Summary. The extracellular coat surrounding fish (vitelline envelope; VE) and mammalian (zona pellucida; ZP) eggs is composed of long, interconnected filaments. Fish VE and mammalian ZP proteins that make up the filaments are highly conserved groups of proteins that are related to each other, as well as to their amphibian and avian egg counterparts. The rainbow trout (O. mykiss) egg VE is composed of 3 proteins, called VEα (~58 kDa), VEβ (~54 kDa), and VEγ (~47 kDa). The mouse (M. musculus) egg ZP also is composed of 3 proteins, called ZP1 (~200 kDa), ZP2 (~120 kDa), and ZP3 (~83 kDa). Overall, trout VE and mouse ZP proteins share ~25% sequence identity and have features in common; these include an N-terminal signal sequence, a ZP domain, a consensus furin cleavage-site, and a C-terminal tail. VEα, VEβ, and ZP1 also have a trefoil or P-type domain upstream of the ZP domain. VEα and VEβ are very similar in sequence (~65% sequence identity) and are related to ZP1 and ZP2, whereas VEγ is related to ZP3 (~25% sequence identity). Mouse ZP proteins are synthesized and secreted exclusively by growing oocytes in the ovary. Trout VE proteins are synthesized by the liver under hormonal control and transported in the bloodstream to growing oocytes in the ovary. The trout VE is assembled from VEα/γ and VEβ/γ heterodimers. The mouse ZP is assembled from ZP2/3 heterodimers and crosslinked by ZP1. Despite ~400 million years separating the appearance of trout and mice, and the change from external to internal fertilization and development, trout VE and mouse ZP proteins have many common structural features; as do avian and amphibian egg VE proteins. However, the site of synthesis of trout and mouse egg extracellular coat proteins has changed over time from the liver to the ovary, necessitating some changes in the C-terminal region of the polypeptides that regulates processing, secretion, and assembly of the proteins.

Key words: Trout eggs, Mouse eggs, Vitelline envelope, Zona pellucida, ZP domain, Sequences, Synthesis, Polymerization, Evolution

Introduction

Virtually all eggs from both invertebrates and vertebrates are surrounded by an extracellular coat (EC) composed of (glyco)protein (Dumont and Brummett, 1985; Dietl, 1989; Monne et al., 2006). In many instances the EC appears and increases in thickness as growing oocytes increase in diameter. In general, the EC serves a protective function for eggs and for many organisms plays other significant roles during oogenesis, fertilization, and early embryogenesis.

The EC of most eggs undergoes structural rearrangements after fertilization making the coat much more durable and able to prevent penetration of supernumerary sperm through the EC. In addition, there is considerable evidence to suggest that binding of mammalian sperm to eggs is species-restricted due in large part to the presence of receptors for homologous sperm in the egg EC (Gwatkin, 1977; Wassarman, 1983, 1987, 1999; Hardy, 2002; Florman and Ducibella, 2006). In the case of mammalian eggs, it is clear that following fertilization by a single sperm, changes in the EC prevent further binding and penetration of sperm. On the other hand, fish eggs have a micropyle at the animal pole region that is generally funnel-shaped and which attracts sperm to this region; as a result, the number of sperm entering and attaching to the egg plasma membrane is restricted (Hart and Donovan, 1983; Iwamatsu et al., 1997). Therefore, it is not necessary for fish sperm to bind to the egg EC. Following fertilization by a single sperm, fish eggs form a fertilization cone to plug the micropyle and the EC is hardened (Kudo, 1980; Shibata et al., 2000) as part of the block to polyspermy. These and other features of invertebrate and vertebrate egg ECs have made them the subject of extensive investigation for well over a century.

Rainbow trout (O. mykiss) and mouse (M. musculus) eggs are quite different in appearance. Trout eggs are large (~3 mm diameter), pigmented, and have a thick
(~50 μm; ~600 μg protein) EC, or vitelline envelope (VE). Mouse eggs, on the other hand, are small (~80 μm diameter), transparent, and have a relatively thin (~6 μm; ~3.5 ng protein) EC, or zona pellucida (ZP). While there are significant differences in appearance of mouse and trout eggs, their ECs are composed of proteins bearing remarkably similar structural characteristics. That is, over the last 400 million years or so of evolution, egg EC proteins have retained some essential features of their primary structure that enables them to perform important functions during oogenesis, fertilization, and early development (Dumont and Brummett, 1985; Wassarman, 1988; Wassarman et al., 2001; Monne et al., 2006).

Here, we review briefly some of the key features of rainbow trout VE and mouse ZP proteins, from their synthesis and structure, to their assembly into an egg EC. Reference is made to the synthesis and structure of ECs of amphibian and avian eggs as well.

### Synthesis of mammalian ZP and non-mammalian VE proteins

Ovaries of mammals are populated by small, non-growing oocytes that lack a ZP. Mouse ZP proteins are synthesized exclusively in the ovary by growing oocytes during each reproductive cycle (Bleil and Wassarman, 1980; Greve et al., 1982; Salzmann et al., 1983; Shimizu et al., 1983; Lira et al., 1990, 1993). The ZP increases in thickness as oocytes increase in diameter and surrounding follicle cells increase markedly in number during a ~3-week period, resulting finally in the formation of a Graafian follicle from which an unfertilized egg is ovulated. ZP proteins from all other mammals, including marsupials, also are synthesized exclusively in the ovary by oocytes and/or follicle cells (Table 1). Similarly, VE proteins in amphibia (Kubo et al., 1997; Yang and Hedrick, 1997) and cyprinoid fish (e.g., zebrafish, goldfish, and carp) (Chang et al., 1996, 1997).

#### Table 1. Synthesis of ZP proteins in mammals.

<table>
<thead>
<tr>
<th>TYPE OF MAMMAL</th>
<th>SITE OF SYNTHESIS</th>
<th>REPRESENTATIVE REFERENCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cat (Felis catus)</td>
<td>Ovary</td>
<td>Jewgenow and Fickel, 1999</td>
</tr>
<tr>
<td>Cow (Bos taurus)</td>
<td>Ovary</td>
<td>Noguchi et al., 1994; Topper et al., 1997; Totzauer et al., 1998</td>
</tr>
<tr>
<td>Dog (Canis familiaris)</td>
<td>Ovary</td>
<td>Blackmore et al., 2004</td>
</tr>
<tr>
<td>Hamster (Mesocricetus auratus)</td>
<td>Ovary</td>
<td>Moller et al., 1990; Kinloch et al., 1990</td>
</tr>
<tr>
<td>Human (Homo sapiens)</td>
<td>Ovary</td>
<td>Chamberlin and Dean, 1990; Bauskin et al., 1999; Hughes and Barratt, 1999; Liefievre et al., 2004</td>
</tr>
<tr>
<td>Macaque (Macaca radiata)</td>
<td>Ovary</td>
<td>Kolluri et al., 1995; Gupta et al., 1997; Harris and Pierson, 2003</td>
</tr>
<tr>
<td>Marmoset (Callithrix jacchus)</td>
<td>Ovary</td>
<td>Thillai-Koothan et al., 1993; Bogner et al., 2004</td>
</tr>
<tr>
<td>Mouse (Mus musculus)</td>
<td>Ovary</td>
<td>Bleil and Wassarman, 1980; Greve et al., 1982; Shimizu et al., 1983; Philpott et al., 1987; Kinloch et al., 1988; Epifano et al., 1995</td>
</tr>
<tr>
<td>Pig (Sus scrofa)</td>
<td>Ovary</td>
<td>Dunbar et al., 1981; Nakano et al., 1987; Hedrick and Wardrip, 1987; Yurewicz et al., 1993; Taya et al., 1995</td>
</tr>
<tr>
<td>Possum (Trichosurus vulpecula)</td>
<td>Ovary</td>
<td>Mate and McCartney, 1998; Haines et al, 1999; McCartney and Mate, 1999; Voyle et al., 1999</td>
</tr>
<tr>
<td>Rabbit (Oryctolagus cuniculus)</td>
<td>Ovary</td>
<td>Dunbar et al., 1981; Schwobel et al., 1991; Lee et al., 1993</td>
</tr>
<tr>
<td>Rat (Rattus norvegicus)</td>
<td>Ovary</td>
<td>Hinsch et al., 1994; Akatsuka et al., 1998; Boja et al., 2005</td>
</tr>
</tbody>
</table>

#### Table 2. Synthesis of VE proteins in teleosts.

<table>
<thead>
<tr>
<th>TYPE OF FISH</th>
<th>SITE OF SYNTHESIS</th>
<th>REPRESENTATIVE REFERENCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arctic char (Salvelinus alpinus)</td>
<td>Liver</td>
<td>Westerland et al., 2001; Berg et al., 2004</td>
</tr>
<tr>
<td>Atlantic halibut (Hippoglossus hippoglossus)</td>
<td>Liver</td>
<td>Hyllner et al., 1994</td>
</tr>
<tr>
<td>Carp (Cyprinus carpio)</td>
<td>Ovary</td>
<td>Chang et al., 1996, 1997</td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Liver</td>
<td>Oppen-Bernsten et al., 1990, 1992, 1999</td>
</tr>
<tr>
<td>European sea bass (Dicentarchus labrax)</td>
<td>Liver</td>
<td>Scapigliati et al., 1994, 1999; Hyllner et al., 1995</td>
</tr>
<tr>
<td>Gillhead seabream (Sparus aurata)</td>
<td>Liver/Ovary</td>
<td>Hyllner et al., 1995; Del Giacco et al., 1998; Modig et al., 2006</td>
</tr>
<tr>
<td>Goldfish (Carassius auratus)</td>
<td>Ovary</td>
<td>Chang et al., 1997</td>
</tr>
<tr>
<td>Salmon (Oncorhynchus masou)</td>
<td>Liver</td>
<td>Fujita et al., 2002, 2004</td>
</tr>
<tr>
<td>Pipefish (Syngnathus scovelli)</td>
<td>Ovary</td>
<td>Begovac and Wallace, 1989</td>
</tr>
<tr>
<td>Rainbow trout (Oncorhynchus mykiss)</td>
<td>Liver</td>
<td>Hyllner and Haux, 1992; Hyllner et al., 2001; Brivio et al., 1991</td>
</tr>
<tr>
<td>Winter flounder (Pseudopleuronectes americanus)</td>
<td>Liver</td>
<td>Lyons et al., 1993</td>
</tr>
<tr>
<td>Zebrafish (Danio rerio)</td>
<td>Ovary</td>
<td>Bonsignorio et al., 1996; Wang and Gong, 1999; Del Giacco et al., 2000; Mold et al., 2001</td>
</tr>
</tbody>
</table>
1997; Wang and Gong, 1999; Mold et al., 2001) are synthesized in the ovary (Table 2). On the other hand, VE proteins in rainbow trout (Brivio et al., 1991; Hyllner and Haux, 1992) and a large number of other non-cyprinoid fish (e.g., winter flounder, seabream, cod, and medaka) (Hamazaki et al., 1989; Oppen-Bernsten et al., 1992; Lyons et al., 1993; Murata et al., 1997; Del Giacco et al., 1998) are synthesized in the liver under hormonal (estradiol-17β) control and then transported in the bloodstream to the ovary where they are incorporated into the VE of growing oocytes (Table 2). Therefore, the site of synthesis is a major difference when comparing rainbow trout VE and mouse ZP proteins. In this context, it is of interest to note that the avian (chicken and Japanese quail) egg VE consists of a few proteins, of which at least one is synthesized in the liver and others in the ovary (Waclowek et al., 1998; Takeuchi et al., 1999; Bausek et al., 2000; Pan et al., 2001).

It would appear that during the evolution of animal species the sites of synthesis of egg EC proteins have included the liver and ovary and, in the latter case, either growing oocytes or follicle cells, or both. Mammalian ZP proteins and amphibian VE proteins are synthesized in the ovary, whereas fish and bird VE proteins are synthesized in either the liver or ovary, or in both. It has been suggested that the acquisition of dual sites of synthesis (i.e., ovary and/or liver) of ZP-like proteins is the result of an ancient polyploidization event, followed by additional species-specific amplifications (Conner and Hughes, 2003).

Cloning of trout VE and mouse ZP proteins

To screen for rainbow trout VE proteins, a cDNA expression library was constructed from poly(A)+ RNA prepared from livers of juvenile male and female fish treated with estradiol-17β (Hyllner et al., 2001). Livers of males begin to synthesize VE proteins following treatment of males with estradiol-17β (Hyllner et al., 1991). Clones complementary to mRNAs that encode VEα, VEß, and VEγ were identified by using specific antisera and limited amino acid sequencing, and the primary structures of VE proteins were determined. The primary structures of mouse egg ZP proteins, ZP1, ZP2, and ZP3, were determined by both cDNA and genomic cloning (Ringuette et al., 1988; Kinloch et al., 1988; Liang et al., 1990; Epifano et al., 1995). ZP1, ZP2, and ZP3 are encoded by single-copy genes located on chromosomes 19, 7, and 5, respectively.

Characteristics of trout VE and mouse ZP proteins

Trout VE proteins

Rainbow trout VEα, VEß, and VEγ are synthesized as precursor polypeptides consisting of 564, 526, and 441 amino acids, respectively. Secreted polypeptides are shortened by removal of both an N-terminal signal sequence and a C-terminal propeptide. Consequently, secreted VEα, VEß, and VEγ polypeptides incorporated into the VE consist of 533, 493, and 417 amino acids, respectively. In addition, VEγ possesses a single Asn-(N-) linked oligosaccharide. In the end, VEα, VEß, and VEγ present in the VE have apparent Mr's of ~58, ~54, and ~47 kDa, respectively, under reducing conditions on SDS-PAGE.

Mouse ZP proteins

Mouse ZP1, ZP2, and ZP3 are synthesized as precursor polypeptides consisting of 623, 713, and 424 amino acids, respectively. Secreted ZP polypeptides are shortened by removal of both an N-terminal signal sequence and a C-terminal propeptide. Consequently, secreted ZP1, ZP2, and ZP3 polypeptides incorporated into the ZP consist of 528 (per polypeptide), 601, and 331 amino acids, respectively. In addition, ZP1, ZP2, and ZP3 are heterogeneously glycosylated (Ser/Thr- (O-) and Asn-(N-) linked oligosaccharides), sialylated, and sulfated. In the end, ZP1, ZP2, and ZP3 present in the ZP have apparent Mr's of ~200 (2 identical polypeptides connected by disulfides), ~120, and ~83 kDa, respectively, under non-reducing conditions on SDS-PAGE. These apparent Mr's reflect the extensive glycosylation of mouse ZP proteins (Wassarman, 1988) as compared to trout VE proteins. Evidence suggests that

**Fig. 1.** Schematic representation of mature rainbow trout VE and mouse ZP polypeptides. The term "mature" refers to the polypeptide after removal of the signal sequence from the precursor and cleavage of the precursor at the CFCS. Each polypeptide is divided into an N-terminal region (black), a ZP domain (ZPD; red), and a C-terminal region (blue); when present, the trefoil or P-domain is represented (P; yellow). Each colored region is presented as the percentage of the entire mature polypeptide. The percentages are indicated below each region of the schematic of the polypeptides. The N- and C-termini are indicated by N and C, respectively.
all three ZP proteins play a structural role and that, in addition, ZP3 and ZP2 serve as primary and secondary receptors for mouse sperm, respectively, during fertilization (Wassarman, 1999, 2005; Wassarman et al., 2001).

**ZP domain of VE and ZP proteins**

The ZP domain is a key feature of both rainbow trout VE and mouse ZP proteins (Fig. 1). First identified by Bork and Sander (1992), the ZP domain is a sequence of ~260 amino acids that contains 8 conserved Cys residues as 4 intramolecular disulfides. A ZP domain is characteristic of all vertebrate egg ECs described to date. A ZP domain is also found in protein components of the ECs of some invertebrate eggs, such as abalone (H. rufescens, H. corrugata, and H. fulgens) and ascidian (H. roretzi) eggs (Galindo et al., 2002; Sawada et al., 2002; Aagaard et al., 2006). However, the ZP domain is not restricted to egg ECs, but is found in hundreds of extracellular proteins having diverse functions from a wide variety of tissues in mammals, amphiibia, birds, fish, flies, worms, molluscs, and tunicates (Jovine et al., 2002a, 2005). Most proteins possessing a ZP domain are modular proteins (e.g., possessing CUB, EGF, or PAN domains) that have been shown to act as receptors and/or to have mechanical functions. While it is likely that these biological functions can be ascribed to sequences that lie outside the ZP domain (see below), this remains an issue of considerable interest.

Initially proposed as playing a major role in polymerization of extracellular proteins (Killick et al., 1995; Legan et al., 1997), it is now well documented that the ZP domain does indeed function as a "polymerization module" (Jovine et al., 2002b, 2006a,b). Various biochemical and functional information strongly suggests that the ZP domain consists of two independently folding subdomains, an N-terminal and a C-terminal subdomain (Jovine et al., 2005, 2006a,b). The N-terminal subdomain is responsible for polymerization of ZP domain proteins and the C-terminal subdomain apparently is responsible for regulating protein-protein interactions between ZP domain proteins (discussed below). It should be noted that several proteins consisting of only the N-terminal subdomain of the ZP domain have been identified in flies, mouse, and man (e.g., Papillote, Oosp1, and PLAC1), but none have been found consisting of only the C-terminal subdomain (Jovine et al., 2006b). This strongly suggests that the C-terminal subdomain is found only within the context of a complete ZP domain.

**Primary structure of trout VE and mouse ZP proteins**

**General features of trout VE and mouse ZP proteins**

In Fig. 1, rainbow trout VE and mouse ZP mature polypeptides (i.e., after removal of the signal sequence and proteolytic cleavage of precursor polypeptides at the consensus furin cleavage-site (CFCS) are presented schematically as percentages of the polypeptides represented by the N-terminal region, ZP domain, and C-terminal tail. Overall, VEα and VEβ are most similar to ZP1 and ZP2, whereas VEγ is most similar to ZP3. VEα, VEβ, and ZP1 all possess a trefoil (P-type) domain (Thim, 1989; Carr, 1992; Carr et al., 1994) and their ZP domain (Bork and Sander, 1992; Jovine et al., 2005) represents a significantly lower percentage of the mature polypeptide (~44-61%) than the ZP domain of VEγ and ZP3 (~73 and 78%, respectively). Similarly, the region N-terminal to the ZP domain of VEα, VEβ, ZP1, and ZP2 is a significantly higher percentage of the mature polypeptide (~39-55%) than the same region of VEγ and ZP3 (~21 and 7%, respectively). Finally, a region C-terminal to the ZP domain is either present minimally (~1%) or missing completely in VEα, VEβ, ZP1, and ZP2, but is present in VEγ and ZP3 (~6 and 15%, respectively).

**Trout VEγ and mouse ZP3**

As indicated above, trout VEγ is most similar to mouse ZP3. In Fig. 2, the amino acid sequences of trout VEγ and mouse ZP3 are compared. Overall, they share ~25% sequence identity. Since it can be estimated that the frequency of amino acid change per unit time in evolution is 0.1-1.0 amino acid/10⁶ residues per 10⁶ years (i.e., 0.1-1.0% sequence divergence/10⁶ years) (Doolittle, 1986), the extent of amino acid conservation for rainbow trout VEγ and mouse ZP3 (~25%) over ~400 million years is well within the expected range. Each polypeptide has an N-terminal signal sequence (22 amino acids), a ZP domain (~260 amino acids) containing 8 conserved Cys residues as 4 intramolecular disulfides (Cysγ₁-Cysγ₅, Cysγ₃-Cysγ₅, Cysγ₃-Cysγ₇, and Cysγ₇-Cysγ₈) (Boja et al., 2003; Dairie et al., 2004), and a CFCS (VEγ-ArgLysGlyArg; ZP3-ArgAsnArgArg) close to the C-terminus. VEγ also has a long N-terminal extension (72 amino acids) that is particular rich (~45% of extension residues) in Pro (19 residues) and Gln (14 residues) residues and ZP3 has a long (23 amino acids) hydrophobic peptide (transmembrane domain; TMD) close to the C-terminus. It is noteworthy that just downstream of the ZP domain of VEγ and ZP3 is a region containing 4 Cys residues in rather close proximity to one another (VEγ-Cysγ₁₉₉GlyCysγ₄₀₂AspSerThrCySγ₄₀₆; ZP3-Cys₃₂₁AspCyS₃₃२CyS₃₄₆SerHisGlyAsn-Cys₃₃₅). From such sequence comparisons it is apparent that rainbow trout VEγ and mouse ZP3 are closely related to each other.

**Trout VEα and VEβ and mouse ZP1 and ZP2**

A similar comparison of trout VEα and VEβ polypeptides reveals that, overall, they share ~65% sequence identity. Additionally, they each have a trefoil
or P-type domain that contains 6 conserved Cys residues as intramolecular disulfides (Cys₁–Cys₄, Cys₂–Cys₅, and Cys₃–Cys₆), long N-terminal extensions rich in Pro and Gln residues (>50% of extension residues), and a CFCS close to the C-terminus of the precursor polypeptides. When the primary structures of these polypeptides are compared with mouse ZP1 and ZP2, the overall sequence identity is quite low, although the ZP domain Cys residues, trefoil domain (present in VEα, VEß, and ZP1), and CFCS are conserved. Therefore, despite a relatively low sequence identity, it is clear that trout VEα and VEß are related to mouse ZP1 and ZP2.

**ZP domain of trout VE and mouse ZP proteins**

Aside from the conserved Cys residues, the primary structure of ZP domains from different proteins bear very little resemblance to one another (Jovine et al., 2005). In general, this is not unusual for structural domains. However, there are a relatively large number of amino acid positions along the polypeptides with conserved physiochemical character (polar residues, ~15%; small residues, ~15%; hydrophobic residues, ~7.5%). Currently it is not possible to reliably predict the conformation of the ZP domain, but it seems likely that it adopts a novel protein fold and that this fold is a common feature of all ZP domain proteins.

Insofar as Cys residues of the ZP domain of VEα, VEß, ZP1, and ZP2, while they have the same Cys₁–Cys₄, Cys₂–Cys₅, and Cys₃–Cys₆ disulfides as in VEγ and ZP3, their disulfides differ from VEγ and ZP3 in the second half of the ZP domain that contains two additional Cys residues (Cys₇ and Cys₈) (Boja et al., 2003; Darie et al., 2004). For example, for VEß and ZP2 the Cys residues are linked Cys₅–Cys₆, Cys₇–Cys₈, and Cys₉–Cys₁₀, not Cys₅–Cys₇, and Cys₉–Cys₁₀ for VEγ and ZP3. These different disulfide arrangements could explain why the presence of both ZP2 and ZP3 is required for assembly of the mouse ZP (Liu et al., 1996; Rankin et al., 1996, 2001) and this same situation could apply to VÉα/ß and VEγ in trout. Considering that VEγ and ZP3-like proteins always form heterocomplexes with VEα/ß and ZP1/2-like proteins, respectively, whereas the latter can also form complexes in the absence of VEγ and ZP3-like proteins, it is likely that Cys residue connectivity plays an important role in specifying recognition between ZP domain proteins.

**Trout VE protein ProGln-rich region**

It is known that fish VE proteins can heterodimerize in a covalent manner through their N-terminal ProGln-

---

**Fig. 2. Primary structures of rainbow trout VEγ and mouse ZP3.** Rainbow trout VEγ and mouse ZP3 consist of 441 and 424 amino acids, respectively. Changes in amino acids are indicated by using the single letter amino acid code. The amino acid sequences depicted by different colors include the signal sequence (1-22; pink), ProGln-rich region of VEγ (23-94; brown), ZP domain (blue), Cys residues (red), CFCS (green), TMD of ZP3 (386-409; aqua), and remainder of the polypeptides (black). VEγ and ZP3 share ~25% sequence identity.
rich region (Fig. 2) (Darie et al., 2004, 2005). This takes place by a reaction catalyzed by a transglutaminase (i.e., formation of an amide bond between the $\gamma$-carbonyl of Gln residues and the $\epsilon$-amino group of Lys residues with release of NH$_3$) (Oppen-Bernsten et al., 1990). Recently, it was proposed that formation of the outer layer of the carp egg fertilization envelope involves cross-linking of two VE proteins (called ZP2 and 3) by transglutaminase (Chang et al., 2002). Mammalian ZP proteins lack an N-terminal ProGln-rich region and covalently linked heterodimers between ZP proteins have not been detected in either eggs or embryos (Wassarman, 1988).

**Assembly of trout VE and mouse ZP proteins**

There are significant differences between VE assembly in rainbow trout and ZP assembly in mice. For example, ZP protein precursors are synthesized in the mouse ovary by growing oocytes and have a C-terminal TMD that anchors the proteins in egg plasma membrane prior to secretion and assembly (Jovine et al., 2004). On the other hand, rainbow trout VE protein precursors are synthesized in the liver, are transported to the ovary, and do not have a TMD (Darie et al., 2004, 2005). [However, it should be noted that in the case of two other fish, carp and zebrafish, whose egg VE proteins are synthesized in the ovary, there is a TMD within the C-terminal region of the proteins (Chang et al., 1997; Mold et al., 2001).] Despite the differences, nascent mammalian ZP proteins and fish VE proteins are deposited on the inside margin of the respective ECs (Hamazaki et al., 1989; Qi et al., 2002). Interestingly, when ZP proteins are truncated just upstream of the TMD so that they resemble trout VE proteins, the proteins lacking a TMD are secreted, but are not cleaved at the CFCs or incorporate into the ZP (Jovine et al., 2002b, 2004). The evidence suggests that the TMD is not involved in specific interactions, but ensures proper localization and/or topological orientation of nascent proteins so that proteolytic processing and assembly can take place. Unlike ZP precursor proteins, VE precursor proteins that lack a TMD undergo both cleavage at the CFCs and assembly into the VE once they reach the ovary (Sugiyama et al., 1999; Darie et al., 2005).

Like ZP protein precursors, VE precursor polypeptides possess both an external hydrophobic patch (EHP) in the C-terminal tail and an internal hydrophobic patch (IHP) in the ZP domain (Table 3). The presence of an IHP and EHP appears to be characteristic of all ZP domain proteins and are predicted to form $\beta$-strands (Jovine et al., 2004). These elements could prevent premature polymerization of VE proteins into filaments in the bloodstream in the manner proposed for ZP proteins (Jovine et al., 2004, 2005, 2006a). As a result of proteolytic processing at the CFCs in the ovary (Darie et al., 2005), the EHP is lost with the C-terminal tail, concomitantly the IHP in the ZP domain is exposed, and mature VE proteins are rendered able to polymerize around eggs (Fig. 3; Jovine et al., 2004, 2005, 2006a). A similar proteolytic processing mechanism that regulates polymerization of proteins has been reported for several other kinds of proteins as well (Taylor et al., 1997; Bourne et al., 2000; Handford et al., 2000; Mosesson et al., 2001).

**Final comments**

Although they are highly related groups of proteins possessing a ZP domain, it is likely that the biological roles of fish VE proteins and mouse ZP proteins differ somewhat. Since fish sperm are attracted to and enter a micropyle on the egg, fish VE proteins may play solely a structural role during fertilization. On the other hand, mouse sperm bind to and then penetrate the egg ZP in a relatively species-restricted manner. This suggests that mammalian ZP proteins play not only a structural role, but also serve as receptors for homologous sperm during the initial steps of fertilization. Following fertilization, both fish VE proteins and mouse ZP proteins are modified appropriately by cortical granule proteases, transglutaminases, and/or other enzymes in order to provide a block to polyspermy.

Many questions about VE and ZP assembly and structure remain unanswered. Among these is the question of how some fish VE protein precursors (e.g., in rainbow trout) are targeted specifically to the ovary from the bloodstream. In an analogous situation in fish, birds, and amphibians, the yolk precursor protein, vitellogenin, is synthesized and secreted by the liver and transported in the bloodstream to the ovary (Wallace,

**Table 3.** EHP and IHP of mouse ZP and trout VE proteins.

<table>
<thead>
<tr>
<th></th>
<th>EHP</th>
<th>IHP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZP1</td>
<td>Pro-Gly-Ala-Val-Gly-Phe-Glu$_{572}$</td>
<td>Ser-Gly-Pro-Leu-Arg-Leu-Glu$_{400}$</td>
</tr>
<tr>
<td>ZP2</td>
<td>Pro-Gly-Pro-Ile-Leu-Leu-Leu$_{64}$</td>
<td>Pro-Gly-Pro-Leu-Val-Leu-Val$_{499}$</td>
</tr>
<tr>
<td>ZP3</td>
<td>Val-Gly-Pro-Leu-Ile-Phe-Glu$_{699}$</td>
<td>Glu-Glu-Lys-Ala-Leu-Arg$_{1172}$</td>
</tr>
<tr>
<td>Trout</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VE$_{\alpha}$</td>
<td>Ser-Glu-Lys-Val-Ile-Met-Ile$_{557}$</td>
<td>Asp-Ala-Val-Leu-Asp-Glu$_{576}$</td>
</tr>
<tr>
<td>VE$_{\beta}$</td>
<td>Ser-Gly-Gln-Leu-Ile-Thr$_{317}$</td>
<td>Pro-Gly-Pro-Leu-Ile-Val-Glu$_{336}$</td>
</tr>
<tr>
<td>VE$_{\gamma}$</td>
<td>Leu-Gly-Pro-Ile-Phe-Ile-Ser$_{436}$</td>
<td>Glu-Glu-Leu-Leu-Tyr-Phe-Ser$_{233}$</td>
</tr>
</tbody>
</table>
There it is taken up into growing oocytes in a receptor-mediated fashion by micropinocytosis. In the chicken, vitellogenin receptors also import very-low-density lipoprotein, riboflavin-binding protein, and α-2-macroglobulin into growing oocytes (Schneider, 1996). Whether or not a similar receptor-mediated mechanism applies to uptake of fish VE protein precursors remains to be determined. In a similar vein, it will be of interest to determine the source of the furin-like enzyme that cleaves VE precursor proteins at the CFCS in the ovary. Possibly the enzyme is associated with oocyte plasma membrane that is close to the innermost layer of the VE into which nascent, processed VE proteins are incorporated.

Much is known about the structure of the mouse egg ZP. For example, the ZP is composed of long filaments (~70 Å in width) containing a ZP2/3 repeat (every 140 Å or so) and the filaments are interconnected by ZP1 (Greve and Wassarman, 1985; Wassarman and Mortillo, 1991; Green, 1997). However, recent evidence suggests that each filament may actually consist of two protofilaments wound tightly around each other, as reported for another ZP domain protein, uromodulin (Jovine et al., 2002b), and the location of the crosslinker ZP1 (either incorporated into the filaments with ZP2 and ZP3 or simply associated with the surface of filaments)

**Fig. 3.** A general mechanism for assembly of ZP domain proteins, including rainbow trout VE and mouse ZP proteins. In all ZP domain precursors, the ZP domain is followed by a C-terminal propeptide that contains a basic cleavage site (such as a CFCS), and EHP, and, in most cases, a TMD or GPI-anchor (top panel). Precursors do not polymerize within the cell either as a result of direct interaction between the EHP and IHP or because they adopt an inactive conformation dependent on the presence of both patches (middle left panel). C-Terminal processing at the CFCS by a proprotein convertase (middle right panel) would lead to dissociation of mature proteins from the EHP (bottom left panel), activating them for assembly into filaments and matrices (bottom right panel).
remains to be determined. In this context, it is clear that results of high-resolution imaging, X-ray crystallographic, and NMR studies of ZP proteins and filaments could have significant impact upon our understanding of structure-function relationships for these proteins. Hopefully, such structural information will become available in the near future.

The ECs of vertebrate and invertebrate eggs continue to be of great interest to investigators because of their very important biological functions. Clearly, key features of the unique class of proteins (i.e., ZP domain proteins) that constitutes egg ECs have been conserved over many hundreds of millions of years of evolution. Further research on these proteins is bound to reveal important new information about their expression, synthesis, assembly, and functions during animal development.

Acknowledgements. We are very grateful to our recent colleagues, Drs. Luca Jovine and Costel Darie, for enhancing our understanding of the relationships between trout VE proteins and mouse ZP proteins. We also thank Norman Soule and the staff of the Cold Spring Harbor Fish Hatchery for providing us with generous hospitality and buckets of trout eggs. We had to limit the number of references in the review and apologize for all omissions. Our research was supported in part by National Institutes of Health Grant HD-35105.

References


**Egg extracellular coat proteins**


Egg extracellular coat proteins


Egg extracellular coat proteins


Accepted October 16, 2006