# SHORT COMMUNICATION

Luca Del Giacco · Silvia Diani · Franco Cotelli

# Identification and spatial distribution of the mRNA encoding an egg envelope component of the Cyprinid zebrafish, *Danio rerio*, homologous to the mammalian ZP3 (ZPC)

Received: 1 April 1999 / Accepted: 8 June 1999

Abstract Using degenerate reverse transcription polymerase chain reaction (RT-PCR) techniques we have isolated a cDNA encoding a putative component of the zebrafish Danio rerio egg chorion, homologous to the mammalian ZP3 (ZPC). The predicted protein (zfZPC) has a calculated molecular mass of 58.4 kDa and contains a signal peptide (located in the N-terminal region) composed of 11 hydrophobic amino acid residues followed by a signal peptide cleavage site. The zfZPC contains the ZP domain, a characteristic amino acid sequence shared by all ZP proteins of the mammalian zona pellucida and of both amphibian and bird egg envelope components. The zfZPC also exhibits certain unique features including five N-terminal Q-rich tandem repeats presumably involved in the hardening of the chorion after the fertilization of the egg and a long C-terminal tail containing two potential sites of N-linked type glycosylation. RT-PCR and in situ hybridization revealed a restricted pattern of tissue distribution: the gene encoding zfZPC is transcribed only in the growing oocyte of sexually mature female fish.

**Key words** Fish  $\cdot$  Egg envelope  $\cdot$  ZP domain  $\cdot$  N-terminal repeats

# Introduction

The teleost egg envelope – the chorion – is an extracellular matrix that surrounds the fully grown oocyte. The chorion, like the egg envelope of all the other animals, participates in the main fertilization steps (regulation of the sperm-egg interaction and prevention of polyspermy) and protects the embryo during the successive stages of development. The chorion is a thick structure of protein

Edited by J. Campos-Ortega

L. Del Giacco · S. Diani · F. Cotelli (💌) Dipartimento di Biologia, Università di Milano, via Celoria, 26, I-20133 Milan, Italy e-mail: franco.cotelli@unimi.it and glycoprotein nature (Cotelli et al. 1988). While the structural complexity and macromolecular composition differ from species to species, the chorion is in general composed of a wave-shaped fibrillar component embedded in an amorphous matrix. The structure is normally multilayered, with the exception of the monolaminar chorion of salmonids (Brivio et al. 1991).

Transmission electron microscopy of the zebrafish *Danio rerio* egg chorion showed that it is an acellular three-layered envelope, of which 75–80% is represented by the inner layer (Bonsignorio et al. 1996). SDS-PAGE of the protein and glycoprotein pattern of the zebrafish egg chorion revealed the presence of at least 20 bands (Bonsignorio et al. 1996) confirming, as in other cyprinids, the highly complex nature of the egg envelope (Cotelli et al. 1988; Chang et al. 1996). Furthermore, assays performed with lectins indicated the highly glycosylated nature of the majority of the zebrafish egg chorion components. All the glycoproteins detected belong to the *N*-linked type, and *O*-linked glycosylations are not present (Bonsignorio et al. 1996).

Recently, cDNAs encoding chorion components in the teleosts winter flounder Pseudopleuronectes americanus (Lyons et al. 1993), medaka Oryzias latipes (Murata et al. 1997; Sugiyama et al. 1998), carp Cyprinus carpio (Chang et al. 1997) and gilthead sea bream Sparus aurata (Del Giacco et al. 1998) have been identified, and the amino acid sequences of these components show a primary structure homologous to ZP1(ZPB), ZP2(ZPA) and ZP3(ZPC), the three major glycoproteins composing the zona pellucida, the envelope surrounding the mammalian egg. Moreover, they show homology with some of the components of the egg envelope of other species of vertebrates, including Xenopus laevis (Kubo et al. 1997, Yang et al. 1997) and birds (Waclawek et al. 1998). In particular, all these components are characterized by a conserved motif named the ZP domain, first identified by Bork and Sander (Bork and Sander 1992). This domain is a common feature shared by different filament forming proteins, such as ZP1, ZP2 and ZP3, composing the fibrillogranular material forming the zona pellucida (Greve and Wassarman 1985), or the Tectorins, the glycoprotein component of the tectorial membrane of the inner ear of the mouse and the chicken (Killick et al. 1995; Legan et al. 1997).

The sites and the regulation of the synthesis of the chorion components vary from species to species. In some species of teleosts some chorion components are synthesized by the liver under estrogen estradiol- $17\beta$  regulation (Sugiyama et al. 1998; Del Giacco et al. 1998), while the ovarian origin of some of the egg envelope components has been demonstrated in some other species (Chang et al. 1997).

This evidence indicates that the teleosts are the only group of vertebrates for which an extraovarian organ is involved in the construction of the egg envelope; for all the other vertebrates studied until now, the expression of the genes encoding egg envelope components is restricted to the ovary (Kubo et al. 1997, Yang et al. 1997; for a review see Wassarman 1990).

We isolated three different mRNAs encoding zebrafish egg chorion components, and examine the spatial distribution of one of them (zfZPC mRNA), demonstrating its presence in only the oocyte.

# **Materials and methods**

## Animals

Sexually mature specimens of zebrafish (*Danio rerio*) were maintained in aquaria in our laboratories under the previously described conditions (Westerfield 1996). Male fish were reared for1–10 days under normal conditions in estradiol-17 $\beta$  (100 mg/l)-containing water or, alternatively, injected with estradiol-17 $\beta$  dissolved in cocoa butter as previously described (Del Giacco et al. 1998).

#### RNA isolation

Total RNA was isolated from the liver, the ovary and the muscle of sexually mature females and from the liver of normal and estradiol-17 $\beta$ -treated adult males of *Danio rerio* using a nucleic acid extractor (RNAzol B, Cinna/Biotecx).

#### Primers

YZTT: 5'-caggaaacagctatgacTAAGCTTGTCGACG(T)17-3'. Y Z -Y: 5'-TGACTAAGCTTGTCGACG-3'; Z: 5'-CAGGAAACAGC TATGAC-3'; Z1: 5'-CT<u>GGATCC</u>NGACCTCACYYTNGG-3' (409-424). *Bam*HI site underlined. Z2: 5'-GT<u>GAATTC</u>AR-RTGRCAGGTGA TRTA-3' (1039-1056); EcoRI site underlined forward zfZPC primers: Zeb0: 5'-TTGTGTTCATCGACTG-GTG3′ (834–853); Zeb1: 5′-CTGATG TGGATGCCAGTGA-3′ (870–889); Zeb2: 5′-TATAAAGCATCATGGATG-3′ (903–921); Zeb3: 5'-CCTCTCCAGTCCAGCA GC-3' (1221-1239); reverse zfZPC primers: ZebAR: 5'-GACAGGTATGGCTGCACCA-3' (617-598); ZebBR: 5'-ACCGTC TCCGGTATTGGC-3' (638-620); 5'-AGCAGGCCTCAC TGCTGCGC-3' (669-649);ZebCR: ZebDR: 5'-CAAATGGAACCCAAG CAG-3' (683-665); zfAS: 5'-TTGAGATTTTATAAAACATTTTATT-3' (1736-1711). Vitellogenin primers: ZfVtg-forward: 5'-CTGGCTTTGAACAAAT-GCAG-3'; ZfVtg-reverse: 5'-TCTGGA TTGATGGGAACAGC-3'.

All the positions of the zfZPC-specific primers are referred to the sequence submitted to the GenBank/EMBL database under the accession number U55863; the vitellogenin primers were designed on the basis of the partial cDNA clone submitted to the NCBI GenBank/EMBL database under the accession number AA658675.

## Cloning

#### cDNA internal fragment isolation

The isolation of the cDNA fragment corresponding to the mRNA region encoding the zfZPC ZP domain was carried out as previously described (Del Giacco et al. 1998).

#### 3' end isolation of cDNA (RACE)

The 3' end isolation of the cDNA was carried out as previously described (Del Giacco et al. 1998) on the basis of existing protocols (Frohman et al. 1988).

## 5' end isolation of cDNA (RACE)

The 5' end isolation of the cDNA was carried out as described previously (Del Giacco et al. 1998). Briefly, the reverse transcription (RT)was carried out using 200 ng of primer ZebDR. A 1-µl aliquot of the RT reaction mixture was used as template in the second-strand synthesis reaction using the following conditions: 95°C for 2 min, room temperature for 5 min, 72°C for 20 min (1 cycle); the residual ZebDR used in RT reaction was enough to promote DNA amplification at low stringency annealing temperature (room temperature) recognizing its specific site and nonspecific sites distributed on the cDNA. An aliquot of this reaction (1 µl) was amplified with ZebDR primer and specific primer ZebCR using standard PCR conditions. The specific band, obtained after two rounds of subsequently semi-nested PCR performed with primers ZebDR/ZebBR and ZebDR/ZebAR, was gel extracted by QIAquick gel extraction kit (QIAGEN), cloned into pMOSBlue T-vector (Amersham) and sequenced.

Expression analyses

## RT-PCR

For RT, 2  $\mu$ g of RNA were reverse transcribed in presence of 1  $\mu$ g of random primers. PCR was carried out in presence of primers Zeb0 and ZebAS using as template 1  $\mu$ l of the first-strand reaction mixture. Rounds of semi-nested PCRs were performed using Zeb2/ZebAS and Zeb3/ZebAS primers.

To control the efficiency of the procedures of estrogenization we investigated the presence of the mRNA encoding vitellogenin in the samples corresponding to the liver of estrogenized male fish using the primers zfVtgF and zfVtgR. To avoid false-positive signals due to genomic contamination of the RNA samples, a control PCR was also performed using genomic DNA as template. The absence of reaction products of the expected size suggests the presence of introns between the primers used for the PCR and allows discrimination between cDNA and genomic DNA (data not shown).

## In situ hybridization

PCR double strand Zeb3/ZebAS cDNA fragment was gel extracted and used as the template in PCR containing digoxigenin-11dUTP dNTP mixture, where either only the sense (primer Zeb3) or the antisense (primer ZebAS) oligonucleotide was used to prime the reaction. The amount of the single-strand cDNA fragments obtained was quantified by agarose gel electrophoresis and spectophotometric analysis. Ovary specimens from sexually mature female *Danio rerio* were fixed in 4% paraformaldehyde in  $2 \times PBS$  and routine histological techniques were employed for paraffin embedding. Sections, 8 µm thick, were deparaffinized by two washes of 10 min each in xylene and rehydrated with ethanol (100–50%). The hybridization was then performed as described previously (Del Giacco et al. 1998).

# **Results and discussion**

Previous results indicate that the synthesis of the components of the fish egg envelope is regulated by two different control mechanisms involving two different organs, the oocyte and the liver (see Introduction). Therefore, we extracted total RNA from these two organs of a sexually mature Danio rerio female fish. The total RNA was then reverse-transcribed and the cDNA subjected to PCR using a pair of degenerate primers, Z1 and Z2, derived from the most conserved regions among the ZP domain of ZPCs of different vertebrates (Del Giacco et al. 1998). Using this procedure we obtained an mRNA of 1744 nucleotides [poly(A) tail excluded] (the sequence has been submitted to the GenBank/EMBL database under the accession number U55863) encoding a protein that we named zfZPC. The nucleotide sequence has the AUG start codon located between nucleotides 47 and 49. We deduced that this codon was the initiation codon on the basis of the (-3)ANNAUGG(+4) canonical motif (Kozak 1991). The A in position –3 and the G in position +4 have the strongest positive effect on the translation efficiency; our mRNA sequence follows this rule. Furthermore, the sequence contains a stop codon (UAG) in frame with the predicted first AUG codon, 15 nucleotides upstream of the open reading frame indicating that the 47-49 AUG codon is the start codon. The mRNA has a UAA stop translation codon at nucleotide 1643. The 3' UTR is 92 nucleotides long, with one polyadenylation signal (AA-TAAA) located between nucleotides 1712 and 1717.

The mRNA encodes a protein, zfZPC, with a predicted molecular mass of 58.4 kDa. The protein shows a putative signal peptide (located in the N-terminal region), a common feature of secreted proteins. The signal is composed of 11 hydrophobic amino acid residues (AWFISVFLIAE) followed by a putative signal peptide cleavage site, deduced on the basis of the -3, -1 rule (von Heijne 1986). The cleavage could occur between (-1)Ser21 and (+1)Arg22, giving a resultant protein of 512 residues with a predicted molecular mass of 56.1 kDa.

The putative protein (named zfZPC) contains the ZP domain (residues 94–373; Fig. 1), a region that characterizes most of the egg envelope components of the vertebrates studied until now (see Introduction). The domain contains 8 cysteine residues characteristic of ZP domains (10–12 limiting the analysis only to the ZP domain of the ZPC homologues; Fig. 1) and conserved hydrophobic and polar patterns (Bork and Sander 1992), which play an important role in the secondary and tertiary structures of the ZP proteins. Moreover, the presence of the cysteines strongly suggests the formation of disul-

fide bonds. Although their formation has still to be demonstrated, previous work described the presence of intradisulfide bridges in many of the egg chorion components of different species of teleosts (Brivio et al. 1991), zebrafish included (Bonsignorio et al. 1996).

Several authors have indicated that the ZP domain is the main structural unit in the assembly of the fibrillar structures in specialized extracellular matrices. Indeed, ZP domain containing proteins, such as zona pellucida glycoproteins,  $\alpha$ - and  $\beta$ -tectorins, GP2 and uromodulin, can form extracellular matrices or gels (Killick et al. 1995).

Even if there is a remarkable similarity of amino acid sequence between zfZPC and the other ZPC homologues, there are also some significant differences. For example, it is important to note that the ZP domains of all the ZPC homologues identified until now show a putative site of *N*-glycosylation, Asn-X-Ser, but zfZPC lacks this signal (Fig. 1). Our previous work demonstrated that glycosylations of the *O*-linked type do not occur in zebrafish (Bonsignorio et al. 1996).

The N-terminal region of the mature zfZPC contains five XFQ repeats, where in four of the five cases X is glutamine; moreover, this repeat region is characterized by four longer tandem repeats composed of seven amino acid residues, (+1)XFQAAVX(+7), where X at position +1 is glutamine in three of the four cases, and the X in position +7 is proline in three of the four cases. The fourth repeat is tandemly followed by the amino acid motif QFQVTIP, quite similar to the first four repeats. zfZPC is the only ZPC homologue that contains repeated amino acid sequences. In fact, only the fish ZPA-B homologues identified in winter flounder (Lyons et al. 1993), carp, goldfish (Chang et al. 1997) and medaka (Murata et al. 1997; Sugiyama et al. 1998) contain N-terminal repeated sequences, although the tandem repeats are more complex and more frequently represented than in the zfZPC. The function of these repeats is not clear, but some previous studies suggest their involvement in the hardening of the fish egg chorion involving the formation of intermolecular cross-links between the  $\gamma$ -carbonyl group of the glutamine and the  $\epsilon$ amino group of lysine after the fertilization (Oppen-Berntsen et al. 1990). Consequently, zfZPC contains not only the typical ZPC ZP domain but also shares the presence of Nterminal tandem repeats with the fish ZPA-B homologues.

Preliminary data obtained in our laboratory indicate the presence of additional ZPC homologs in zebrafish, as also found in the teleosts carp (Chang et al. 1996) and gilthead sea bream (Del Giacco et al., unpublished results). Through mechanisms such as gene duplication or tetraploidy, a family of *zpc* genes originated from a common ancestor teleost *zpc* gene. These genes could encode proteins with slightly different structural roles in the formation of the complex fibrillar matrix of the egg chorion. In fact, as often occurs after gene duplication, one or more of the homologs are free to mutate, gaining or losing function(s). Our data suggest that this may be the case of the *z*fZPC encoding gene: firstly, the ZP domain of *z*fZPC shows approximately 36–40% homology to the domain of the ZPCs of the other fish as well as to amphibians and

zfZPC saZPC ZI-3 caZP3 ccZP3 xlZPC mmZP3 hsZP3 ckZP3 consensus		94 VIVYCRAEAIELVINPDILAGGLPVFAEELWLGPEASPTCGVV   89 VAVECRERDAVVEVKKDMFGIGQFINPADLTLGNCAAV   81 VAVECREDLAHVEAKKDLFGIGQFINPADLTLGTCAAV   81 VAVECREDLAHVEAKKDLFGIGQFINPADLTLGTCPPS   118 VALQCGENRVHVEVKKDLFGIGQFINPADLTLGTCPVV   119 VAVQCSENRVHVEVKKDLFSNGQLLOPSGLSMGG
zfZPC saZPC ZI-3 caZP3 ccZP3 xlZPC mmZP3 hsZP3 ckZP3 consensus	137 127 119 157 158 105 82 82 82 103	S T G P G P L T I R A S L Q D C G T Q L S V N A D S L L Y S N V V V Y S P L P S P D G V G E D N NAH - V L I F Q A E L H Q C G S T L T T T D K S L L Y T F L L N Y N P Q P L G A A P A E D P A A Q - V L I F E S P L Q N C G S V L T M T E D S L V Y T F T L N Y N P K P L G S A P G L V P G S K - V L F F D N E L Q D C N S V L MM T K D E L V Y T F A L T Y T P E A F A G T P R E E D S A S G V L I F E Y E L Q D C N S V L MM T E N E L V Y T F A L T Y T P V A F A G T P P Q S T S T A - V L - F Q V A L Q D C G N S L Q M T P N F L V Y S T N L T Y A P A P R N V P I V S V D T D V - V R - F N A Q L H E C S S R V Q M T K D A L V Y S T F L L H D P R P V S G L S V S M D T E D V V R - F E Y G L H E C G N S M Q V T D D A L Y Y S T F L L H D P R P V G N L S S L N A A H N T V T - F A A G L H E C G S V V Q V T P D T L I Y R T L I N Y D P S P A S N P V L C L C L Y Y R T L I N Y D P S P A S N P V
zfZPC saZPC ZI-3 caZP3 ccZP3 xlZPC mmZP3 hsZP3 ckZP3 consensus	181   173   165   203   205   150   127   128   149	I Y T N G A A - I P VQ C Q Y R R Y S VD S A A V R P A W Y P FD A S Y S A T D Y L D F S L V V R T S S A V V I V E C H Y Q R H H N V S S L P L D P L W I P F S A V K V S E E F L Y F T L Y V R T S Q A V V I V E C H Y P R K H N V S S L A L D P L W V P F S A A K M A E E F L Y F T L I T R A G G A V I G V Q C H Y Q R F Q N V S G N A L K P T W V P Y A S T E A G E E V L L F S L I T R A E G A L V G Y Q C H Y Q R F E N V S S N A L K P I W V P Y A S T E V G E D L L Q F S L I R M G A K V L I - Q C F Y P R N G N V S S K A I K P T W V P F S S T I S A E D R L A F S L I R M N G A K V L I - Q C F Y P R N G N V S S K A I K P T W V P F S S T I S A E D R L A F S L I L R T N R V E V P I E C R Y P R Q G N V S S H P I Q P T W Y P F R A T V S S E E K L A F S L I V R T N R A E I P I E C R Y P R Q G N V S S Q A I L P T W L P F R T T Y F S E E K L T F S L I I R T N P A V I P I E C H Y P R R E N V S S N A I R P T W S P F K L R T V S R G E A V F S L C Y R V P W P
zfZPC saZPC ZI-3 caZP3 ccZP3 xtZPC mmZP3 hsZP3 ckZP3 consensus	227 220 212 251 252 196 174 175 196	R LMTDDWQEERGSNVEFLGDEIHLQAAVRLAYHLPLLVFIDWCVATP KLMTDDWQYERPIYQYFLGDIINIEASVKQFYHVPLRVYLDNCVATL KLTTDDEQFERPSYQYFLGDLIHIEATVKQYEHVPLRVYVDRCVATL KLMMDDWSFERPSNSYELGDVINVEASVKVYNHVPLRVFVDSCVATQ NLMMDDWSYQRPSNVYYLGDIINIEASVKVYNHVPLRVFVDSCVATR RLMTDDWSSPRASNLFQLGDIFHIEASINIANHAPMTIYADSCVATV RLMEENWNTEKSAPTFHLGEVAHLQAEVQTGSHLPLQLFVDHCVATP RLMEENWNAEKRSPTFHLGDAAHLQAEIHTGSHVPLRLFVDSCVAT LMSDDWSTERPFTGEQLGDILNIQAEVSTENHVPLRLFVDSCVAAL
zfZPC saZPC ZI-3 caZP3 ccZP3 xlZPC mmZP3 hsZP3 ckZP3 consensus	274 267 259 298 299 243 221 222 243	TP D YD A S E V K - Y SF I K - H H G C LAD S R S P N SN SM F M R - R T E G N H L N L Q S P D R - S S N P N Y A F I D N - H G C L Y D A R L T G S D S K F M - A R T A D N Y L R F Q S P D A - N S S P S Y A F I D N - Y G C L L D G R I T G S D S K F Y - S R P A E N K L D F Q V P D V - N A L P R Y L F I E N - H G C L V D A K V T A S S S R F M - P R S Q E D K I R F Q V P D V - N S V P R Y S E I E N - H G C L V D A K A T A S S S R F M - P R T Q E D K I R F Q T P D V - N S N P R Y S E I E N - H G C L V D G K L D D S S S A F R S P R P O P D K L Q F S S P L P D P N S S P - Y H F I V D F H G C L V D G K L D D S S A F R S P R P O P D K L Q F S S P L P D P N S S P - Y H F I V D F H G C L V D G - L S E S F S A F Q V P R P R P E T L Q F T T P D Q - N A S P - Y H T I V D F H G C L V D G - L S E S F S A F K V P R P G P D T L Q F T S P D G D - S S P H Y A I I - D F N G C L V D G R V D D T S S A F I T P R P R E D V L R F R P
zfZPC saZPC ZI-3 caZP3 cc2P3 xlZPC mmZP3 hsZP3 ckZP3 consensus	318 310 302 341 342 287 266 265 287	L D A F R F H K F T G N L V Y I H C H M K A L P A A Y S V SAK N R A C S F I D Q R W R S L E A F R F K G A D S G L L Y I T C H L K A T S V A Y P I D S D H R A C S Y I G G W K E L E A F R F Q G A D S G M I Y I T C H L K A T S A A Y P L D A E H R A C S Y I Q G W K E L E A F M F Q G G S S P S I Y M T C V L K A T L A S A P S D A L H K S C S F A N G W L A L E A F M F Q G G D S P S I Y M T C V L K A T L A S A P S D A L H K S C S F S N G W L A V D A F R F T T S D S A V L Y I T C T V K A T L A S A P S D A L H K S C S F S N G W L A V D A F R F T T S D S A V L Y I T C N L R A A T T Q V P D P M N K A C S F S K S A N S W S P V D V F H F A N S S R N T L Y I T C H L K V A P A N Q I P D K L N K A C S F N K T S Q S W L P V D V F H F A N D S R N M I Y I T C H L K V T L A E Q D P D E L N K A C S F S K P S N S W F P I D V F R F A G D N R N L I Y I T C H L K V T P A D Q G P D P Q N K A C S F N K A R N T W V P C C S W
zfZPC saZPC ZI-3 caZP3 ccZP3 xlZPC mmZP3 hsZP3 ckZP3 consensus	363 354 346 385 386 354 313 312 334	VD G N D DV C N T C EP S K P A S G A D A A C G S C E S G L P378 369V S G A D P L C A S C E S G C F A D G N H Q V C G C C D S T - C A D G N H Q A C G C C D S T - C369 361A D G N H Q V C G C C D S T - C A D G N H Q A C G C C D S T - C399 400A D G N H Q A C G C C D S T - C V E G D A D I C D C C S H G N C V E G R R D V C N C C C E T G N C349 327V E G R R D V C N C C C C D S T G N C V E G R R D V C N C C C T G N C349 328V E G R R D V C N C C C C N K G D C G C C C N K G D C349 327V E G R R D V C N C C C T G N C G C C C N K G D C349 328V E G R R D V C N C C C T G N C C C C C N K G D C349 349V E G R R D V C N C C C T G N C C C C C N K G D C349 349V E G R R D V C N C C C T G N C C C C N K G D C349 349V E G R R D V C N C C C T G N C C C C N K G D C349 349V E G R R D V C N C C C T G N C C C C N K G D C349 349V E G R R D V C N C C C T G N C C C C N K G D C349 349V E G R R D V C N C C C T G N C C C C N K G D C349 349V E G R R D V C N C C C T G N C C C C N C C C C N K G D C 349349V E G R R D V C N C C C T G N C C C C N C C C C C C C S N G N C



**Fig. 2** RT-PCR analysis of the distribution of the zfZPC and zfVtg (vitellogenin) transcripts. Ethidium bromide-stained gels of RT-PCRs performed with Zeb2/ZebAS (*upper panel*) or VtgF/VtgR (*lower panel*) primers: *lane* –negative control (containing no cDNA); *lane 1*liver from non-treated male zebrafish; *lane 2* full grown oocytes; *lanes 3–5* liver from estradiol-17β treated male zebrafish, *lane 3* 1-day treatment, *lane 4* 5-day treatment, *lane 5* 

10-day treatment; *lane 6* liver and *lane 7* ovary from sexually mature *Danio rerio* female fish. To avoid false-positive signals due to genomic contamination of the RNA samples, a control PCR was also performed using genomic DNA as template (data not shown). The presence of introns allows discrimination between cDNA and genomic DNA (see Results). Molecular weight markers (M)



**Fig. 3A–C** In situ hybridization of the ovary of a sexually mature *Danio rerio* female fish. **A**, **B** Digoxigenin-dUTP-labeled anti-sense zfZPC cDNA probe (see Materials and methods) was hybridized to formaldehyde-fixed, paraffin embedded tissue sections (7 µm); signals appear as *violet staining; black arrow* indicates a full-grown oo-cyte; *black arrowheads* indicate follicle cells. **C** No signals were present when a sense zfZPC probe was used as control. *Bars* **A**, **B** 100 µm, **C** 200 µm

mammals, suggesting a high mutation rate due to weak selective pressure on the zfZPC ZP domain. In addition, our preliminary results on two other ZPC clones that we have recently isolated suggest these homologues are more closely related to the other ZPCs than zfZPC; secondly, zfZPC is the only ZPC that contains the Q-rich tandem repeats in its N-terminal region, a characteristic feature of the fish ZPA-B; and thirdly, the C-terminal region of zfZPC is not homologous to any other ZP protein and contains two sites of glycosylation of *N*-linked type. These observations may indicate that events of recombination and/or mutation have put the ZP domain, the basic domain for the formation of the filaments, together with other regions having different functions (i.e., the suggested involvement of the Q-rich repeats in the chorion hardening or the possible role of the glycosylations in the protein-protein interactions that could be important to form the network of interconnected filaments).

The sites of biosynthesis of the fish egg chorion components vary with species. As already mentioned, previous studies indicate two organs involved in the chorion construction: the ovary and the liver. RT-PCR experiments were used to investigate the presence of the mRNA encoding zfZPC in the liver, the ovary and the eggs (obtained by gentle abdominal squeezing) of a spawning female and in the liver of estradiol-17 $\beta$ -treated and non-treated male fish. The efficiency of the procedures of estrogenization was investigated by RT-PCR for the presence of the mRNA encoding vitellogenin in the samples corresponding to the liver of spawning female and estrogenized male fish. In fact, the expression of the vitellogenin gene in zebrafish (Peute et al. 1978) and in general in all the teleosts occurred in the liver under the control of the estradiol- $17\beta$ . All the samples were vitellogenin mRNA positive, confirming the inducing activity of the estradiol-17 $\beta$  on the hepatocytes (Fig. 2).

The expression analysis of zfZPC indicated an organrestricted pattern (Fig. 2). In the case of zfZPC the mRNA is only detected in the oocytes. Using in situ hybridization we detected zfZPC mRNA in the oocytes at different stages of maturation (Fig. 3), but not in the fullgrown oocyte. The lack of signal is not a consequence of the increased dimensions of the oocyte and the subsequent dilution of the zfZPC mRNA (Fig. 2). In some species of teleosts, including Cyprinids carp and goldfish (Chang et al. 1996) and in other vertebrates such as Xenopus (Kubo et al. 1997; Yang et al. 1997) and mammals (Wassarman 1990; Epifano et al. 1995), the oocyte is the only site of synthesis of the ZPC, suggesting the existence of a common vertebrate ancestor where the primitive *zpc* gene was expressed only in the egg. By contrast, other species of teleosts probably evolved a new mechanism to synthesize chorion components: the expression of such genes became activated in the liver under the control of estradiol-17 $\beta$ . For this reason we extended the expression analysis to the liver of estrogenized male fish; the negative result confirmed the absence of the hepatic expression of the zfZPC gene.

More information about the complex mechanism of the choriogenesis will be available from screening of mutants in the pathways of the production of the egg envelope structure.

Acknowledgements We wish to thank Dr. Linda Cairns and Dr. Khalid Islam for their reading of the manuscript and helpful suggestions and Dr. Stefano 'Pupo' Duga for automated sequencing. This work was supported by IV Piano Triennale – Ministero per le Risorse Agricole and CNR Progetto Finalizzato Biotecnologie.

## References

- Bonsignorio D, Perego L, Del Giacco L, Cotelli F (1996) Structure and macromolecular composition of the zebrafish egg chorion. Zygote 4:101–108
- Bork P, Sander C (1992) A large domain common to sperm receptors (ZP2 and ZP3) and TGF- $\beta$  type III receptor. FEBS Lett 300:237–240

- Brivio MF, Bassi R, Cotelli F (1991) Identification and characterization of major component of the Oncorhynchus mykiss egg chorion. Mol Reprod Dev 28:85–93
- Chang YS, Wang SC, Tsao CC, Huang FL (1996) Molecular cloning, structural analysis and expression of carp ZP3 gene. Mol Reprod Dev 44:295–304
- Chang YS, Hsu CC, Wang SC, Tsao CC, Huang FL (1997) Molecular cloning, structural analysis and expression of carp ZP2 gene. Mol Reprod Dev 46:258–267
- Cotelli F, Andronico F, Brivio MF, Lora Lamia C (1988) Structure and composition of the fish egg chorion (*Carassius auratus*). J Ultrastruct Mol Struct Res 99:70–78
- Del Giacco L, Vanoni C, Bonsignorio D, Duga S, Mosconi G, Santucci A, Cotelli F (1998) Identification and spatial distribution of the mRNA encoding gp49 component of the gilthead sea bream, *Sparus aurata*, egg envelope. Mol Reprod Dev 4:101–108
- Epifano O, Liang L, Familari M, Moos MC Jr, Dean J (1995) Coordinate expression of the three zona pellucida genes during mouse oogenesis. Development 121:1947–1956
- Frohman MA, Dusk MK, Martin GR (1988) Rapid production of full-length cDNAs from rare transcript: amplification using a single gene-specific oligonucleotide primer. Proc Natl Acad Sci USA 85:8998–9002
- Greeve JM, Wassarman PM (1985) Mouse egg extracellular coat is a matrix of interconnetted filaments possessing a structural repeat. J Mol Biol 181:253–264
- Killick R, Legan PK, Malenczac C, Richardson GP (1995) Molecular cloning of chick β-tectorin, an extracellular matrix molecule of the inner ear. J Cell Biol 129:535–547
- Kozak M (1991) Structural features in eukaryotic mRNAs that modulate the initiation of translation. J Biol Chem 266: 19867–19870
- Kubo H, Kawano T, Tsubuki S, Kawashima S, Katagiri C, Suzuki A (1997) A major glycoprotein of *Xenopus* egg vitelline envelope, gp41, is a frog homolog of mammalian ZP3. Dev Growth Differ 39:405-417
- Legan PK, Ray A, Keen JN, Richardson GP (1997) The mouse tectorins: modular matrix proteins of the inner ear homologous to components of the sperm-egg interaction system. J Biol Chem 272:8791–8801
- Lyons CE, Payette KL, Price JL, Huang RC (1993) Expression and structural analysis of a teleost homologue of a mammalian zona pellucida gene. J Biol Chem 268:21351–21358
- Murata K, Sugiyama H, Yasamasu S, Iuchi I, Yasumasu I, Yamagami K (1997) Cloning of cDNA and estrogen-induced hepatic gene expression for choriogenin H, a precursor protein of the fish egg envelope (chorion). Proc Natl Acad Sci USA 94: 2050–2055
- Oppen-Berntsen DO, Helvik JV, Walther BT (1990) The major structural protein of cod (*Gadus morhua*) eggshell and protein crosslinking during teleost egg hardening. Dev Biol 137: 258–265
- Peute J, Gaag MA van der, Lambert JG (1978) Ultrastructure and lipid content of the liver of the zebrafish, *Brachydanio rerio*, related to vitellogenin synthesis. Cell Tissue Res 186:297–308
- Sugiyama H, Yasumasu S, Murata K, Iuchi I, Yamagami K (1998) The third egg envelope subunit in fish: cDNA cloning and analysis, and gene expression. Dev Growth Differ 40:35–45
- Von Heijne G (1986) A new method for predicting signal sequence cleavage sites. Nucleic Acids Res 14:4683–4690
- Waclawek M, Foisner R, Nimpf J, Schneider WJ (1998) The chicken homologue of zona pellucida protein-3 is synthesized by granulosa cells. Biol Reprod 59:1230–1239
- Wassarman PM (1990) Profile of mammalian sperm receptor<sup>1, 2</sup>. Development 108:1–17
- Westerfield M (1996) The zebrafish book: a guide for the laboratory use of zebrafish (*Danio rerio*). University of Oregon Press, Eugene
- Yang JC, Hedrick JL (1997) cDNA cloning and sequence analysis of the *Xenopus laevis* egg envelope glycoprotein gp43. Dev Growth Differ 39:457–467