J. Biochem. 125, 469-475 (1999)

Formation of Mature Egg Envelope Subunit Proteins from Their Precursors (Choriogenins) in the Fish, *Oryzias latipes*: Loss of Partial C-Terminal Sequences of the Choriogenins[1]

Hitoshi Sugiyama,*,†,[2] Kenji Murata,*,[3] Ichiro Iuchi,* Kohji Nomura,‡ and Kenjiro Yamagami*

*Life Science Institute, Sophia University, Tokyo 102-8554; †Department of Physiology, St. Marianna University School of Medicine, Kawasaki 216-8511, and ‡Department of Biochemistry, Tokyo Metropolitan Institute of Gerontology, Tokyo 173-0015

Received August 25, 1998; accepted November 19, 1998

The inner layer of egg envelope of the medaka, *Oryzias latipes*, comprises two major groups of glycoprotein subunits, ZI-1,2 and ZI-3. Their precursor proteins, choriogenin H (Chg H) and choriogenin L (Chg L), respectively, are synthesized in spawning female liver. In the present study, the primary structures of the precursors and the corresponding mature subunits were compared by peptide mapping and amino acid sequencing to find what difference in their molecular structures is relevant to the assembly of the soluble precursors into the insoluble inner layer. The primary structures of the solubilized subunits were mostly identical to those of the respective precursors, but they lacked C-terminal partial sequences that their precursors possessed, namely, ZI-1,2 subunit was shorter than Chg H by 34 amino acid residues and ZI-3 was shorter than Chg L by 27 residues. In addition, a consensus amino acid sequence, Arg-Lys-X-Arg, was found at the putative cleavage sites in the C-terminal region of the precursors. It is conjectured that the truncation of the precursor proteins is prerequisite for formation of mature chorion subunit proteins and their assembly into chorion.

Key words: choriogenin, fish egg envelope, limited cleavage, medaka (teleost fish), molecular architecture.

The envelope of fish eggs plays important roles, including interaction with spermatozoa at fertilization and protection of the developing embryo it encloses (*1*). The egg envelope of teleosts consists mainly of an inner layer, which is composed of three major subunit proteins (*2-6*). In the medaka, *Oryzias latipes*, the inner layer of the egg envelope consists of two major groups of glycoprotein subunits, ZI-1,2 and ZI-3. The former apparently comprises three similar subunit bands on SDS-PAGE, while the latter reveals a single band (*7-9*).

Immunochemical studies using antibodies specific to ZI-1,2 and ZI-3 strongly suggest that these glycoproteins are formed as two groups of precursor proteins. They are, namely, high-molecular-weight spawning female-specific substances [H-SF, now named choriogenin H, Chg H (*10, 11*)] and low-molecular-weight spawning female-specific substance [L-SF, now named choriogenin L, Chg L (*10, 11*)], respectively, in the hepatic cells under the influence of estradiol-17β. They are transported to the surface of the vitellogenic oocytes through the blood stream and assembled into the inner layer (*12-16*). Recently, cDNAs for H-SF (Chg H) and L-SF (Chg L) were cloned from liver cell mRNAs of vitellogenic medaka and analyzed to elucidate the structure of the protein moieties of the precursors (*9, 11*). Moreover, northern blot analyses using the cDNAs as probes revealed that mRNAs for both the precursor proteins were expressed in the liver cells of both the spawning female and male after administration of estradiol-17β (*9, 11*).

Thus, it is necessary to clarify the mechanism of molecular assembly of the soluble precursor proteins into the insoluble egg envelope by comparing the molecular structures of the precursor proteins with those of the assembled subunits. Although a major difficulty in biochemical analysis of egg envelopes of medaka as well as many other teleosts comes from their poor solubility in conventional buffers (*1*), it has been found that solubilization with guanidium chloride is suitable for the study of primary structure of the egg envelope subunits of medaka (*17*). In the present study, primary structures of the two egg envelope subunits, ZI-1,2 and ZI-3, of medaka were

[1] The present work was partly supported by Grants-in-Aid (Nos. 06454688 and 08680799) to K.Y. from the Ministry of Education, Science, Sports and Culture of Japan.

[2] To whom correspondence should be addressed at: Department of Physiology, St. Marianna University School of Medicine, 2-16-1 Sugao, Miyamae-ku, Kawasaki 216-8511. Tel: +81-44-977-8111 (Ext. 3405), Fax: +81-44-977-0172, E-mail: sugiyama@marianna-u.ac.jp

[3] Present address: Life Science Division, Lawrence Berkeley Laboratory, University of California, Berkeley, CA 94720, USA.

Abbreviations: Con A, concanavalin A; DBA, *Dolichos biflorus* agglutinin; LCA, *Lens culinaris* agglutinin; PNA, peanut agglutinin; PHA, *Phaseolus vulgaris* agglutinin; UAE-1, *Ulex europaeus* agglutinin I; WGA, wheat germ agglutinin.

compared with those of their precursor proteins, Chg H and Chg L. The results suggest that post-translational modification of the precursor proteins occurs in their C-terminal regions prior to their incorporation into the inner layer of egg envelope.

MATERIALS AND METHODS

Subunit Proteins and the Precursor Proteins—Unfertilized egg envelopes were solubilized with 7 M guanidium chloride (GuHCl)-3.4 mg/ml DTT-1 mM EDTA-0.5 M Tris-HCl (pH 8.5) (solubilizing buffer) and following *S*-carboxymethylation with iodoacetic acid as described in the previous paper (*18*). Then purification of the subunits of egg envelope was performed as described in the previous paper (*18*). The precursor proteins, choriogenins (Chg H and Chg L), were purified from the ascites fluid of male fish treated with estrogen according to the method of Hamazaki *et al.* (*7*), modified by Murata *et al.* (*8*). The purified choriogenins (Chgs) were dialyzed against distilled water and lyophilized. Each of the lyophilizate was dissolved in 500 μl of the solubilizing buffer, incubated for 2 h at room temperature, then *S*-carboxymethylated as described above. Purity of the proteins was confirmed by SDS-PAGE according to Laemmli (*19*) after removing GuHCl by dialysis. The purified egg envelope subunits and Chgs were dialyzed against the buffers used for each enzymatic digestion. Amount of proteins was quantified by use of BCA Protein Assay Reagent (Pierce, Rockford, USA) using bovine serum albumin dissolved in each buffer as standard.

Peptide Mapping—Approximately 50-100 μg of the reduced and *S*-carboxymethylated (rcm) ZI-3 and rcmChg L were digested with 20 pmol lysylendoprotease, Lys-C (Endoprotease, Lys-C; Promega), in 0.2 ml of buffer [1 mM EDTA-50 mM Tris-HCl (pH 8.9)] at 37°C for 16-20 h. Approximately 80-160 μg of the rcmZI-1,2 and rcmChg H were digested with 40 pmol of V8-protease (Endoprotease, Glu-C; Promega) in 0.2 ml of buffer [0.3 M GuHCl-50 mM Tris-HCl (pH 8.0)] at 37°C for 18-24 h. Each digest was diluted with 1 ml of 0.15% trifluoroacetic acid (TFA) and fractionated by HPLC with a reverse phase column (RP-HPLC, Shodex C18-5A, Showa Denko K.K., Tokyo). The elution was carried out with a linear gradient of 0-80% acetonitrile in 0.15% TFA for 80 min at a flow rate of 1 ml/min, and the elution pattern was monitored by absorbance at 215 nm.

Amino Acid Sequencing—The proteins and peptides isolated by RP-HPLC were lyophilized. Amino acid sequences were analyzed with an automatic gas-phase protein sequencer (PSQ-1, Shimadzu, Kyoto) by Edman degradation in an automatic mode as recommended by the manufacturer. The amino acid sequences of Chg H and Chg L were identical to the data of the amino acid sequences deduced from the respective cDNAs (*9, 11*).

Analysis of C-Terminus—The C-terminal amino acids of rcmZI-3 and Chg L (non-reduced and *S*-carboxymethylated) were analyzed with carboxypeptidase (CPase, Worthington Biochemical, Washington, DC). About 4 mg of rcmZI-3 was dissolved in 250 μl of 1.0% SDS-0.1 M triethanolamine-HCl (pH 8.0) and incubated overnight at room temperature, then 750 μl of 0.1 M triethanolamine-HCl (pH 8.0) was added. As an internal standard, norleucine at a final concentration of 40 nmol/ml was added to the solution. The rcmZI-3 and Chg L were treated with 1/150 mol CPase A alone or together with 1/300 mol CPase B against the amount of the substrate. The amount of the CPases was calculated based on the absorption coefficient of CPase A ($A^{1\%}_{280}=19.4$) and CPase B ($A^{1\%}_{280}=21.4$). The digestion was performed at 25°C for various periods of time as shown in Fig. 6. The reaction was stopped by the addition of 75 μl of 1 N HCl to 75 μl of the sample. The supernatant after centrifugation at 10,000×*g* for 15 min was lyophilized and subjected to amino acid analysis on a JLC-300 amino acid analyzer (JEOL).

Detection of Sugar Chain and Deglycosylation—For detection of sugar chain(s) of the purified proteins or isolated peptides, various peroxidase-conjugated lectins (Con A, DBA, LCA, PHA-E4, PNA, UEA-1, WGA; Honen, Tokyo) were used. The purified proteins or isolated peptides were reacted with the lectins in a 96-well plate for enzyme-linked immunosorbent assay. The rcmZI-3, rcm-Chg L and isolated peptides in 50 mM Tris-HCl (pH 7.5) were immobilized to the well wall by incubation for 3 h at room temperature. The wells were washed three times with a washing buffer [150 mM NaCl-20 mM Tris-HCl (pH 7.2)-1% Tween-20], then about 50 μl of the washing buffer containing 150 ng of either DBA, PHA-E4, PNA or UEA-1, about 50 μl of the washing buffer containing 150 ng of either Con A or LCA together with 1 mM $CaCl_2$, 1 mM $MgSO_4$, and 1 mM $ZnCl_2$, or about 50 μl of 500 mM NaCl-20 mM Tris-HCl (pH 7.2)-1% Tween-20 containing 150 ng of WGA was added. After incubation for 2 h at room temperature, the wells were again washed three times with the washing buffer. Color development was performed with Immunostain HRP-1000 Konica (Konica, Tokyo) as a substrate of peroxidase. For deglycosylation, the fractions containing glycopeptides were lyophilized, dissolved in 50 μl of 50 mM Tris-HCl (pH 8.0), and treated with 0.5 mU *N*-deglycanase (Recombinant *N*-Glycanase™, Takara Shuzo, Otsu). Deglycosylation was performed by incubation for 18 h at 37°C and terminated by the addition of 1 ml of 0.15% TFA. The deglycosylated peptide was subjected to RP-HPLC and eluted as described above.

RESULTS

The egg envelope subunits, ZI-1,2 and ZI-3, of medaka were purified after *S*-carboxymethylation as described in a previous paper (*18*). The purified Chg H and Chg L were also reduced and *S*-carboxymethylated.

The N-termini of these proteins were probably blocked, because none of their N-terminal amino acids were determined by the sequencing.

Comparative Analysis of Peptides from ZI-1,2 Subunit and Its Precursor, Chg H—The rcmZI-1,2 and rcmChg H were digested with V8-protease and then subjected to RP-HPLC. The resultant peptide maps are shown in Fig. 1. The maps showed a high degree of similarity with respect to major peptide peaks, with the exception of Z12-3 in the ZI-1,2 digest, and amino acid sequence analysis revealed that many corresponding peaks (H2 and Z12-1, H3 and Z12-2, H5 and Z12-4, and H6 and Z12-5) were identical peptides (Fig. 1, also see later). The peptides H5 and Z12-4 were identified as the N-terminal peptides of Chg H and ZI-1,2, respectively, since their N-terminal amino acids were blocked and they were found, on amino acid sequenc-

ing after partial proteolysis, to contain a Pro-X-Y repeat region of Chg H and ZI-1,2, as described in the previous paper (*18*). A notable difference in the minor peaks between the two peptide maps was that peaks H1 and H4 on the map of rcmChg H had no corresponding peaks on the map of rcmZI-1,2 (Fig. 1). Examination of the amino acid sequences of the peaks H1 and H4 revealed that they were derived from the C-terminal region of rcmChg H: the sequence of H4 was coincident with the C-terminal amino acid sequence, Val^{579}-Gln^{591}, deduced from the cDNA for Chg H (*11*); and that of H1 was identical to the preceding sequence, Arg^{570}-Glu^{578} (Fig. 2).

The amino acid sequence of the peptide Z12-2 in the digests of ZI-1,2 was identical to that of H3 in the Chg H digest. This peptide corresponded to Lys^{530}-Gly^{557} deduced from the cDNA for Chg H, and was the closest peptide to the C-terminus of Chg H among the peptides obtained from the rcmZI-1,2 digest (Fig. 2). Therefore, it was strongly suggested that the peptide Z12-2 was the C-terminal peptide of the ZI-1,2 subunit, and the C-terminal amino acid of the ZI-1,2 subunit was Gly^{557}. The peptide of peak Z12-3 in the digest of the ZI-1,2 subunit was derived from the third egg envelope subunit protein as reported in the previous paper (*18*). Peaks Z12-1, Z12-4, and Z12-5 on the peptide map of rcmZI-1,2 correspond to the peaks H2, H5, and H6, respectively, on the peptide map of rcmChg H as described above. The locations of these peptides in the amino acid sequence deduced from cDNA for Chg H were identified (Fig. 2).

Comparative Analysis of Peptides from ZI-3 Subunit and Its Precursor, Chg L—The rcmZI-3 and rcmChg L were digested with Lys-C, and peptide maps of these proteins were produced by RP-HPLC of the digests. The peptide map obtained from rcmZI-3 was almost the same as that from rcmChg L (Fig. 3). Amino acid sequences of all the corresponding peptides except L6 in the Chg L digest and Z3-6 in the ZI-3 digest were identical, and the locations of these peptides in the deduced amino acid sequence of the Chg L cDNA (*9*) were identified (Fig. 4). The peak Z3-6 from rcmZI-3 was found to contain a single peptide that corresponded to His^{183}-Lys^{201} of the deduced amino acid sequence of Chg L cDNA, as shown in Fig. 4. Although the second amino acid of Z3-6 was Asn^{184} in the deduced amino acid sequence of the cDNA, the second cycle of amino acid sequencing of this peptide did not give any signal for this position. This is probably because the second amino acid was *N*-glycosylated, since Asn^{184} was in a consensus sequence of an *N*-linked glycosylation site, as shown in Fig. 4. However, the peak L6 in the Chg L digest was found to be composed of two peptides, and they were separable as described below.

Separation of Two Peptides in the Peak L6 in the Chg L Digest—When the isolated peak L6 was subjected to amino acid sequencing, two different signals were obtained in every cycle of the sequencing, suggesting that L6 comprised two different peptides. To identify the peptides, their separation was required. Since the retention time of

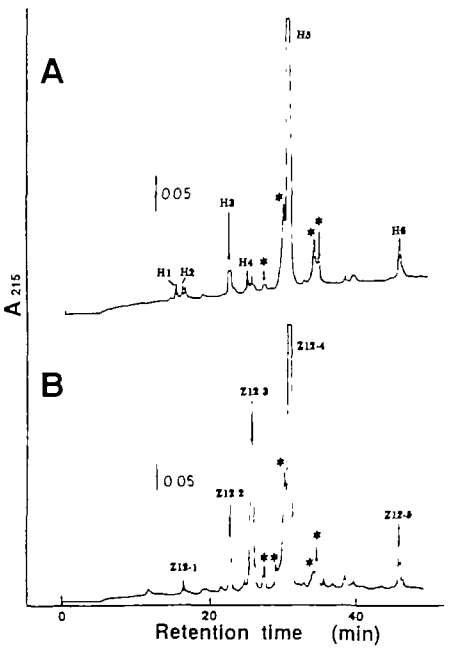

Fig. 1. **Comparison of peptide maps of the ZI-1,2 subunit and its precursor, Chg H.** The egg envelope subunit ZI-1,2 and its precursor, Chg H, were *S*-carboxymethylated and digested with V8-protease. The peptide maps for Chg H (A) and ZI-1,2 (B) were obtained by RP-HPLC as described in "MATERIALS AND METHODS." Peptides H5 and Z12-4 were N-terminal polypeptide of Chg H and ZI-1,2, respectively. The peptide Z12-3 was derived from a novel protein of the egg envelope as described in a previous paper (*18*). Amino acid sequences of the peptides in the peaks marked with asterisks could not be determined since the peaks probably composed multiple peptides.

```
MARHWSITVFSALALLCSFLGTEVDAQKGNPQDPKVPYPPYYPQPKPQDP      50

QHVSPPYYPGKPQNPPQKPSNPQYPSYPQTPQNPQVPQNPQVPQNPQYPS     100

YPQNPSYPQNPSYPQYPSNPPTSQNPSYPQNPKLFQDGKPSNPQQPQVPQ     150

YPSKPQPPQNPQVPQYPSKPQPPQNPQVPQYPSKPQNPQYPSKPQYPQAP     200

QQPQNPQYPSKPQDPGKNPNTPPIGPPPPKSCEVPRDVRVPCGVPDISPS     250

ACDAIDCCHDGQSCYFGTGATVQCTKDGHFIVVVAKDVTLPHIDLETISL     300

LGQGQDCGPADSNSAFAIYYFPVTYCGTVVMEEPGVIVYENRMTSSYEVG     350
                                          H2
                                          Z12-1

VGPLGAITRDSSFELLFQCRYRATSVETLVVEVQPPDSPLSIAELGPLNV     400

YLQIANGQCQTKGCDEAAAAYTSFYTDADYPVTKVLRDPVYVDVQILGRT     450
               H6
               Z12-5

DPNLVLTLGRCVATTSPNAFSLPQVDILIDGCPYADDRYLSALVPIDHSS     500

GLPFPTHHSRFLFKMFTFVDPHSMEPLREKVYIHCSTAACVPGQGVSCEP     550
                                  H3
                                  Z12-2

SCSRRKGRDTEAVAIRTDERRVVVSSGEVLNLAAADEPSEQ              591
                    H1        H4
```

Fig. 2. **Amino acid sequences of the peptides in the V8-protease digests of the egg envelope subunit ZI-1,2 and its precursor protein, Chg H.** The whole amino acid sequence is predicted from the cDNA for Chg H (*11*). The amino acid sequences of the peptides in Fig. 1, A and B, were determined, and their locations in the whole sequence are shown by underlines. The peptides H1 and H4, found only in the Chg H digest, were derived from the C-terminal region of Chg H, H4 being C-terminal peptide of Chg H.

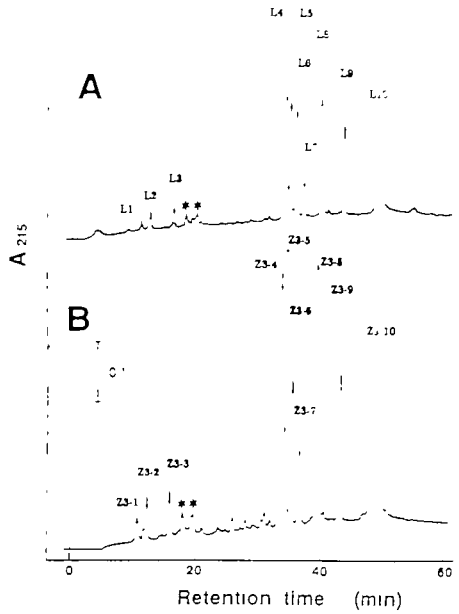

Fig. 3. **Comparison of peptide maps of the ZI-3 subunit and its precursor, Chg L.** The egg envelope subunit ZI-3 and its precursor, Chg L, were *S*-carboxymethylated and digested with Lys-C. The peptide maps for Chg L (A) and ZI-3 (B) were obtained through RP-HPLC as described in "MATERIALS AND METHODS." A peak Z3-6 included only a single peptide, while a peak L6 was composed of two peptides (see "RESULTS"). Amino acid sequences of the peptides in the peaks marked with asterisks could not be determined since any PTH-amino acids were not detected as the results of amino acid sequencing.

```
MNKFTAVCLVVLALLDGFCDAQHNYGKPSYPPTGSKTPQDPTQQKQLHEK      50
                                          L2    L1
                                          Z3-2  Z3-1

ELTWKYPADPQPEAKPVVPFEQRYPVPAATVAVECREDLAHVEAKKDLFG     100
L5                                                L10
Z3-5                                              Z3-10

IGQFIDPADLTLGTCPPSAEDPAAQVLIFESPLQNCGSVLTNTEDSLVYT     150

FTLNYNPKPLGSAPVVRTSQAVVIVECHYPRKHNVSSLALDPLWVPFSAA     200
                                      Ldg
                                  Z3-6, Zdg

KMAEEFLYFTLKLTTDDFQFERPSYQYFLGDLIHIEATVKQYFHVPLRVY     250
L7        L9                      L8
Z3-7      Z3-9                    Z3-8

VDRCVATLSPDANSSPSYAFIDNYGCLLDGRITGSDSKFVSRPAENKLDF     300
                                      L3
                                      Z3-3

QLEAFRFQGADSGNIYITCHLKATSAAYPLDAEHRACSYIQGWKEVSGAD     350
                    L4
                    Z3-4

PICASCESGGFEVHANAVVSHGTSTLSGGGHGTGKPSDPSRKTREAAKTE     400
                                                  Lct

VLEWEGDVTLGPIPIEERRV                                    420
```

Fig. 4. **Amino acid sequences of the peptides in the Lys-C digests of the egg envelope subunit, ZI-3, and its precursor protein, Chg L.** The whole amino acid sequence is predicted from the cDNA for Chg L (*9*). The amino acid sequences of the peptides in Fig. 3, A and B, were determined, and their locations in the whole sequence are shown by underlines. An *N*-linked glycosylation site, Asn^{184}, is shown by an asterisk. For the peptides Ldg, Zdg, and Lct, see Fig. 5 and "RESULTS."

L6 was almost the same as that of Z3-6 in HPLC fractionation, and one of the two amino acids obtained in every cycle of sequencing was identical to that of Z3-6, it was conjectured that one of the two peptides in peak L6 was identical to the peptide in Z3-6. Sugar chains of rcmChg L, rcmZI-3, and peptides in their digests were detected using various peroxidase-conjugated lectins, and both rcmChg L and rcmZI-3 were reactive with Con A, PHA-E4, and WGA, but not with DBA, LCA, PNA, nor UAE-1. Among the peptides derived from the Lys-C digests of rcmZI-3 and rcmChg L, only Z3-6 and L6 were reactive with Con A, PHA-E4, and WGA, showing that they were glycopeptides. Moreover,

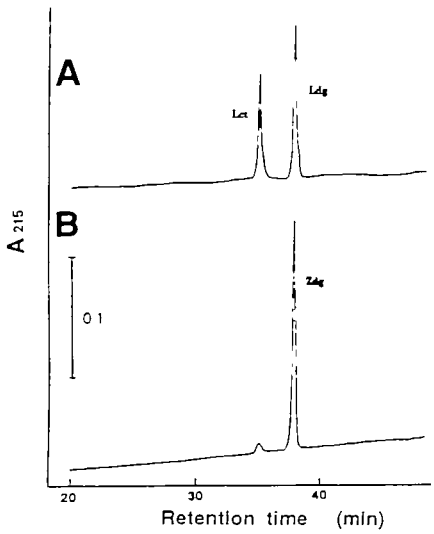

Fig. 5. **Separation of two peptides included in peak L6.** Peptides included in peaks L6 and Z3-6 were subjected to RP-HPLC after treatment with *N*-glycanase. A: HPLC pattern of deglycosylated L6. B: HPLC pattern of deglycosylated Z3-6. Amino acid sequences of the peptides Ldg, Lct, and Zdg were determined, and their locations in the whole amino acid sequence of Chg L deduced from the cDNA are shown by underlines in Fig. 4.

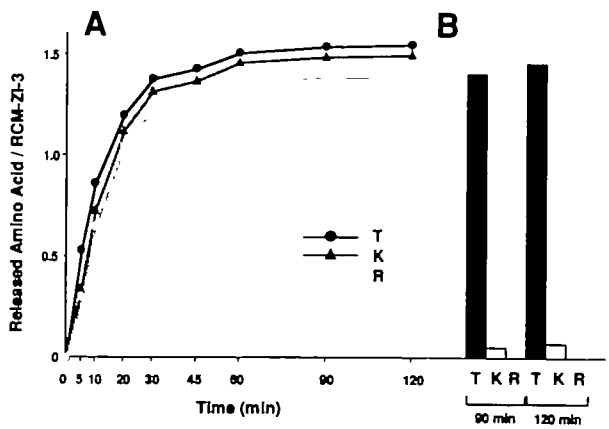

Fig. 6. **Determination of the C-terminus of ZI-3.** The C-terminus of ZI-3 was determined by use of CPases. Amino acids released from the C-terminus of rcmZI-3 by incubation with CPases A and B together for 0, 5, 10, 20, 30, 45, 60, 90, and 120 min were measured (A). Amino acids released from the C-terminus of rcmZI-3 by incubation with CPase A alone for 90 and 120 min were measured (B). The amino acids are expressed by one-letter symbols.

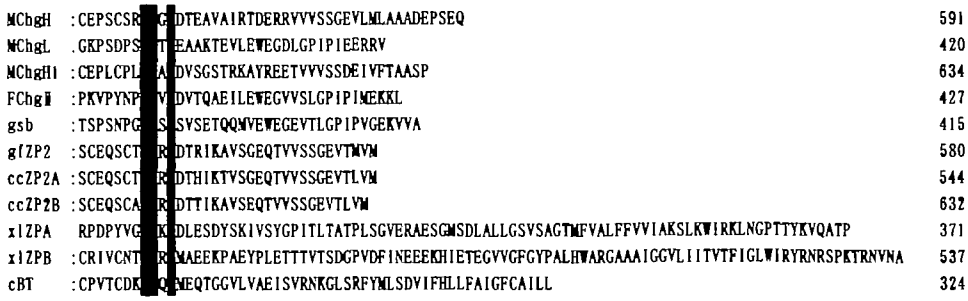

Fig. 7. **Comparison of C-terminal regions between medaka choriogenins and some proteins homologous to zona pellucida.** The putative consensus sequence for the modification, Arg-Lys-X-Arg, is indicated by shading. The proteins: are Chg H (MChgH), Chg L (MChgL), and the novel protein precursor [MChgH1 (*18*)] of medaka; choriogenin II of *Fundulus heteroclitus* (FChg II); chorion protein of gilthead sea bream (gsb); ZP-2 of goldfish (gfZP-2); two ZP-2s of common carp (ccZP-2A and ccZP-2B); egg envelope subunits of *Xenopus laevis* (xlZPA and xlZPB); and chick beta-tectorin (cBT).

peptide Z3-6 had a consensus sequence for an *N*-linked glycosylation site. It was expected, therefore, that the two peptides present in L6 might be separable by RP-HPLC after deglycosylation with *N*-glycanase. The RP-HPLC pattern of the deglycosylated L6 exhibited two peaks: one (Lct) was eluted at the same retention time as the non-deglycosylated L6 and Z3-6, and the other (Ldg) was eluted at a longer retention time than the non-deglycosylated L6 and Z3-6 (Fig. 5A), in comparison, the deglycosylated Z3-6 (Zdg) was eluted as an almost single peak by RP-HPLC at a longer retention time than non-deglycosylated Z3-6, and at the same retention time as that of Ldg (Fig. 5B). The amino acid sequence of the peptide Ldg was identical to that of Zdg and to the sequence His^{183}-Lys^{201} deduced from the nucleotide sequence of cDNA for Chg L, except for the second amino acid residue (Fig. 4). The second amino acid of these peptides was aspartic acid, while the second amino acid deduced from cDNA for Chg L was asparagine in the consensus sequence of *N*-linked glycosylation site as described above (Fig. 4). The difference was probably due to the enzymatic deglycosylation by *N*-glycanase. The amino acid sequence of the peptide Lct was identical to the sequence Thr^{399}-Val^{420} deduced from cDNA for Chg L. It is known from the cDNA for Chg L that Val^{420} is the C-terminus of Chg L, and therefore the peptide Lct is identified as the C-terminal peptide of Chg-L. No shorter or longer peptide corresponding to the C-terminal peptide, Lct, was found in the ZI-3 digest.

Determination of the C-Terminus of ZI-3—HPLC fractionation of the digest of rcmZI-3 gave peak Z3-4 as the most C-terminal of the peptides obtained, and part of the amino acid sequence of the Z3-4 peptide was identical to Ala^{323}-Val^{363} deduced from cDNA for Chg L (Fig. 3 and Fig. 4). Although peptide Z3-4 was suspected to be the C-terminal peptide of ZI-3, it was too long for its full amino acid sequence to be determined. To determine the C-terminus of ZI-3, the C-terminal region of rcmZI-3 was analyzed by use of carboxypeptidases (CPases). CPases A and B together released Thr, Lys, and Arg (1:1:1) from rcmZI-3 (Fig. 6A), while CPase A alone release a large amount of Thr and a small amount of Lys, but no Arg (Fig. 6B). These results indicated that the C-terminal amino acid residue of ZI-3 was Thr, and that the amino acid sequence of the C-terminal region was Arg-Lys-Thr. This was identical to the sequence Arg^{391}-Lys^{392}-Thr^{393} deduced from the cDNA for Chg L (Fig. 4).

In preliminary examinations, Thr and Val (1:1) were released from Chg L by CPase A alone, and Thr, Val, Lys, and Arg (1:1:1:3) were released from Chg L by CPase A and B together (data not shown). These results and the presence of Lct suggested that the purified Chg L contained both a full-length protein and a protein lacking its C-terminal region. However, it was found from the results of C-terminal analyses of ZI-3 that most of the assembled subunit ZI-3 lacked the C-terminal region.

DISCUSSION

To clarify the mechanism of molecular assembly of the soluble Chgs into the insoluble inner layer of egg envelope, it is necessary to compare the molecular structures of Chgs with those of the respective subunits after assembly. Although the primary structures of Chg H and Chg L have been clarified by cDNA analyses (*9, 11*), the Chg proteins isolated from ascites fluid were treated in the same manner as the purified subunits, and their peptide mapping and amino acid sequencing were performed for comparison with the subunits ZI-1,2 and ZI-3.

The primary structures of the egg envelope subunit proteins were very similar to those of the corresponding Chgs, because the patterns of the peptide maps and the amino acid sequences of the peptides of the egg envelope subunits, rcmZI-1,2 and rcmZI-3, were almost the same as those of rcmChg H and rcmChg L, respectively. We also found that of the two potential *N*-linked glycosylation sites in Chg L, Asn^{184}, and Asn^{263} (*9*), only Asn^{184} was glycosylated, and that the same was also the case in ZI-3. However, we suspected that the structure of the C-terminal regions of the egg envelope subunits was different from that of Chgs. The C-terminal residues of the presumed C-terminal peptides obtained from rcmChg H and rcmChg L were Gln and Val, identical with the C-termini deduced from cDNA for Chg H and Chg L, respectively. However, it was determined that the C-terminus of the egg envelope subunit ZI-1,2 was Gly^{557} and that of ZI-3 was Thr^{393} by analysis using CPases. Thus, ZI-1,2 subunit was shorter than the precursor, Chg H, by 34 amino acids and, ZI-3 subunit was shorter than Chg L by 27 amino acids.

Most interestingly, it appears that before they are incorporated into the inner layer of the egg envelope as its subunits Chg L undergoes proteolytic cleavage at the C-terminal side of Thr^{393}, which is situated in the sequence

Arg^{391}-Lys^{392}-Thr^{393}-Arg^{394}, and that Chg H undergoes similar cleavage at the C-terminal side of Gly^{557}, which is situated in the sequence Arg^{555}-Lys^{556}-Gly^{557}-Arg^{558}. Therefore, it is highly probable that the sequence Arg-Lys-X-Arg is a consensus sequence for the position of processing cleavage. The sequence Arg-Lys-Ala-Arg was also found in a portion of approximately 30 amino acids from the C-terminus of the newly found precursor of the third egg envelope subunit of medaka (*18*). Recently, amino acid sequences of egg envelope (or chorion) proteins were deduced from the cDNAs in other fish species, such as winter flounder [wf ♀ (*20*)], *Fundulus heteroclitus* [choriogenin I and II (*21*)], *Sparus aurata* (gilthead sea bream; accession, X93306), *Carassius auratus* (goldfish; ZP-2 and ZP-3; accession, Z72495 and L41636), and *Cyprinus carpio* (common carp; three ZP-2s and ZP-3; accession, Z72491, Z72492, Z72494, and L41638). The putative consensus sequence, Arg-Lys-X-Arg, was also found in the C-terminal regions of the deduced amino acid sequences of Chg II of *F. heteroclitus*, egg envelope protein of gilthead sea bream, ZP-2 of goldfish and two ZP-2s of common carp, but not in those of Chg I of *F. heteroclitus*, wf ♀ of winter flounder, ZP-3 of goldfish, or ZP-3 of common carp (Fig. 7). Moreover, the egg envelope subunits of *Xenopus laevis* (ZPA and ZPB; accession, U44949 and U44950) and chick beta-tectorin, which is an extracellular matrix molecule of the inner ear (*22*), also have the putative consensus sequence Arg-Lys-X-Arg in the C-terminal regions. Figure 7 shows a comparison of C-terminal regions among zona pellucida homologous proteins containing the putative consensus sequence, Arg-Lys-X-Arg, in their C-terminal regions. It is reported that predicted the amino acid sequences of all mammalian zona pellucida proteins produced in oocyte or follicle cells contain a sequence called a "putative furin cleavage site" in the C-terminal regions, which is rich in basic amino acids (*23*). The putative consensus sequence, Arg-Lys-X-Arg, seems to be comparable to the furin cleavage site, although we have no information about the nature of a protease responsible for the cleavage in fish choriogenins.

From preliminary analyses of the C-terminal portion of the purified Chg L by use of CPases, and from peptide mapping and amino acid sequencing, two different C-termini were found: one was that of the full-length protein, and the other was probably that of a protein lacking the C-terminal region. It is considered that on prolonged administration of estrogen to male medaka fish, Chg L [formerly called SF-substance (*14*)] is synthesized in liver cells, secreted into blood, then accumulated in the ascites fluid. During accumulation in the ascites fluid under non-physiological conditions, some molecules of Chg L might be truncated by the protease(s). This was also the case with Chg H (data not shown).

We have not yet identified of the site in the spawning female fish where Chg H and Chg L are truncated prior to their assemblage into the inner layer of egg envelope under physiological conditions. Clarification of the site of the molecular modification seems to be crucial for elucidation of the mechanism of the subunit assembly.

REFERENCES

1. Yamagami, K., Hamazaki, T.S., Yasumasu, S., Masuda, K., and Iuchi, I. (1992) Molecular and cellular basis of formation, hardening, and breakdown of the egg envelope in fish. *Int. Rev. Cytol.* **136**, 51-92
2. Begovac, P.C. and Wallace, R.A. (1989) Major vitelline envelope proteins in pipefish oocytes originate within the follicle and are associated with the Z3 layer. *J. Exp. Zool.* **251**, 56-73
3. Oppen-Berntsen, D.O., Helvik, J.V., and Walther, B.T. (1990) The major structural proteins of Cod (*Gadus morhua*) egg shells and protein crosslinking during teleost egg hardening. *Dev. Biol.* **137**, 258-265
4. Brivio, M.F., Bassi, R., and Cotelli, F. (1991) Identification and characterization of the major components of the *Oncorhynchus mykiss* egg chorion. *Mol. Rep. Dev.* **28**, 85-93
5. Hyllner, S.J., Oppen-Berntsen, D.O., Helvic, J.V., Walther, B.T., and Haux, C. (1991) Oestradiol-17β induces the major vitelline envelope proteins in both sexes in teleosts. *J. Endocrinol.* **131**, 229-236
6. Scapigliati, G., Carcupino, M., Taddei, A.R., and Mazzini, M. (1994) Characterization of main egg envelope proteins of the sea bass, *Dicentrarchus Labrax* L. (Teleostea, Serranidae). *Mor. Rep. Dev.* **38**, 48-53
7. Hamazaki, T.S., Iuchi, I., and Yamagami, K. (1987) Isolation and partial characterization of a "spawning female-specific substance" in the teleost, *Oryzias latipes*. *J. Exp. Zool.* **242**, 343-349
8. Murata, K., Iuchi, I., and Yamagami, K. (1993) Isolation of H-SF substances, the high-molecular-weight precursors of egg envelope proteins, from the ascites accumulated in the oestrogen-treated fish, *Oryzias latipes*. *Zygote* **1**, 315-324
9. Murata, K., Sasaki, T., Yasumasu, S., Iuchi, I., Enami, J., Yasumasu, I., and Yamagami, K. (1995) Cloning of cDNAs for the precursor protein of a low-molecular-weight subunit of the inner layer of the egg envelope (chorion) of the fish, *Oryzias latipes*. *Dev. Biol.* **167**, 9-17
10. Murata, K., Hamazaki, T.S., Iuchi, I., and Yamagami, K. (1991) Spawning female-specific egg envelope glycoprotein-like substances in *Oryzias latipes*. *Dev. Growth Differ.* **33**, 553-562
11. Murata, K., Sugiyama, H., Yasumasu, S., Iuchi, I., Yasumasu, I., and 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
12. Hamazaki, T., Iuchi, I., and Yamagami, K. (1984) Chorion glycoprotein-like immunoreactivity in some tissues of adult female medaka. *Zool. Sci.* **1**, 148-150
13. Hamazaki, T.S., Iuchi, I., and Yamagami, K. (1985) A spawning female-specific substance reactive to anti-chorion (egg envelope) glycoprotein antibody in the teleost, *Oryzias latipes*. *J. Exp. Zool.* **235**, 269-279
14. Hamazaki, T.S., Iuchi, I., and Yamagami, K. (1987) Production of a "spawning female-specific substance" in hepatic cells and its accumulation in the ascites of the estrogen-treated adult fish, *Oryzias latipes*. *J. Exp. Zool.* **242**, 325-332
15. Hamazaki, T.S., Nagahama, Y., Iuchi, I., and Yamagami, K. (1989) A glycoprotein from the liver constitutes the inner layer of the egg envelope (zona pellucida interna) of the fish, *Oryzias latipes*. *Dev. Biol.* **133**, 101-110
16. Yamagami, K., Murata, K., Hamazaki, T.S., and Iuchi, I. (1994) Intrahepatic formation of SF substances, the precursors of egg envelope proteins, in response to estrogen administration in fish in *Perspectives in Comparative Endocrinology* (Davey, K.G., Peter, R.E., and Tobe, S.S., eds.) pp. 316-324, Natl. Res. Council of Canada, Ottawa
17. Sugiyama, H., Murata, K., Iuchi, I., and Yamagami, K. (1996) Evaluation of solubilizing methods of the egg envelope of the fish, *Oryzias latipes*, and partial determination of amino acid sequence of its subunit protein, ZI-3. *Comp. Biochem. Physiol.* **114B**, 27-33
18. Sugiyama, H., Yasumasu, S., Murata, K., Iuchi, I., and Yama-

gami, K. (1998) The third egg envelope subunit in fish; cDNA cloning and analysis, and the gene expression. *Dev. Growth Differ.* **40**, 35-45
19. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685
20. Lyons, C.E., Payette, K.L., Price, J.L., and Huang, R.C.C. (1993) Expression and structural analysis of a teleost homologue of a mammalian zona pellucida gene. *J. Biol. Chem.* **268**, 21351-21358
21. LaFleur, G.J. Jr., Byrne, B.M., Haux, C., Greenberg, R.M., and Wallace, R.A. (1995) Liver-derived cDNAs: vitellogenins and vitelline envelope protein precursors (choriogenins) in *Reproductive Physiology of Fish* (Goetz, F.W. and Thomas, P., eds.) pp. 336-338, Austin, Texas
22. Killick, R., Legan, P.K., Malenczak, C., and Richardson, G.P. (1995) Molecular cloning of chick beta-tectorin, an extracellular matrix molecule of the inner ear. *J. Cell Biol.* **129**, 535-547
23. Harris, J.D., Hibler, D.W., Fontenot, G.K., Hsu, K.T., Yurewicz, E.C., and Sacco, A.G. (1994) Cloning and characterization of zona pellucida genes and cDNAs from a variety of mammalian species: the ZPA, ZPB and ZPC gene families. *DNA Sequence-J. Sequencing Mapping* **4**, 361-393