by a Thr in the same strand and/or a "facing" Tyr in Bpc401.
6. Cysteines are found in favorable positions to form disulfide bonds, stabilizing the dimer and may also promote efficient cross-linking in three dimensions with other monomers or dimers.

These models were found intuitively. However, since there are eight possible ways of packing for each dimer (each molecule has two "faces" and the molecules may pack in a parallel or antiparallel fashion), efficient packing was checked in each case, by "sliding" surfaces on top of each other and scoring favorable interactions (gain in free energy by the burial of hydrophobic chains, counterbalancing of buried polar or charged residues, possible formation of disulfide bonds, etc.) and it was found that the models shown represent the most stable possible dimers. Furthermore, an additional check was performed utilizing an interactive graphics system.

A full account of these findings will be given elsewhere (Orfanidou C and Hamodrakas SJ, in prep.). Here, we propose, however, that these dimers may represent our starting point to fully understand the molecular interactions leading to helicoidal self-assembly. Refined biochemical work is needed along the lines suggested by Regier and Kafatos (1985) and also further modeling studies to define their possible roles in the formation of higher order structure and also verify their validity.

7 Solution Structure and Assembly of Helicoidal Eggshell Proteins

One of our future aims is to study the in vitro assembly process of chorion proteins in detail and determine the structural parameters of reconstituted units, making comparisons with native chorion.

As a first step, lepidopteran chorions were solubilized in the presence of denaturing and reducing agents (6 M urea, 0.4 M Tris-HCl, 1% 2-mercaptoethanol). Extracted chorion proteins were reassembled by a 12–18 h dialysis at room temperature against double-distilled water. Reassembled (polymerized) units were studied by negative staining (Fig. 37), dark field electron microscopy and circular dichroism (CD) spectroscopy (Fig. 38) to determine: (1) size and shape of reconstituted units (2) secondary structure of their constituent proteins.

Negative staining (Fig. 37) shows that chorion polypeptides assemble to form globular structures of varying diameters, 50–200 Å, which are cross-linked, at the low pH conditions of the experiment (pH ~3). Dark field microscopy reveals that similar in diameter globular structures result from polymerization of dissolved chorion components, even at near normal pH (~7) (data not shown). Nevertheless, the globules are not cross-linked in this case.

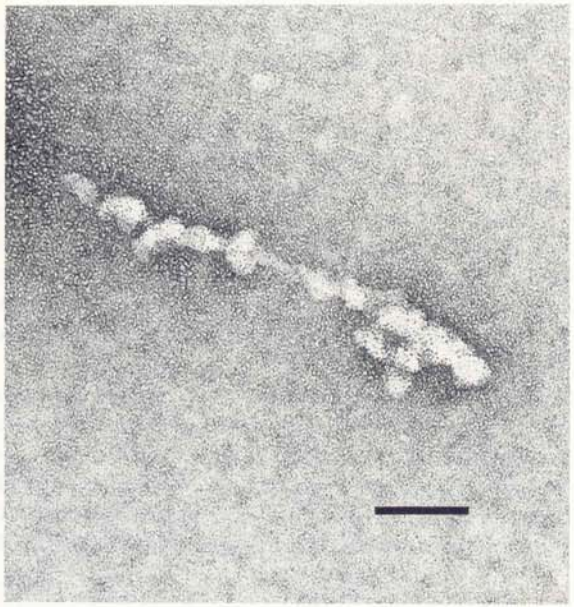


Fig. 37. Transmission electron micrographs of extracted silkmoth chorion proteins (6 M urea, 0.4 M Tris-HCl, 1% 2-mercaptoethanol), reassembled by dialysis at room temperature against double distilled water and contrasted by negative staining (1% uranyl acetate). Cross-linked globular structures of varying diameters (5–20 nm) are seen. *Bar* 100 *nm*

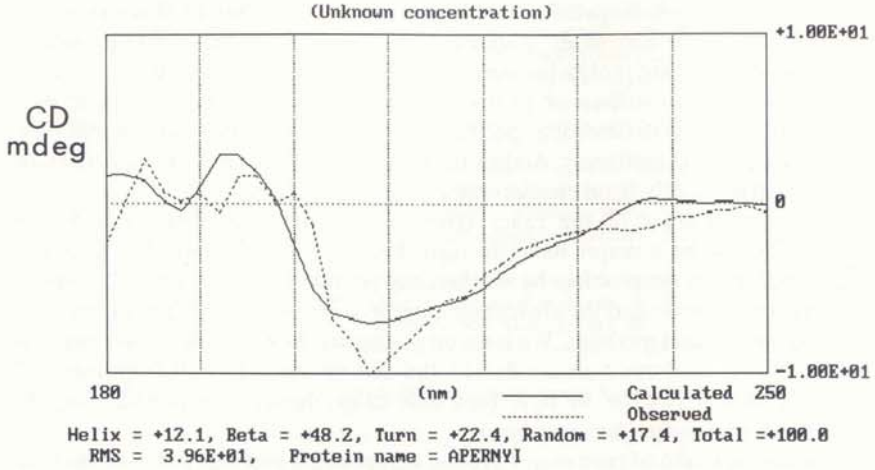


Fig. 38. Circular dichroism (*CD*) spectrum of a water solution containing reassembled units from extracted *A. pernyi* chorion proteins (details as in Fig. 37). The observed spectrum is shown by the *dotted line*. The *solid line* represents a (fitted) calculated spectrum by expressing the observed spectrum as a linear combination of reference spectra of four types of secondary structure: α-helix (α), β-sheet (β), β-turn (t), and coil (c; random). The linear coefficients representing secondary structure estimates were calculated by a linear least-squares method. The reference spectra of the four types of secondary structure were calculated from the CD spectra of four proteins with crystallographically known secondary structures: lysozyme (41%α, 16%β, 23%t, 20%c), ribonuclease A (23%α, 40%β, 13%t, 24% c), papain (28%α, 14%β, 17%t, 41%c), chymotrypsin A (9%α, 34%β, 34%t, 23%c)

The most surprising finding of these experiments, however, was that, in solution, the proteins which constitute these units, retain a very high percentage of β-sheet structure (usually more than 50%), in a ratio with β-turns of approximately 2:1, in many respects reminiscent of the secondary structure percentages obtained for the proteins of native chorions (Fig. 38; Sect. 4.3). It is, therefore, tempting to make the reasonable hypothesis that the prominent regular β-sheet structure of chorion proteins may persist in solution, dictating the aggregation and polymerization process in solution as well. A full account of these preliminary but instructive experiments will be given elsewhere (Hamodrakas SJ, Wellman SE, Case ST, and Ottensmeyer FP, in prep.).

8 Future Aims

Our molecular understanding of helicoidal proteinaceous eggshell structure is still quite primitive due to the complexity of the system and of its constituent proteins. However, many interesting questions will be experimentally and theoretically approachable in the near future.

Reconstitution studies are likely to furnish new insights into fibril structure and assembly (in space and time). Variation of several parameters in vitro will be attempted to simulate molecular events occurring in eggshell assembly.

Our theoretical studies of protein-protein interactions will continue, with particular emphasis in detecting specific sites of recognition between protein pairs and higher possible multimers. Additional helicoidal eggshell protein sequences are desirable, particularly from species other than the silkmoths.

The identification of the exact types of β-turns in eggshell proteins unequivocally will be a major breakthrough. Perhaps it will be possible to obtain partial answers to the problem by synthesizing peptides representative of segments of the central domain and the arms of chorion proteins and studying their structure utilizing biophysical methods. We have no guarantee, however, that these peptides will fold to a conformation similar to the one in vivo. In collaboration with Professor Steven T. Case, we have been able to synthesize such peptides and the study of their structure has just begun.

Finally, isolation of pure major protein components and study of their molecular structure may provide the necessary clues towards unraveling in molecular detail the secrets of helicoidal eggshell architecture.

9 Synopsis

I will not attempt to summarize this chapter, which is already a summary. I merely wish to point out that experimental and theoretical evidence to date clearly suggests that antiparallel β-pleated sheet dictates self-assembly in helicoidal proteinaceous eggshells. Molecular details of this process have started to become clear after the development of the specific, most probably correct, protein structural models in the case of the silkmoths, where amino acid information is available and with the help of several experimental techniques. However, for people seeking universal mechanisms the picture should still be far from complete. Several analogous systems should be studied before providing final answers.

Acknowledgments. I am indebted to all those whose publications are reviewed. I wish to thank Professor Lucas H. Margaritis, my Ph.D students Paraskevi Papadopoulou and Constance Orfanidou, and my son John for their invaluable help during the preparation of this manuscript.

References

Aebi U, Fowler WE, Rew P, Sun TT (1983) The fibrilar substructure of keratin filaments unraveled. J Cell Biol 97:1131–1143

Aggeli A, Hamodrakas SJ, Kanitopoulou K, Konsolaki M (1991) Tandemly repeating peptide motifs and their secondary structure in *Ceratitis capitata* eggshell proteins Ccs36 an Ccs38. Int J Biol Macromol 13:307–315

Anderson E (1967) The formation of the primary envelope during oocyte differentiation in teleosts. J Cell Biol 35:193–212

Anfinsen CB (1973) Principles that govern the folding of protein chains. Science 181:223–230

Barker WC, Hunt LT, George DG, Yeh LS, Chen HR, Blomquist MC, Seibel-Rose EI, Elzanowski A, Hong MK, Ferrick DA, Bair JK, Chen SL, Ledley RS (1986) Protein sequence database. National Biomedical Research Foundation, Georgetown University, Washington, DC

Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasoumi M (1977) The protein data bank: a computer-based archival file for macromolecular structures. J Mol Biol 112:535–542

Blau HM, Kafatos FC (1979) Morphogenesis of the silkmoth chorion: patterns of distribution and insolubilization of the structural proteins. Dev Biol 72:211–225

Blundell TL, Johnson LN (1976) Protein crystallography. Academic Press, New York

Bouligand Y (1972) Twisted fibrous arrangements in biological materials and cholesteric mesophases. Tissue Cell 4:189–217

Bouligand Y (1978a) Cholesteric order in biopolymers. Am Chem Soc Symp Ser 74:237–247

Bouligand Y (1978b) Liquid crystalline order in biological materials. In: Blumstein A (ed) Liquid crystalline order in polymers. Academic Press, New York, pp 261–297

Burke WD, Eickbush TH (1986) The silkmoth late chorion locus. I. Variation within two paired multigene families. J Mol Biol 190:343–356

Carey PR (1982) Biochemical applications of Raman and resonance Raman spectroscopies. Academic Press, New York

Chothia C, Janin J (1982) Orthogonal packing of β-pleated sheets in proteins. Biochemistry 21:3955–3965

Chou PY, Fasman GD (1977) β-turns in proteins. J Mol Biol 115:135–175

Chou PY, Fasman GD (1978) Prediction of the secondary structure of proteins from their amino acid sequence. Adv Enzymol 47:45–148

Cohen FE, Sternberg MJE, Taylor WR (1981) Analysis of the tertiary structure of protein β-sheet sandwiches. J Mol Biol 148:253–272

Crick FHC (1953) The packing of α-helices in simple coiled-coils. Acta Cryst 6:689–697

Davenport J, Lonning S, Kjorsvik E (1986) Some mechanical and morphological properties of the chorions of marine teleost eggs. J Fish Biol 29:289–301

Eickbush TH, Kafatos FC (1982) A walk in the chorion locus of *Bombyx mori*. Cell 29:633–643

Eickbush TH, Rodakis GC, Lekanidou R, Kafatos FC (1985) A complex set of early chorion DNA sequences from *Bombyx mori*. Dev Biol 112:368–376

Fehrenbach H, Dittrrich V, Zissler D (1987) Eggshell fine structure of three lepidopteran pests: *Cydia pomonella* (Tortricidae), *Heliothis virescens* and *Spodoptera littoralis* (Noctuidae). Int J Insect Morphol Embryol 16(3):201–219

Filshie BK, Rogers GE (1962) An electron microscope study of the fine structure of feather keratin. J Cell Biol 13:1–12

Filshie BK, Smith DS (1980) A proposed solution to a fine-structural puzzle: the organisation of gill cuticle in a crayfish (panulirus). Tissue Cell 12(1):209–226

Flugel H (1967) Licht- und elektronenmikroskopiche Untersuchungen an Oozyten und Eiern einiger Knochenfische. Z Zellforsch Mikrosk Anat 83:82–116

Fraser RDB, McRae TP (1959) Molecular organization in feather keratin. J Mol Biol 1:387–397

Fraser RDB, McRae TP (1973) Conformation in fibrous proteins. Academic Press, New York

Fraser RDB, McRae TP (1976) The molecular structure of feather keratin. In: Frith HJ, Calaby JH (eds) Proc 16th Int Ornithological Congress, Canberra. Australian Academy of Science, Canberra, pp 443–451

Friedel MG (1922) Les états mésomorphes de la matière. Ann Phys (Paris) 18:273–474

Frushour BJ, Koenig JL (1975) Raman Spectroscopy of proteins. In: Clark RJH, Hester RE (eds) Advances in infrared and Raman spectroscopy, vol I. Heyden, London, p 35

Furneaux PJS, Mackay AL (1972) Crystalline protein in the chorion of insect eggshells. J Ultrastrŭct Res 38:343–359

Geddes AJ, Parker KD, Atkins EDT, Beighton E (1968) Cross-β conformation in proteins. J Mol Biol 32:343–358

Giraud MM, Castanet J, Meunier FJ, Bouligand Y (1978) The fibrous structure of coelacanth scales: a twisted "plywood". Tissue Cell 10:671–686

Goldsmith MR, Kafatos FC (1984) Developmentally regulated genes in silkmoths. Annu Rev Genet 18:443–487

Green NM, Wrigley NG, Russel WC, Martin SR, McLachlan AD (1983) Evidence for a repeating cross-β sheet structure in the adenovirus fibre. EMBO J 2:1357–1365

Gregg K, Wilton SD, Parry DAD, Rogers GE (1984) A comparison of genomic coding sequences for feather and scale keratins: structural and evolutionary implications. EMBO J 3:175–181

Grierson JP, Neville AC (1981) Helicoidal architecture of fish eggshell. 13:819–830
Groot EP, Alderdice DF (1985) Fine structure of the external egg membrane of five species of Pacific salmon and steelhead trout. Can J Zool 63:552–566
Gubb D (1975) A direct visualisation of helicoidal architecture in *Carcinus maenas* and *Halocynthia papillosa* by scanning electron microscopy. Tissue Cell 7:19–32
Hagenmaier HE, Schmitz J, Fohles J (1976) Zum Vorkommen von Isopeptidbindungen in der Eihülle der Regenbogenforelle (*Salmo gairdneri* Rich). Hoppe-Seyler's Z Physiol Chem 357:1435–1438
Hamodrakas SJ (1984) Twisted β-pleated sheet: the molecular conformation which possibly dictates the formation of the helicoidal architecture of several proteinaceous eggshells. Int J Biol Macromol 6:51–53
Hamodrakas SJ (1988) A protein secondary structure prediction scheme for the IBM PC and compatibles. CABIOS 4:473–477
Hamodrakas SJ, Kafatos FC (1984) Structural implications of primary sequences from a family of Balbiani ring-encoded proteins in *Chironomus*. J Mol Evol 20:296–303
Hamodrakas SJ, Jones CW, Kafatos FC (1982a) Secondary structure predictions for silkmoth chorion proteins. Biochim Biophys Acta 700:42–51
Hamodrakas SJ, Asher SA, Mazur GD, Regier JC, Kafatos FC (1982b) Laser-Raman studies of protein conformation in the silkmoth chorion. Biochim Biophys Acta 703:216–222
Hamodrakas SJ, Paulson JR, Rodakis GC, Kafatos FC (1983) X-ray diffraction studies of a silkmoth chorion. Int J Biol Macromol 5:149–153
Hamodrakas SJ, Kamitsos EI, Papanicolaou A (1984) Laser-Raman spectroscopic studies of the eggshell (chorion) of *Bombyx mori*. Int J Biol Macromol 6: 333–336
Hamodrakas SJ, Etmektzoglou T, Kafatos FC (1985) Amino acid periodicities and their structural implications for the evolutionary conservative central domain of some silkmoth chorion proteins. J Mol Biol 186:583–589
Hamodrakas SJ, Margaritis LH, Papasideri I and Fowler A (1986) Fine structure of the silkmoth *Antheraea polyphemus* chorion as revealed by X-ray diffraction and freeze fracturing. Int J Biol Macromol 8:237–242
Hamodrakas SJ, Kamitsos EI, Papadopoulou PG (1987) Laser-Raman and infrared spectroscopic studies of protein conformation in the eggshell of the fish *Salmo gairdneri*. Biochim Biophys Acta 913:163–169
Hamodrakas SJ, Bosshard HE, Carlson CN (1988) Structural models of the evolutionarily conservative central domain of silk-moth chorion proteins. Prot Eng 2:201–207
Hamodrakas SJ, Batrinou A, Christoforatou T (1989) Structural and functional features of *Drosophila* chorion proteins s36 and s38 from analysis of primary structure and infrared spectroscopy. Int J Biol Macromol 11:307–313
Hinton H (1981) Biology of insect eggs. Pergamon Press, Oxford
Hojrup P, Andersen SO, Roepstorff P (1986) Isolation, characterization and N-terminal sequence studies of cuticular proteins from the migratory locust *Locusta migratoria*. Eur J Biochem 154:153–159
Hurley DA, Fischer KC (1966) The structure and development of the external membrane in young eggs of the brook trout, *Salvelinus fontinallys* (Mitschill). Can J Zool 44:173–190
Iatrou K, Tsitilou SG, Kafatos FC (1984) DNA sequence transfer between two high-cysteine chorion gene families in *Bombyx mori*. Proc Natl Acad Sci USA 81:4452–4456
Jones CW, Kafatos FC (1980a) Coordinately expressed members of two chorion multi-gene families are clustered, alternating and divergently oriented. Nature 284:635–638
Jones CW, Kafatos FC (1980b) Structure, organization and evolution of developmentally regulated chorion genes in a silkmoth. Cell 22:855–867
Jones CW, Kafatos FC (1982) Accepted mutations in a gene family: evolutionary diversification of duplicated DNA. J Mol Evol 19:87–103
Kafatos FC, Regier JC, Mazur GD, Nadel MR, Blau HM, Petri WH, Wyman AR, Gelinas RE, Moore PB, Paul M, Efstratiadis A, Vournakis JN, Goldsmith MR, Hunsley JR, Baker B, Nardi J, Koehler M (1977) The eggshell of insects: differentiation-specific proteins and the control of their synthesis and accumulation during development. In: Beerman W (ed) Results and problems in cell differentiation, vol 8. Springer, Berlin Heidelberg New York, pp 45–145
Kakudo M, Kasai N (1972) X-ray diffraction by polymers. Elsevier, Amsterdam
Kawasaki H, Sato H, Suzuki M (1971) Structural proteins in the silkworm eggshells. Insect Biochem 1:130–148
King RC, Aggarwal SK (1965) Oogenesis in *Hyalophora cecropia*. Growth 29:17–83
Kobayashi W (1982) The fine structure and amino acid composition of the envelope of the chum salmon egg. J Fac Sci Hokkaido Univ Ser 6 23:1–12

Lecanidou R, Rodakis GC, Eickbush TH, Kafatos FC (1986) Evolution of the silkmoth chorion gene superfamily: gene families CA and CB. Proc Natl Acad Sci USA 83:6514–6518
Lipman DJ, Pearson WR (1985) Rapid and sensitive protein similarity searches. Science 227:1435–1441
Livolant F, Bouligand Y (1989) Freeze-fractures in cholesteric mesophases of polymers. Mol Cryst Liq Cryst 166:91–100
Lonning S, Kjorsvik E, Davenport J (1984) The hardening process of the egg chorion of the cod, *Gadus morhua* L., and lumpsucker, *Cyclopterus lumpus* L. J Fish Biol 24:505–522
Lotz B, Gouthier-Vassal A, Brack A, Magoshi J (1982) Twisted single crystals of *Bombyx mori* silk fibroin and related model polypeptides with β-structure. J Mol Biol 156:345–357
Margaritis LH (1985) Structure and physiology of the eggshell. In: Gilbert LI, Kerkut GA (eds) Comprehensive insect biochemistry, physiology and pharmacology, vol I. Pergamon, Oxford, pp 153–230
Marsh RE, Corey RB, Pauling L (1955) The structure of silk fibroin. Biochim Biophys Acta 16:1–34
Mazur GD, Regier JC, Kafatos FC (1980) The silkmoth chorion: Morphogenesis of surface structures and its relation to synthesis of specific proteins. Dev Biol 76:305–321
Mazur GD, Regier JC, Kafatos FC (1982) Order and defects in the silkmoth chorion, a biological analogue of a cholesteric liquid crystal. In: Akai H, King RC (eds) Insect ultrastructure, vol I. Plenum, New York, pp 150–183
McLachlan AD (1977) Analysis of periodic patterns in amino acid sequences: collagen. Biopolymers 16:1271–1297
McLachlan AD, Stewart M (1976) The 14-fold periodicity in α-tropomyosin and the interaction with actin. J Mol Biol 103:271–298
Neville AC (1975) Biology of the arthropod cutide. Springer, Berlin Heidelberg New York
Neville AC (1981) Cholesteric proteins. Mol Cryst Liq Cryst 76:279–286
Neville AC (1986) The physics of helicoids: multidirectional "plywood" structures in biological systems. Phys Bull 37:74–76
Ohzu E, Kusa M (1981) Amino acid composition of the egg chorion of rainbow trout. Annot Zool Jpn 54:241–244
Orcutt BC, George DG, Dayhoff MO (1983) Protein and nucleic acid sequence database systems. Annu Rev Biophys Bioeng 12:419–441
Papanicolau AM, Margaritis LH, Hamodrakas SJ (1986) Ultrastructural analysis of chorion formation in the silkmoth *Bombyx mori*. Can J Zool 64:1158–1173
Parker FS (1971) Applications of infrared spectroscopy in biochemistry, biology and medicine. Plenum, New York
Parry DAD (1979) Determination of structural information from the amino acid sequences of fibrous proteins. In: Parry DAD, Creamer LK (eds) Fibrous proteins: scientific, industrial and medical aspects, vol I. Academic Press, London, pp 393–427
Parry DAD, Fraser RDB, McRae TP (1979) Repeating patterns of amino acid residues in the sequences of some high-sulphur proteins from α-keratin. Int J Biol Macromol 1:17–22
Pau RN (1984) Cloning of cDNA for a juvenile hormone-regulated oothecin mRNA. Biochim Biophys Acta 782:422–428
Rashin AA, Honig B (1984) On the environment of ionizable groups in globular proteins. J Mol Biol 173:515–521
Regier JC (1986) Evolution and higher-order structure of architectural proteins in silkmoth chorion. EMBO J 5:1981–1989
Regier JC, Kafatos FC (1985) Molecular aspects of chorion formation. In: Gilbert LI, Kerkut GA (eds) Comprehensive insect biochemistry, physiology and pharmacology, vol I, Pergamon, Oxford pp 113–151
Regier JC, Vlahos NS (1988) Heterochrony and the introduction of novel modes of morphogenesis during the evolution of moth choriogenesis. J Mol Evol 28:19–31
Regier JC, Wong JR (1988) Assembly of silkmoth proteins: in vivo patterns of disulphide bond formation. Insect Biochem 18:471–482
Regier JC, Kafatos FC, Goodfliesh R, Hood L (1978a) Silkmoth chorion proteins: sequence analysis of the products of a multigene family. Proc Natl Acad Sci USA 75:390–394
Regier JC, Kafatos FC, Kramer KJ, Heinrikson RL, Keim PS (1978b) Silkmoth chorion proteins: their diversity, amino acid composition and the NH_2-terminal sequence of one component. J Biol Chem 253:1305–1314
Regier JC, Mazur GD, Kafatos FC (1980) The silkmoth chorion: morphological and biochemical characterization of four surface regions. Dev Biol 76:286–304
Regier JD, Mazur GD, Kafatos FC, Paul M (1982) Morphogenesis of silkmoth chorion: initial framework formation and its relation to synthesis of specific proteins. Dev Biol 92:159–174

Regier JC, Kafatos FC, Hamodrakas SJ (1983) Silkmoth chorion multigene families constitute a superfamily: comparison of C and B family sequences. Proc Natl Acad Sci USA 80:1043–1047
Richardson JS (1981) The anatomy and taxonomy of protein structure. Adv Prot Chem 34:167–339
Rill RL, Livolant F, Aldrich HC, Davidson MW (1989) Electron microscopy of liquid crystalline DNA: direct evidence for cholesteric-like organisation of DNA in dinoflagellate chromosomes. Chromosoma (Berl) 98:280–286
Rodakis GC, Kafatos FC (1982) Origin of evolutionary novelty in proteins: how a high-cysteine chorion protein has evolved. Proc Natl Acad Sci USA 79:3551–3555
Rodakis GC, Moschonas NK, Kafatos FC (1982) Evolution of a multigene family of chorion proteins in silkmoths. Mol Cell Biol 2:554–563
Rodakis GC, Lekanidou R, Eickbush TH (1984) Diversity in a chorion multigene family created by tandem duplications and a putative gene conversion event. J Mol Evol 20:265–273
Rudall KM (1956) Protein ribbons and sheets. In: Lectures on the scientific basis of medicine 5. Athlone Press, London, pp 217–230
Schulz GE, Schirmer RH (1978) Principles of protein structure. Springer, New York Heidelberg Berlin
Siamwiza MN, Lord RC, Chen MC, Takamatsu T, Harada I, Matsuura H, Shimanouchi T (1975) Interpretation of the doublet at 850 and 830 cm^{-1} in the Raman spectra of tyrosyl residues in proteins and certain model compounds. Biochemistry 14:4870–4876
Sibanda BL, Thornton JM (1985) β-hairpin families in globular proteins. Nature 316:170–174
Smith DS, Telfer WH, Neville AC (1971) Fine structure of the chorion of a moth, *Hyalophora cecropia*. Tissue Cell 3:477–498
Spiro TG, Gaber BP (1977) Laser-Raman scattering as a probe of protein structure. Annu Rev Biochem 46:553–572
Spoerel NA, Nguyen HT, Eickbush TH, Kafatos FC (1989) Gene evolution and regulation in the chorion complex of *Bombyx mori*: hybridization and sequence analysis of multiple developmentally middle A/B chorion gene pairs. J Mol Biol 209:1–19
Squire JM, Vibert PJ (1987) Fibrous protein structure. Academic Press, London
Stewart M (1977) The structure of chicken scale keratin. J Ultrastruct Res 60:27–33
Sugeta H, Go A, Miyazawa T (1972) S–S and C–S stretching vibrations and molecular conformations of dialkyl disulphides and cystine. Chem Lett 1:83–86
Taylor WR (1987) Protein structure prediction In: Bishop MJ, Rawlins CJ (eds) Nucleic acid and protein sequence analysis: a practical approach. IRL Press, Oxford, pp 285–322
Telfer WH, Smith DS (1970) Aspects of egg formation. Symp R Entomol Soc Lond 5:165–185
Tesoriero JV (1977) Formation of the chorion (zona pellucida) in the teleost *Oryzias latipes*.I. Morphology of early oogenesis. J Ultrastruct Res 59:282–291
Vanderlei R, Chaudhri M, Knight M, Meadows H, Chambers A, Taylor W, Kelly C, Simpson AJG (1989) Predicted structure of a major *Schistosoma mansoni* eggshell protein. Mol Biochem Parasitol 32:7–14
Walker ID, Bridgen J (1976) The keratin chains of avian scale tissues. Eur J Biochem 67:283–293
Williams RW, Dunker AK (1981) Determination of the secondary structure of proteins from the amide I band of the laser-Raman spectrum. J Mol Biol 152:783–813
Wourms JP (1976) Annual fish oogenesis. I. Differentiation of the mature oocyte and formation of the primary envelope. Dev Biol 50:338–354
Xu M, Lewis RV (1990) Structure of a protein superfiber: spider dragline silk. Proc Natl Acad Sci USA 87:7120–7124
Young EG, Inman WR (1938) The protein casing of salmon eggs. J Biol Chem 124:189–193
Yu NT (1977) Raman spectroscopy: a conformational probe in biochemistry. CRC Crit Rev Biochem 4:229–280