Mass Spectrometric Evidence That Proteolytic Processing of Rainbow Trout Egg Vitelline Envelope Proteins Takes Place on the Egg*

Received for publication, June 20, 2005 and in revised form, August 24, 2005 Published, JBC Papers in Press, September 12, 2005, DOI 10.1074/jbc.M507670200

Costel C. Darie†, Martin L. Biniossek‡, Mary A. Gawinowicz§, Yelena Milgrom¶, Joerg O. Thumfart‡, Luca Jovine¶, Eveline S. Litscher‡, and Paul M. Wasserman§†

From the †Brookdale Department of Molecular, Cell and Developmental Biology, Mount Sinai School of Medicine, New York, New York 10029-6574, ‡Institut fuer Molekulare Medizin und Zellforschung/Physiologie II, Albert-Ludwigs Universitaet, Stefan-Meier-Strasse 19, 79104 Freiburg, Germany, §Protein Core Facility, Columbia University College of Physicians and Surgeons, New York, New York 10032-3702

The rainbow trout egg vitelline envelope (VE) is constructed of three proteins, called VEα, VEβ, and VEγ, that are synthesized and secreted by the liver and transported in the bloodstream to the ovary, the site of VE assembly around eggs. All three proteins possess an N-terminal signal peptide, a zona pellucida domain, a consensus furin-like cleavage site (CFLCS) close to the C terminus, and a short propeptide downstream of the CFLCS. Proteolytic processing at the CFLCS results in loss of the short C-terminal propeptide from precursor proteins and enables incorporation of mature proteins into the VE. Here mass spectrometry (matrix-assisted laser desorption ionization-time-of-flight-mass spectrometry and liquid chromatography-mass spectrometry) with a micromass-quadrupole TOF hybrid mass and a QSTAR Pulsar i mass spectrometer was employed with VE proteins isolated from rainbow trout eggs in a peptidomics-based approach to determine the following: 1) the C-terminal amino acid of mature, proteolytically processed VE proteins; 2) the cellular site of proteolytic processing at the CFLCS of VE precursor proteins; and 3) the relationship between proteolytic processing and limited covalent cross-linking of VE proteins. Peptides derived from the C-terminal region were found for all three VE proteins isolated from eggs, indicating that processing at the CFLCS occurs after the arrival of VE precursor proteins at the egg. Consistent with this conclusion, peptides containing an intact CFLCS were also found for all three VE proteins isolated from eggs. Furthermore, peptides derived from the C-terminal propeptides of VE protein heterodimers VEα-VEγ and VEβ-VEγ were found, suggesting that a small amount of VE protein can be covalently cross-linked on eggs prior to proteolytic processing at the CFLCS. Collectively, these results provide important evidence about the process of VE formation in rainbow trout and other non-cyprinoid fish and allow comparisons to be made with the process of zona pellucida formation in mammals.

All vertebrate eggs are surrounded by an extracellular coat, referred to as a chorion, vitelline envelope (VE), or zona pellucida (ZP) (1). The coatings surrounding fish and mammalian eggs, often designated as VE and ZP, respectively, are located between the oolemma and surrounding follicle cells and frequently exhibit distinguishable layers. These egg coats have been shown to play important roles during oogenesis, fertilization, and early embryogenesis.

Proteins that constitute the VE and ZP are highly conserved and related to each other (2). For example, a feature common to VE and ZP proteins is the presence of a so-called ZP domain (3–5), a sequence of ~260 amino acids. Also, proteins into long fibrils or filaments (6). Rainbow trout (Oncorhyncus mykiss) VEs consist of at least three proteins, called VEα (58 kDa), VEβ (~52 kDa), and VEγ (~47 kDa), that possess an N-terminal signal peptide (SP), a proline-glutamine- (PQ-) rich (PQ) region, a ZP domain, a consensus furin-like cleavage site (CFLCS) close to the C terminus of the polypeptides (7, 8), and a short propeptide downstream of the CFLCS. VEα and VEβ also have a trefoil (P) domain, just upstream of the ZP domain, that has six cysteine residues present as three intramolecular disulfides (Cys1–Cys1, Cys1–Cys2, and Cys1–Cys2) (9). The Zip domain of each VE protein has eight conserved Cys residues present as four intramolecular disulfides (7). VEα and VEβ exhibit high homology with each other and are related to mammalian egg coat proteins ZP1 and ZP2, whereas VEγ is homologous with ZP3 (8).

Mouse ZP proteins (ZP1–3) are synthesized in the ovary by growing oocytes (5, 10). ZP proteins from other mammals also are synthesized exclusively in the ovary by oocytes and/or follicle cells (2). Similarly, VE proteins in amphibian (11, 12) and cyprinoid fish (e.g., Zebrasfish, goldfish, and carp) (13–16) are synthesized in the ovary. On the other hand, VE proteins in rainbow trout (17) and a large number of other non-cyprinoid fish (e.g. winter flounder, sea bream, cod, and medaka) (18–22) are synthesized in the liver under hormonal (estradiol-17β) control and then transported in the bloodstream as proproteins to the ovary where they form the inner layer of the egg VE. In most instances it would appear that polymerization of fish VE proteins takes place in the ovary (8, 17, 21, 23–27).

The site of proteolytic processing of trout VE precursor proteins that occurs during VE formation around eggs is of considerable interest. Here we used mass spectrometry (MS) measurements of individual VE membrane; ER, endoplasmic reticulum; HMWP, high molecular weight protein; FA, formic acid; ACN, acetonitrile; DTT, dithiothreitol; Met, oxidized methionine; CM, carbamidomethyl; MS, mass spectrometry; CID, collision-induced dissociation; MALDI-TOF-MS, matrix-assisted laser desorption/ionization-time of flight-mass spectrometry; LC-MS, liquid chromatography-mass spectrometry; LC-MS(QTOF), LC-MS with a Micromass-QTOF hybrid mass spectrometer; LC-MS(QSTAR), LC-MS with a QSTAR Pulsar i mass spectrometer; HPLC, high pressure liquid chromatography.

All vertebrate eggs are surrounded by an extracellular coat, referred to as a chorion, vitelline envelope (VE), or zona pellucida (ZP) (1). The
proteins isolated from rainbow trout eggs to assess whether or not the CFLCS and C-terminal propeptide were retained following synthesis and secretion of the proteins by the liver and transport to the ovary and eggs. Because only VE proteins that had recently arrived in the vicinity of eggs would be expected to retain a CFLCS and C-terminal propeptide, i.e. a small fraction of total VE protein associated with eggs. MS provides the level of sensitivity necessary to make such an assessment. The evidence obtained strongly suggests that proteolytic cleavage at the CFLCS of VE precursor proteins, which results in removal of the short C-terminal propeptide, takes place on the egg. Occasionally this occurs after formation of covalently linked heterodimers (VEα-VEγ and VEβ-VEγ) but before incorporation of VE proteins into the VE. This would be consistent with the suggestion that soluble VE precursor proteins are converted into insoluble VE proteins at the site of envelope assembly in the ovary, not in the liver or bloodstream (8, 17, 21, 23–27). Furthermore, the evidence is consistent with a recently proposed mechanism for ZP assembly in mammals that involves short hydrophobic patches near the C terminus (external hydrophobic patch (EHP)) and within the ZP domain (internal hydrophobic patch (IHP)) of ZP precursor proteins (28).

MATERIALS AND METHODS

Reagents

Reagents were obtained from the following commercial sources: MALDI-TOF-MS standards, chymotrypsin, AspN, Glu(V8), trifluoroacetic acid, formic acid (FA), and acetonitrile (ACN) from Sigma; trypsin from Roche Applied Science; SDS-polyacrylamide gels from Bio-Rad; and molecular weight markers from Bio-Rad, In vitrogen, and New England Biolabs.

Isolation of VE Proteins

The procedure described by Brivio et al. (29) was used to isolate VE proteins from rainbow trout (O. mykiss) eggs. Briefly, eggs were obtained from females by abdominal squeezing in a bloodless procedure that avoided contamination of eggs with proteins from blood. Eggs were frozen immediately by immersion in an ethanol and dry ice mixture. To prepare proteolytic digests of VE proteins, VE proteins were subjected to SDS-PAGE, and gel pieces containing individual VE proteins were excised and then digested with various proteases. Analyses of VE protein digests by MS were performed using several VE preparations.

Gel Electrophoresis

VE proteins were separated on 10% gels by SDS-PAGE (30) under reducing and nonreducing conditions (Fig. 1A). For nonreducing conditions, samples were dissolved in the absence of DTT. Gels were then stained with Coomassie Brilliant Blue. In most cases, a small portion of unstained gel just above the Coomassie-stained band (i.e. higher molecular weight) was used for MS experiments. As a result, under reducing conditions, there was a small amount of cross-contamination of VE proteins (e.g. VEβ peptides in VEγ samples and VEα peptides in VEβ samples), but this did not interfere with interpretation of the MS analyses. For molecular weight determination by MS, individual VE proteins were obtained by preparative SDS-PAGE on 12% gels under nonreducing conditions, and the position of each protein was determined by transient zinc staining (see below). Gel bands containing individual VE proteins were excised, electroeluted (Biotrap, Schleicher & Schuell) in 25 mM Tris, 192 mM glycine, and 0.025% (w/v) SDS, pH 8.8, for 12–16 h at room temperature, and concentrated to 0.5–1 mg/ml. To remove SDS, individual proteins were dialyzed twice against 6 l 1,6 urea (16 h each dialysis), followed by dialysis against distilled water six times (12–16 h each dialysis).

Enzymatic Digestion of VE Proteins

Digestion of gel pieces containing individual VE proteins with proteases was carried out in the manner of Hellman et al. (31) with certain modifications (32). Gel pieces containing 0.5–2 μg of protein were incubated with 60% (v/v) ACN for 20 min, dried completely in a SpeedVac evaporator, and rehydrated for 10 min with digestion buffer (25 mM ammonium bicarbonate, pH 8.0). This procedure was repeated three times. After drying, gel pieces were again rehydrated in digestion buffer containing 10 mM DTT and incubated for 1 h at 56 °C. Reduced Cys residues in VE proteins were blocked by replacing the DTT solution with 100 mM iodoacetamide in 25 mM ammonium bicarbonate, pH 8.0, for 45 min at room temperature with occasional vortexing. In some experiments, the reduced Cys residues were not blocked by iodoacetamide. Gel pieces were dehydrated, dried, and rehydrated twice. Dried gel pieces were then digested for 16–18 h at 37 °C in digestion buffer containing 15 ng/μl trypsin (cleavage at the C terminus of Arg and Lys residues), chymotrypsin (cleavage at the C terminus of Phe, Tyr, Trp, Met, and Leu residues), endoprotease AspN (cleavage at N terminus of Asp and Glu residues), or endopeptidase Glu(V8) (cleavage at C terminus of Glu and Asp residues). For tryptic and chymotryptic digests, 5 mM calcium chloride was added to the digestion buffer. Following digestion, peptides were extracted twice from gel pieces by addition of 300 μl of 60% ACN, 5% FA in 25 mM ammonium bicarbonate, pH 8.0, and shaking for 60–90 min at room temperature. Solutions containing peptides of VE proteins were pooled, dried, and used for MALDI-TOF-MS in either linear or reflective modes and MS or MS/MS by LC-MS(QTOF and QSTAR).

Matrix-assisted Laser Desorption Ionization-Time-of-Flight-Mass Spectrometry (MALDI-TOF-MS)

MALDI-TOF-MS analyses of VE peptides were carried out on a TOF Voyager-DERP mass spectrometer (PerSeptive Biosystems, Foster City, CA) in a linear mode and on a Reflex III TOF system (Bruker Daltonics, Leipzig, Germany) in a reflective mode, both equipped with a nitrogen laser (337 nm).

Samples were dissolved in 50% ACN (linear mode) and co-crystallized with α-cyano-4-hydroxycinnamic acid matrix on a stainless steel target. In a reflective mode, the samples were prepared by a modified thin layer method (33). The dry samples were dissolved with 2% trifluoroacetic acid, and a mixture of α-cyano-4-hydroxycinnamic acid and nitrocellulose was used as matrix. After drying, samples were analyzed either undiluted or at suitable dilutions. After ionization, peptide data were collected in the mass/charge (m/z) 800–8000 (linear mode) and 700–3200 (reflective mode) range, and time-to-mass conversion was achieved by using external or internal calibration with bradykinin (m/z 1061.2), angiotensin (m/z 1297.5), insulin (m/z 5734.5), and tryptic fragments. Peaks detected by MALDI-TOF-MS in a linear mode corresponded to average m/z peptides. The peaks detected by MALDI-TOF-MS in a reflective mode corresponded to monoisotopic m/z peptides. Here we refer to the MALDI-TOF-MS peaks (m/z) as average/monoisotopic, singly charged (protonated) peptides. For MALDI-TOF-MS measurements in a linear mode, m/z differences (from the calculated value) up to 1 Da were observed, whereas calibration differences for measurements in a reflective mode were generally up to 100 ppm. Interpretation of mass spectra was performed using Mascot (www.matrixscience.com), FindMod, FindPep, PeptideMass (www.expasy.ch), as well as manually.

Electrospray Ionization (ESI) Quadrupole (Q) TOF-MS

Nanoflow HPLC-MS(LC-MS(QTOF))—Analyses were carried out with a Micromass-QTOF hybrid mass spectrometer (Waters Corp.,
Milford, MA) with a nanoelectrospray source. A fused silica tip mounting adaptor, fitted with a 75-µm (inner diameter) fused silica tip (New Objective), was connected through a 50-µm (inner diameter) fused silica tubing to the liquid chromatography (LC) detector outlet. An LC Packings system (Dionex Corp., Sunnyvale, CA), equipped with an ultimate micro pump and solvent organizer and a Switchos loading pump and Famos autosampler, was used for LC-MS. Separation was carried out on a 75-µm × 15-cm column (LC Packings C18 PepMap; 5-µl injection volume) at a flow rate of 200 nl/min, using a gradient of 2–80% ACN in 0.1% FA. The mass spectrometer was operated in the data-dependent mode and automatically switched between MS and MS/MS. The peaks detected by LC-MS(QTOF) corresponded to peptides with a mass determination of the C-terminal propeptides, which increased the probability of detecting C-terminal peptides. In MALDI-TOF-MS and tandem MS experiments, possible modifications of the peptides (e.g. cyclization of Gln or Glu residues to pyroglutamate or oxidation of Met to Met) were examined. In MS/MS experiments, the collision-induced dissociation (CID) spectra corresponding to a C-terminal peptide were analyzed using MS-Product, MS-Fit, and MS-Tag (prosector.ucsf.edu), as well as manually. Because of the presence of many Pro and Asp residues in the C-terminal propeptides, fragments that resulted in CID from the cleavage of the peptide bond of these amino acids, and the internal fragments that resulted from subsequent fragmentation of these fragments (MS/MS/MS), were taken into account. Because the C-terminal peptides contained many positively charged amino acids (e.g. Arg, His, Lys, and Asn), in MS and MS/MS experiments we searched for multiply [(2+), (3+), (4+)] charged ions, in addition to the doubly charged ions classical for tryptic digests.

RESULTS

Electrophoretic and Mass Spectrometric Analyses of Rainbow Trout VE Proteins—Under reducing conditions (+DTT) on SDS-PAGE, rainbow trout VE proteins contain several major proteins, called VEα, VEβ and VEγ, with masses of ~58, ~52, and ~47 kDa, respectively (Fig. 1A, lane a). In addition, a 4th band with a mass of ~120 kDa is present and contains VEα-VEγ and VEβ-VEγ heterodimers, as reported previously (7). No other proteins (e.g. proteins from blood) were detected either on Coomassie-stained gels or by MS experiments. Under nonreducing conditions (−DTT) on SDS-PAGE, the band containing VEα-VEγ and VEβ-VEγ heterodimers separates into two bands; the top band contains VEα-VEγ heterodimers and the bottom band contains VEβ-VEγ heterodimers (Fig. 1A, lanes b and d). In addition to the VEα, VEβ, and VEγ monomers observed under reducing (Fig. 1A, lane a) and nonreducing (Fig. 1A, lanes b and d) conditions, additional protein bands were occasionally observed with molecular weights close to that of VEα (Fig. 1A, lane b), VEγ (Fig. 1A, lane d), and VEβ (data not shown). MALDI-TOF-MS analysis of trypsin-digested gel bands marked as VEα (Fig. 1A, lane b) identified both bands as VEα (data not shown), suggesting the presence of VE isopeptides. Each VE propolypeptide contains an N-terminal SP, a PQ region, a P domain (absent in VEγ), a ZP domain, a CFLCS, and a C-terminal propeptide downstream of the CFLCS (Fig. 1B). The amino acid sequence of VE proteins containing the last Cys residue upstream of the CFLCS (VEα–Cys528, VEβ–Cys519, and VEγ–Cys519), the CFLCS, and the C-terminal propeptide are shown in Fig. 1B. Sometime after secretion by the liver, the C-terminal propeptide is removed by a furin-like propeptide processing enzyme.
FIGURE 1. SDS-PAGE of VE proteins under reducing and nonreducing conditions. A, under reducing conditions (lane a), only four main bands are resolved as follows: VEα, VEβ, VEγ, and a band with a higher molecular weight (HMWP) that corresponds to VEα-VEγ and VEβ-VEγ heterodimers. Under nonreducing conditions (lanes b and d) VE heterodimers are resolved. In lane b of the gel (−DTT) there are two VEα polypeptides (identified by MALDI-TOF-MS; data not shown). In lane d of the same gel there are two VEγ polypeptides. The proteins in lanes b and d are from two different VE preparations. The position of each VE protein and molecular weight standards (lane c) are shown. It should be noted that the apparent molecular weights of nonreduced VE proteins (lanes b and d) are lower than those of reduced proteins (lane a). B, schematic diagram of VE precursor proteins. Each VE proprotein contains an N-terminal SP and a PQ region, a P domain (absent in VEγ), a ZP domain, a CFLCS, and a C-terminal propeptide downstream of the CFLCS. During maturation, VE precursor proteins are post-translationally modified by removal of the SP, glycosylation of the ZP domain of VEγ, oxidation of Cys residues to disulfides, and removal of the C-terminal propeptide by cleavage at the CFLCS. The region containing the last Cys residue (boldface C in ZP domain) from mature VEα(Cys528), VEβ(Cys487), and VEγ(Cys406) that is upstream of CFLCS, the CFLCS, and the C-terminal propeptide are shown together with their respective sequences. The consensus sequence of the CFLCS of VEα and VEβ is RXXRX. The consensus sequence of the CFLCS of VEγ is RKXR. Each VE protein contains an additional potential proprotein convertase cleavage site (furin family) downstream of CFLCS that has the consensus sequence RX(0,2)R(K) (n = 0 or 2, highlighted in boldface). C, LC-MS(QTOF) analysis of the molecular weight of intact VEγ. The processed spectrum contains a peak with m/z 51,996 that corresponds to mature VEγ, starts with Gln21 and ends with Arg489 or Lys490. D, LC-MS(QTOF) analysis of the molecular weight of intact VEβ. The peak with m/z 47,600 that corresponds to mature VEβ, starts with Gln23, contains a 3-kDa molecular mass N-linked oligosaccharide, and ends within the CFLCS. E, diagram of the types of peptides detected by MS. Left panel indicates that peptide(s) detected by MS are part of mature VE proteins and end within the CFLCS. Middle panel indicates that peptide(s) detected by MS are part of the C-terminal propeptide from VE precursor proteins. Right panel indicates that peptide(s) detected by MS are part of VE precursor proteins that contain part of the protein upstream of the CFLCS, the CFLCS, and part of the C-terminal propeptide downstream of the CFLCS.
protein convertase that cleaves at the CFLCS. The predicted molecular masses of secreted VE proteins lacking the SP are 61, 56, and 47 kDa for VE/H9251, VE/H9252, and VE/H9253, respectively (not accounting for glycosylation of VE/H9253). Based on amino acid sequences after removal of C-terminal propeptides, mature VE proteins should have molecular masses of 58 kDa (VE/H9251), 52 kDa (VE/H9252), and 44 kDa (VE/H9253). In addition, VE/H9253 has 3 kDa of covalently linked N-linked oligosaccharide. By comparing theoretical molecular weights of VE proteins with and without the C-terminal propeptide and the molecular weights of VE proteins run under reducing conditions (Fig. 1A, lane a), it appears that VE proteins are proteolytically processed and do not contain the C-terminal propeptide. However, because molecular weights determined by SDS-PAGE may not be accurate, the molecular weights of individual polypeptides also were measured by LC-MS(QTOF) to confirm that VE proteins are processed at the C terminus. The molecular masses determined were 51,996 Da (VE/H9252) and 47,600 Da for VE/H9253 (Fig. 1D), in good agreement with results of SDS-PAGE analyses.

FIGURE 2. Analysis of the C-terminal region of mature VE proteins. LC-MS(QSTAR) analysis of AspN digests of VE proteins with peaks representing the monoisotopic m/z. A, the doubly charged peak with m/z 685.95 (calculated 685.80) that corresponds to a peptide with the sequence 481DSCEPRCYRK490 (Cys modified to CM-C) from VE/H9252, with a C-terminal Lys490 residue. The doubly charged peak with m/z 573.86 (calculated 573.80) that corresponds to a peptide with the sequence 484EPRCYRKR491 (Cys modified to CM-C; E as pyroglutamate) from VE/H9252, with a C-terminal Arg491 residue. B, the (2+) charged precursor ion with m/z 917.05 (calculated 916.84) (inset) that corresponds to a peptide with the sequence 396DGVCGCCDSTCSNRKG411 (Cys modified to CM-Cys) from VE/H9253, with a C-terminal Gly411 residue. The CID fragmentation spectrum of the parent ion confirms the identity of the peptide. C, the two doubly charged peaks with m/z 419.25 (calculated 419.17) and 427.28 (calculated 427.16) that correspond to a peptide with the sequence 525EQMCNR530 (Cys modified to CM-Cys and Met unmodified or modified to Met) from VE/H9253, with Arg530 as the C-terminal amino acid. The diagram in A represents the position of peptides detected within VE proteins from A to C.

FIGURE 2. Analysis of the C-terminal region of mature VE proteins. LC-MS(QSTAR) analysis of AspN digests of VE proteins with peaks representing the monoisotopic m/z. A, the doubly charged peak with m/z 685.95 (calculated 685.80) that corresponds to a peptide with the sequence 481DSCEPRCYRK490 (Cys modified to CM-C) from VE/H9252, with a C-terminal Lys490 residue. The doubly charged peak with m/z 573.86 (calculated 573.80) that corresponds to a peptide with the sequence 484EPRCYRKR491 (Cys modified to CM-C; E as pyroglutamate) from VE/H9252, with a C-terminal Arg491 residue. B, the (2+) charged precursor ion with m/z 917.05 (calculated 916.84) (inset) that corresponds to a peptide with the sequence 396DGVCGCCDSTCSNRKG411 (Cys modified to CM-Cys) from VE/H9253, with a C-terminal Gly411 residue. The CID fragmentation spectrum of the parent ion confirms the identity of the peptide. C, the two doubly charged peaks with m/z 419.25 (calculated 419.17) and 427.28 (calculated 427.16) that correspond to a peptide with the sequence 525EQMCNR530 (Cys modified to CM-Cys and Met unmodified or modified to Met) from VE/H9253, with Arg530 as the C-terminal amino acid. The diagram in A represents the position of peptides detected within VE proteins from A to C.
**Processing of Fish VE Proteins**

The SignalP prediction program (www.cbs.dtu.dk/services/SignalP/) predicts that the N-terminal amino acids of VE proteins are Gln23, Gln24, and Gln25 for VEα, VEβ, and VEγ, respectively. Based on the SignalP prediction program, previous analyses (7), SDS-PAGE (Fig. 1A, lane a), and LC-MS/MS (QTOF) analyses (Fig. 1, C and D), it appears that mature VEβ terminates within the CFLCS, probably with Arg489 or Lys490, and that VEγ also terminates within the CFLCS. In addition, the N-linked sugar chain of VEγ has an apparent mass of ~3 kDa, in agreement with previous reports (7). It should be noted that we were unable to determine by LC-MS/MS (QTOF) the mass of VE proteins that contain the C-terminal propeptide, probably because of the presence of very low amounts of precursors compared with proteolytically processed VE proteins. Also, when samples containing 1 (Fig. 1A, lane a) or 2 (Fig. 1A, lane b) VEα bands were analyzed, we were unable to assign their molecular weight; this was probably because of either a low signal intensity or breakdown of the proteins during analysis.

**Introduction to Mass Spectrometry Analyses of Rainbow Trout VE Proteins**—It is to be expected that the bulk of VE protein associated with rainbow trout eggs is mature, processed, and assembled, with only a relatively small fraction present as immature, unprocessed VE proprotein. In what follows, MS was used to identify the C-terminal amino acid of VE proteins present on rainbow trout eggs and to determine whether the C-terminal propeptide, downstream of the CFLCS, as well as an intact CFLCS, are present on at least a portion of the VE proteins surrounding eggs. To evaluate the mass spectra that follow (Figs. 2–7), the position of the peptide(s) within the VE proteins is diagrammed schematically in Fig. 1E, with the position of peptide(s) indicated by the arrow at the top of the diagram. For example, in the left panel of Fig. 1E, the peptide is part of the mature protein and ends within the CFLCS. In the middle panel of Fig. 1E, the peptide is part of the unprocessed proprotein and is part of the C-terminal propeptide. In the right panel of Fig. 1E, the peptide is part of the unprocessed proprotein and contains part of the mature protein (upstream of CFLCS), the intact CFLCS, and part of the C-terminal propeptide.

**Determination of the C-terminal Amino Acids of Mature VE Proteins**—Rainbow trout VE precursor proteins are synthesized by the liver and contain an N-terminal SP that directs them to the endoplasmic reticulum (ER). They also contain a C-terminal propeptide that is removed by a proprotein convertase (i.e., a furin-like protease) that acts at a CFLCS having the consensus sequence RXXR. Both the SP and propeptide are missing from the bulk of VE proteins that are assembled into an extra-cellular coat surrounding eggs (Fig. 1, C and D).

The C terminus of mature VEβ was determined by LC-MS(QSTAR) analyses of AspN digests of VEβ. A doubly charged (2+ peak) with m/z 685.95 (calculated 685.80) that corresponds to a peptide with the sequence 484DSCEPRCYR490 (Cys modified to carboxamidomethyl (CM)-C) with Lys490 as terminal amino acid was detected (Fig. 2A, left). Although no fragmentation of this peak was obtained, a peptide containing the same C-terminal amino acid was also detected by MALDI-TOF-MS (linear mode) analyses of AspN digests, where a singly charged peak with m/z 3271.69 (calculated 3271.72) that corresponds to a peptide with the sequence 484ETVFIHCNTAVCLPSLGDSCCEPRCYR490 (Cys modified to CM-C) was found (TABLE ONE). Two other C-terminal amino acids were found for VEβ. LC-MS(QSTAR) analyses of AspN digests of VEβ revealed a doubly charged peak with m/z 573.86 (calculated 573.80) (Fig. 2A, right) that corresponds to a peptide with the sequence 484EPRCYRKR491 (Cys modified to CM-C and Glu modified to pyroglutamate), with Arg489 as a C-terminal amino acid. However, this peak did not produce a fragmentation spectrum and was not confirmed by other analyses. Previously, we determined by MALDI-TOF-MS analyses of AspN digests that the C-terminal amino acid of mature VEβ is Arg489 (7). In this context, it should be mentioned that in another MALDI-TOF-MS analysis of an AspN digest of VEβ, a peak with m/z 1243.44 (calculated 1243.35; TABLE ONE) was found that corresponds to a peptide with the sequence 481DSCEPRCYR489 (Cys modified to CM-C). Therefore, the C-terminal residue of mature VEβ could be Arg489 or Lys490 or Arg491 that lie within the CFLCS.

The C-terminal amino acid of egg VEγ was determined by LC-MS(QSTAR) analyses of AspN digests, where a doubly charged peak with m/z 917.05 (calculated 916.84) (Fig. 2B, inset) was observed that corresponds to a peptide with the sequence 506DGVCCGCDSTCNRRG411 (Cys modified to CM-C), with Gly411 as the last amino acid. The CID produced a fragmentation spectrum that contained some y, b, and a ions (y2-2H2O, b2-H2O, Y4, y9-3H2O-3NH2, y12-12H2O(+2), y12-2NH2(+2), a10), and many internal fragments (e.g., NKr, GCS, DSTC-SNR-2H2O, GVCGCCDSTCSN that confirmed the identity of the peptide (Fig. 2B).

AspN digests of VE proteins were also analyzed, and y12–205 for VEβ, where a peptide with the sequence 1567.87 (calculated 1567.88) (Fig. 2B, inset) was observed that corresponds to a peptide with the sequence 503QRRR531 (Cys modified to CM-C and Met unmodified or modified to Met) (Fig. 2C), with Arg530 as the C-terminal amino acid. On the other hand, these peaks did not produce interpretable fragmentation spectra. These data suggest that the C-terminal amino acid for VEα could be Arg530 and is part of the first CFLCS sequence 503QRRR533. It should be mentioned that no cleavage sites at the second CFLCS sequence 541KKTK544 were found. Collectively, these data indicate that the C terminus of VE proteins lies within the CFLCS consensus sequence RXXR (i.e., RXXR for VEα and Arg542 for VEγ).
the second peptide, which is contiguous with the first, continues in the 
C-terminal propeptide to the end of the proprotein (Asn524). The presence 
of the C-terminal propeptide in VE<sub>α</sub>/H9252 was also confirmed by other MALDI-
TOF-MS analyses using various enzymatic digests (TABLE ONE).

| Protein | Residue # | Peptide Sequence | Enzyme | m/z Obs. | m/z Calc. | Charge | Method |
|---------|-----------|------------------|--------|---------|----------|--------|--------|
| VE<sub>α</sub> | 525-530 | EQMC*NR | AspN | 427.28 | 427.16 | (2+) | LC-MS |
| VE<sub>α</sub> | 525-530 | EQMC*NR | AspN | 419.25 | 419.17 | (2+) | LC-MS |
| VE<sub>α</sub> | 525-545 | EQMC*NRQRRDLASGQKTKTG | AspN | 630.65 | 630.57 | (4+) | LC-MS |
| VE<sub>α</sub> | 526-534 | QMCRQRQRD | GluC-V8 | 1223.15 | 1223.37 | (1+) | MALDI |
| VE<sub>α</sub> | 531-542 | QRRLDLASGQKK | Trypsin | 708.16 | 707.89 | (2+) | MALDI |
| VE<sub>α</sub> | 531-553 | QRRDLASGQKTKGTGVQVSSQK | Trypsin | 2544.40 | 2544.87 | (1+) | MALDI |
| VE<sub>α</sub> | 531-553 | QRRDLASGQKTKGTGVQVSSQK | Trypsin | 2544.38 | 2544.87 | (1+) | MALDI |
| VE<sub>α</sub> | 533-544 | RDLASGQKTKT | Trypsin | 1360.40 | 1360.55 | (1+) | MALDI |
| VE<sub>α</sub> | 534-544 | DLSAQGQKTKT | Trypsin | 602.07 | 602.33 | (2+) | MALDI |
| VE<sub>α</sub> | 534-553 | DLSAQGQKTKGTGVQVSSQK | Trypsin | 2104.83 | 2104.36 | (1+) | MALDI |
| VE<sub>α</sub> | 534-560 | DLSAQGQKTKGTGVQVSSQK | Trypsin | 2929.60 | 2929.40 | (1+) | MALDI |
| VE<sub>α</sub> | 534-560 | DLSAQGQKTKGTGVQVSSQK | Trypsin | 2945.65 | 2945.40 | (1+) | MALDI |
| VE<sub>α</sub> | 546-563 | DSVQSGQKTVIMDPFRYA | AspN | 517.53 | 517.52 | (4+) | MALDI |
| VE<sub>α</sub> | 554-560 | VIMIDPR | Trypsin | 422.46 | 422.24 | (2+) | MALDI |
| VE<sub>β</sub> | 464-490 | ETVFHC*NTAVC*LPSLGDC*EPRC*YRK | AspN | 3271.69 | 3271.72 | (1+) | MALDI |
| VE<sub>β</sub> | 481-489 | DSC*EPRC*YRK | AspN | 1243.44 | 1243.35 | (1+) | MALDI |
| VE<sub>β</sub> | 481-490 | DSC*EPRC*YRK | AspN | 685.95 | 685.80 | (2+) | MALDI |
| VE<sub>β</sub> | 484-491 | *EPRC*YRK | AspN | 573.86 | 573.80 | (2+) | MALDI |
| VE<sub>β</sub> | 487-503 | C*YRRDIPAVQKTR | Trypsin | 697.08 | 697.05 | (3+) | MALDI |
| VE<sub>β</sub> | 490-500 | KRDRIPAVQK | Trypsin | 641.31 | 641.38 | (2+) | MALDI |
| VE<sub>β</sub> | 490-503 | KRDRIPAVQKTR | Trypsin | 805.43 | 805.47 | (2+) | MALDI |
| VE<sub>β</sub> | 492-505 | RDIPAVQKTR | Trypsin | 1568.52 | 1568.87 | (1+) | MALDI |
| VE<sub>β</sub> | 492-505 | RDIPAVQKTR | Trypsin | 1568.72 | 1568.75 | (1+) | MALDI |
| VE<sub>β</sub> | 493-500 | DIPAVQK | Trypsin | 421.27 | 421.23 | (2+) | MALDI |
| VE<sub>β</sub> | 501-520 | TARIKNSLVSSEGELITDPR | Trypsin | 2171.13 | 2171.50 | (1+) | MALDI |
| VE<sub>β</sub> | 506-524 | SNLVSSEGELITDPR | Trypsin | 2059.37 | 2059.28 | (1+) | MALDI |
| VE<sub>β</sub> | 509-524 | VSSEGELITDPR | Trypsin | 872.17 | 872.45 | (2+) | MALDI |
| VE<sub>β</sub> | 464-490 | ETATVC*LPSLGDC*EPRC*YRK | AspN | 3271.69 | 3271.72 | (1+) | MALDI |

By using LC-MS(QSTAR) another peptide was detected from the 
C-terminal propeptide of VEβ (Fig. 3B). A doubly charged peak with 
m/z 805.43 (calculated 805.47) that corresponds to a peptide with the 
sequence<sup>490</sup>KRRDPAPAVQKTR<sup>506</sup> was found (Fig. 3B, inset). The
CID spectrum contained $y_1$, $y_2$, $y_3$, $y_5$, and $y_8$ ions (in intact form or with neutral loss of $H_2O$ or $NH_3$), $b_6$, and $a_{12}$ ions, as well as many internal fragments that confirm the identity of the peptide. Other peptides from the C-terminal propeptide of VEβ are shown in TABLE ONE. Therefore, some VEβ from eggs contains peptides from the C-terminal propeptide, downstream of the CFLCS.

**Determination of the C-terminal Propeptide of VEα**—To determine whether a C-terminal propeptide is present in other VE proteins, VEα and VEγ were analyzed. MALDI-TOF-MS (linear mode) was used to analyze tryptic digests of egg VEα (Fig. 4A). A peak was found with $m/z$ 2544.38 (calculated 2544.87) that corresponds to a peptide from the C-terminal propeptide with the sequence $^{531}$QRRDLASAQQKKKTDVIVVSSQYK$^{553}$. This peptide was also observed in other experiments as a peak with $m/z$ 2544.40 (calculated 2544.87) (TABLE ONE). In the same MALDI-TOF-MS (linear mode) spectrum (Fig. 4A), two additional peaks were found that correspond to a peptide from the C-terminal propeptide of VEγ. These peaks with $m/z$ 2929.80 (calculated 2929.40) and 2945.65 (calculated 2945.40) correspond to a peptide with the sequence $^{53}^DLSAQKTKGTVIVVSSQYKIMIDPR^{660}$, with its Met residue in either an oxidized or oxidized form. The presence of the C-terminal propeptide was also detected in other MALDI-TOF-MS experiments with various enzyme digests (TABLE ONE). Further evidence for the presence of the C-terminal propeptide was provided by tandem MS of LC-MS(QSTAR) analyses of AspN digests of VEα. A quadruply charged peak with $m/z$ 517.52 (calculated 517.52) was found that corresponds to peptide $^{549}^DYVIVVSSQYKIMIDPRFYA^{63}$(Fig. 4B, inset). The CID fragmentation spectrum (Fig. 4B) produced a series of peaks containing doubly charged b ions (b8, b9, b12, b13, and b16) and singly, doubly, or triply charged $y$ ions ($y_5$, $y_6$, $y_7$, $y_{10}$, $y_{12}$, $y_{13}$, $y_{15}$, and $y_{16}$) with neutral loss of $H_2O$, $NH_3$, or both that confirms the identity of this peptide. Other C-terminal peptides from VEα are shown in TABLE ONE. Therefore, some VEα from eggs possesses the C-terminal propeptide downstream of the CFLCS.

**Determination of the C-terminal Propeptide of VEγ**—Evidence for the presence of a C-terminal propeptide in VEγ was obtained by LC-MS(QTOF and QSTAR) analyses of tryptic and AspN digests (Fig. 5). In LC-MS(QTOF) analyses, a doubly charged (2+) ion with $m/z$ 1275.56 (calculated 1275.68) and a triply charged (3+) ion with $m/z$
850.71 (calculated 850.79) were found (Fig. 5A). These ions correspond to the same peptide from the C-terminal propeptide of VE\textsubscript{\textgamma} with the sequence 546 DVVVSSQKVIMIDPRFYA 563 from the C-terminal propeptide. The CID fragmentation spectrum of the doubly charged ion (\textsuperscript{2+}) confirms the identity of the peptide. It is of interest that the CID of these two parent ions (\textsuperscript{m/z} 2929.80) and oxidized (\textsuperscript{m/z} 2945.65) (Fig. 5, inset) that corresponds to C-terminal peptide with the sequence 546 DVVVSSQKVIMIDPRFYA 563 from VE\textsubscript{\textgamma} precursor protein. The CID fragmentation spectrum of the parent ion confirms the identity of the peptide.

**546 DVVVSSQKVIMIDPRFYA 563**

**FIGURE 4.** Analysis of the C-terminal propeptide from VE\textsubscript{\textgamma} precursor protein. A, MALDI-TOF-MS analysis (linear mode) of a tryptic digest of VE\textsubscript{\textgamma}. The singly charged, average peak \textsuperscript{m/z} 2544.38 (calculated 2544.87) (marked with an asterisk) corresponds to peptide \textsuperscript{517}HQKLVWEGDVQLGPIFH2O-NH3; WEGDVQLGPIFH2O-NH3) permitted confirmation of the identity of the peptide. These data were complemented by LC-MS(QSTAR) analysis of AspN digest of VE\textsubscript{\textgamma} and MALDI-TOF-MS analysis (linear mode) of a tryptic digest of VE\textsubscript{\textgamma} (7). MALDI-TOF-MS (linear mode) analyses of tryptic digests of VE\textsubscript{\textgamma} and VE\textsubscript{\textbeta}-VE\textsubscript{\textgamma} heterodimers were also carried out (Fig. 6). A singly charged (1+) peak with \textsuperscript{m/z} 1568.72 (calculated 1567.87) corresponds to a peptide with the sequence 492 RDIPAQQKTARIK 505 from the C-terminal propeptide of VE\textsubscript{\textbeta} in VE\textsubscript{\textbeta}-VE\textsubscript{\textgamma} heterodimers (Fig. 6A). Peaks corresponding to this particular peptide were also observed in MALDI-TOF-MS (linear mode) analyses of tryptic digests of VE\textsubscript{\textbeta} (Fig. 3A), as well as in tryptic digests of VE\textsubscript{\textbeta} analyzed by LC-MS(QSTAR) (Fig. 3B). These findings suggest that...
FIGURE 5. Analysis of the C-terminal propeptide from VEγ precursor protein. A, LC-MS(QSTAR) analysis of a tryptic digest of VEγ. The doubly charged (2+) ion with m/z 1275.56 (calculated 1275.68) and a triply charged (3+) ion with m/z 850.71 (calculated 850.79) that corresponds to peptide with the sequence 417HQKLVNIWEGDVQLGPIFISEK438 from the C-terminal propeptide of VEγ. The diagram in A represents the position of the peptide from the C-terminal propeptide of VEγ. B, CID spectrum of the doubly charged precursor ion (m/z 1275.56) from A produced a fragmentation spectrum that confirms the identity of the peptide. C, LC-MS(QSTAR) analysis of AspN digests of VEγ. A fragmentation spectrum of the doubly charged ion with m/z 741.92 (calculated 741.90) (inset) corresponds to a peptide with the sequence 413DTTKHQKLNIW424 from the C-terminal propeptide of VEγ. D, the triply charged peak with m/z 494.93 (calculated 494.94) (inset) observed in LC-MS(QSTAR) analysis of AspN digests of VEγ produced a fragmentation spectrum that corresponds to the same peptide detected in C, with the sequence 413DTTKHQKLNIW424 from the C-terminal propeptide of VEγ.
some VEβ from VEβ-VEγ heterodimers from eggs contains the C-terminal propeptide.

A singly charged (1+) peak with m/z 2456.10 (calculated 2456.84) from MALDI-TOF-MS (linear mode) analyses (Fig. 6A) corresponds to a peptide with the sequence 493-LVNIWEGDVQLGPIFISEKVAQ441 from the C-terminal propeptide of VEγ in VEβ-VEγ heterodimers. This peptide was also identified by LC-MS(QTOF) of tryptic digests of VEγ, where a triply charged (3+) ion with m/z 2456.10 (calculated 2456.84) (Fig. 6A) corresponds to the same peptide from the C-terminal propeptide of VEγ (TABLE ONE). Another singly charged (1+) peak with m/z 3210.40 (calculated 3210.65) (Fig. 6A) corresponds to a C-terminal peptide from VEγ with the sequence 411-GRDTPKHKLTIWEGDVQLGPIFISEK438. This particular peptide was also detected in another MALDI-TOF-MS (linear mode) analysis of tryptic digests of VEγ heterodimers. A singly charged average peak with m/z 3210.40 (calculated 3210.65) (Fig. 6A) corresponds to a peptide from the C-terminal propeptide of VEγ in VEβ-VEγ heterodimers. The diagram in A represents the position of peptides detected from the C-terminal propeptides of VEα-VEγ and VEβ-VEγ heterodimers.

**Processing of Fish VE Proteins**

**Takes Place on the Egg**—VE precursor proteins that contain a C-terminal propeptide are synthesized by the liver, secreted into the bloodstream, and transported to the ovary. Although unlikely, it is possible that the propeptide could have been cleaved from the protein prior to reaching the ovary and remained attached by noncovalent interactions. Here we used MS analyses of VE proteins isolated from eggs to look for peptides that contain a portion of mature VE proteins (i.e. upstream of the CFLCS), the CFLCS, and a portion of the C-terminal propeptide (i.e. downstream of the CFLCS). Such a peptide was identified by LC-MS(QTOF) analyses of tryptic digests of VEβ (Fig. 7A). A triply charged (3+) peak with m/z 697.08 (calculated 697.05) corresponds to a peptide with the sequence 467-CYRKRDPIAAVQKTAR482 (Cys modified to CM-C) (Fig. 7A, inset). The peptide contains the last conserved Cys residue from the ZP domain, the CFLCS, and a portion of the C-terminal propeptide. CID produced a fragmentation spectrum in which some y ions (y5-H2O; y9-H2O; y15-H2O; y15-2H2O-2NH3), b ions (b4-NH3; b7-NH3; b11-NH3; b15-3NH3; b16-H2O-3NH3), and some internal fragments (RRDIPAAVQKTA-NH3, KRR-
DIPAAVQ-4NH$_3$) (Fig. 7A) confirmed the identity of this peptide. It is of some interest that as a result of fragmentation of the parent ion with $m/z$ 697.08, most of the b ions had single or multiple neutral loss of NH$_3$, whereas most of the y ions had neutral loss of H$_2$O. The presence of a high intensity singly charged (1+ peak of $m/z$ 297.20 (Fig. 7A) could not be assigned.

Peptides that contain a portion of mature VE proteins, the CFLCS, and a portion of the C-terminal propeptide were also detected in VE$_3$, and VE$_y$. In LC-MS(QTOF) analyses of tryptic digests of VE$_y$, a triply charged (3+) peak with $m/z$ 1009.77 (calculated 1009.76) was identified (TABLE ONE) that corresponds to a peptide with the sequence 525EQMCNRQRRDLSAQGQKKTKG545 (Cys modified C-M-C) that contains part of the mature protein, the CFLCS, and a portion of the C-terminal propeptide. However, we were unable to interpret its CID fragmentation spectrum. On the other hand, MALDI-TOF-MS analyses (linear mode) of GluC(V8) digests of VE$_x$ revealed a singly charged peak with $m/z$ 1223.15 (calculated 1223.37) (TABLE ONE) that corresponds to a peptide with the sequence 525EQMCNRQRRDL-SAQGQKKTKG545 (Cys modified C-M-C). This peptide contains the last conserved Cys residue of the ZP domain, the CFLCS, and the amino acid Asp that is part of the C-terminal propeptide.

Another peptide that contains a portion of mature VE proteins, the CFLCS, and a portion of the C-terminal propeptide was detected by LC-MS(QSTAR) analysis of AspN digests of VE$_x$ (Fig. 7B). A quadruply charged peak $m/z$ 630.65 (calculated 630.57) (Fig. 7B, inset) that corresponds to a peptide with the sequence 525EQMCNRQRRDLSAQGQKKTKG545 (Cys modified to CM-C) was found. The CID spectrum (Fig. 7B) contains peaks that correspond to singly or multiply charged b ions (b15, b17, b18, b19, and b20), a ions (a2 and a6), and y ions (y3, y7, y8, y14, y15, y17, y18, and y19) (with neutral loss of H$_2$O, NH$_3$, or both), confirming the identity of the peptide. Collectively, these results suggest that removal of the C-terminal propeptide from VE precursor proteins takes place on the egg.

**DISCUSSION**

VE precursor proteins are secreted from hepatocytes into the bloodstream by female rainbow trout and by estrogen-induced males (17, 21, 34, 35). VE proteins are present in the blood as precursor proproteins (8, 17, 21, 23–27). After transport to the ovary through the bloodstream, these precursors are not internalized by eggs (17) but are incorporated into the innermost layer of the thickening VE (18). Rainbow trout VE$_x$, VE$_B$, and VE$_y$ are
secreted as precursor proteins, with a short C-terminal propeptide downstream of the CFLCS, that undergo various post-translational modifications. Synthesis of the proteins takes place in the liver (34), although VEγ messenger-RNA, but not protein, has been detected in the ovary (8). The initial step in processing of VE precursor proteins is removal of the N-terminal SP that directs them to the ER. After formation of intramolecular disulfides and glycosylation of VEγ, VE precursor proteins are secreted by the liver and transported in the bloodstream as proproteins to the ovary.

The site of proteolytic processing at the CFLCS and assembly of VE precursor proteins following secretion by the liver is not clear. Consequently, we employed a peptidomics-based approach to determine the site of processing at the CFLCS of VE proteins and the relationship between their processing and assembly. In general terms, peptidomics refers to alternative strategies for comprehensive analysis of peptide mixtures that are fractionated by liquid chromatography, analyzed by MS/MS, and identified by database searches (36, 37). To clarify aspects of processing and assembly of the VE precursor proteins, we analyzed cryptic digests (i.e. tryptic, Glu(V8), AspN, and chymotryptic digests) of VEα, VEβ, VEγ, and heterodimers of VE proteins by MS (i.e. MALDI-TOF-MS in both a linear and reflective mode, LC-MS(QTOF), and LC-MS(QSTAR)) (38). Collectively, the MS evidence strongly suggests that rainbow trout VE proteins undergo proteolytic processing at the CFLCS some time after their arrival at eggs in the ovary.

VE precursor proteins, like many other ZP domain-containing proteins, are cleaved at their CFLCS by a proprotein convertase or furin-like protease (39, 40). To assemble into higher order structures such as fibrils or filaments, the C-terminal propeptide downstream of the CFLCS must be removed (23, 28). The C-terminal amino acid of mature (i.e. processed) ZP domain proteins lies within a CFLCS that has the consensus sequence RXR(R/K)(Arg1-X1-X2-Arg2/Lys)). Although most studies have found that the C terminus of ZP domain proteins lies within the CFLCS, the C-terminal amino acid can be Arg1 (7) or before Arg1 (7, 41, 42), X1 (7, 43–45), X2 (23), or Arg2/Lys (46). In addition, the C-terminal amino acid of mature porcine ZPC (ZP3) is Ser532 in the sequence SRK, upstream of the conserved CFLCS but downstream of the ZP domain, suggesting that proteolytic processing may also take place at cleavage sites within dibasic motifs (42). In uromodulin the C-terminal amino acid of the mature polypeptide is Phe548 with ES as the proteolytic cleavage site (47). Most interestingly, this cleavage site lies within the ZP domain between conserved Cys6 and Cys7 residues and is not a CFLCS or a dibasic motif, but it does resemble the substrate of a chymotrypsin-like protease. In a previous report (7), it was found that cleavage of the C-terminal propeptide of fish VE proteins takes place at a proprotein convertase cleavage site having the general consensus sequence (K/R)(Xn,R,K/R), with n = 0, 2, 4, or 6, as suggested by others (48). Differences in reported results are apparently due to trimming by a carboxypeptidase following cleavage at the CFLCS, as suggested for bird (41, 45), amphibian (41), fish (7), and mammalian (42) egg coat proteins. However, if proprotein convertases act at the more general consensus RXR (n = 0), more than one amino acid may be the cleavage site for the enzyme(s). For example, for VEβ that has the cleavage site sequence RKRR (Arg1-X1-X2-Arg2), the proprotein convertases could cleave after X1, X2, or Arg2.

MS and database searches (36, 49–51) allow identification only of peptides that are present in the data base and cannot identify peptides derived from post-translationally modified proteins (52). Consequently, AspN-derived peptides that correspond to the C terminus of mature VE proteins are not identifiable in any database because there are no AspN cleavage sites in the CFLCS. However, by using manual searches of MS and MS/MS data, we found that the major C-terminal amino acid lies within the CFLCS and may be any one of the four CFLCS amino acids (Arg1-X1-X2-Arg2/Lys), confirming a previous report (7). Our results support those of other studies in which most mature ZP domain proteins terminate within the CFLCS (7, 23, 41–46). For example, proteolytic processing of quail ZPC (ZP3) at the CFLCS is mandatory for its secretion and unprocessed ZPC proprotein accumulates in the ER (53). Similar results have been reported for mouse (54) and human (55) ZP3.

On the other hand, proteolytic processing of ZP domain proteins may take place at furin-like cleavage sites other than the CFLCS (42, 56). Furthermore, there is a report that mutation of the CFLCS does not affect secretion or assembly of mouse ZP3 (56), and mutation of the CFLCS of recombinant human ZP3 diminishes but does not abolish its secretion (55). Although there are additional furin-like cleavage sites downstream of the CFLCS in fish VE proteins (Fig. 1), we found no evidence for any additional C-terminal amino acids other than those from the CFLCS.

It has been demonstrated that a small amount of VE protein can be incorporated into covalently cross-linked heterodimers (7), but it is not clear where and when dimerization takes place. In most cases, secreted VE precursor proteins in the bloodstream are monomers (23, 34, 35, 57); however, there is some evidence that heterodimers may be present as well (25, 34, 35, 57). In the gilthead sea bream, sea bass, and Atlantic halibut VE monomers are mainly transported in the bloodstream as HMWPs (25, 27, 57). It was suggested that when HMWPs arrive at the ovary they are cleaved to lower molecular weight VE proteins in preparation for assembly. In the bloodstream of rainbow trout and other fish there are no VE heterodimers (17, 21, 23, 34, 35), and the mononeric VE precursors in the bloodstream are monomers (23, 34, 35); however, there is some evidence that heterodimers may be present as well (25, 34, 35, 57). In the gilthead sea bream, sea bass, and Atlantic halibut VE monomers are mainly transported in the bloodstream as HMWPs (25, 27, 57).

There are significant differences between VE assembly in some fish and ZP assembly in mammals. For example, ZP protein precursors are synthesized in the ovary by growing oocytes and/or follicle cells and have a C-terminal transmembrane (TM) domain that anchors the proteins in egg plasma membrane prior to assembly (28). On the other hand, trout VE protein precursors are synthesized in the liver, are transported to the ovary, and do not have a TM domain (28). Despite the differences, nascent mammalian ZP proteins and fish VE proteins are deposited on the inside margin of the extracellular coats (18, 60). Most interestingly, when ZP proteins are truncated just upstream of the TM domain, the proteins lacking a TM domain are secreted but are neither cleaved at the CFLCS or incorporated into the ZP (6, 28). The evidence suggests that the TM domain 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 protein
precursors that lack a TM domain undergo both cleavage at the CFLCS and assembly into the VE once they reach the ovary (23). Because, like ZP proteins, VE protein precursors possess both an EHP and IHP (28), presumably these elements prevent premature polymerization of VE proteins into fibrils/filaments in the bloodstream. As a result of proteolytic processing at the CFLCS in the ovary, the EHP is lost with the propeptide, the IHP in the ZP domain is exposed, and mature VE proteins are able to polymerize around eggs.

Any of interesting questions about VE assembly remain. Perhaps principal among these is the question of how some fish VE protein precursors 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 (61). 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 α₂-macroglobulin into growing oocytes (62). 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 CFLCS in the ovary. It is possible that the enzyme is associated with the oocyte plasma membrane that is close to the innermost layer of the VE into which nascent, processed VE proteins are incorporated. The enzyme could be the receptor for VE proproteins, which would explain why rainbow trout VE proproteins lack a TM domain. These and other issues will be the focus of future studies of VE formation in fish.

Acknowledgments—We thank Norman Soule and the staff of the Cold Spring Harbor Fish Hatchery (Cold Spring Harbor, NY) for the hospitality and for providing buckets of rainbow trout eggs. C. C. D. also thanks Alisa G. Woods (Padure Biomedical Consulting, Brooklyn, NY) for advice and discussion.

REFERENCES

1. Dumont, J. N., and Brummett, A. R. (1985) in Developmental Biology: A Comprehensive Synopsis (Browder, E. R., ed) pp. 235–288, Plenum Publishing Corp., New York
2. Jovine, L., Litscher, E., and Wassarman, P. M. (2002) in Advances in Developmental Biology and Biochemistry (DePamphilis, M. L., ed) Vol. 12, pp. 31–54, Elsevier Science Publishers B.V., Amsterdam
3. Jovine, L., Darie, C. C., Litscher, E. S., and Wassarman, P. M. (2005) Annu. Rev. Biochem. 74, 83–114
4. Bork, P., and Sander, C. (1992) FEBS Lett. 300, 237–240
5. Wassarman, P. M., Jovine, L., and Litscher, E. S. (2001) Nat. Cell Biol. 3, E59–E64
6. Jovine, L., Qi, H., Williams, Z., Litscher, E., and Wassarman, P. M. (2002) Nat. Cell Biol. 4, 457–461
7. Darie, C. C., Biniossek, M. L., Jovine, L., Litscher, E. S., and Wassarman, P. M. (2004) Biochemistry 43, 7459–7478
8. Hyllner, S. J., Westerlund, L., Olsson, P. E., and Schopen, A. (2001) Biol. Reprod. 64, 805–811
9. Carr, M. D., Bauer, C. J., Gradwell, M. J., and Feeney, J. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 2206–2210
10. Blei, J. D., and Wassarman, P. M. (1980) Proc. Natl. Acad. Sci. U. S. A. 77, 1029–1033
11. Kubo, H., Kawano, T., Tsuki, S., Kawashima, S., Katagiri, C., and Suzuki, A. (1997) Dev. Growth Differ. 39, 405–417
12. Yang, J. C., and Hedrick, J. L. (1997) Dev. Growth Differ. 39, 457–467
13. Wang, H., and Gong, Z. (1999) Biochem. Biophys. Acta 1446, 156–160
14. Mold, D. E., Kim, I. F., Tsai, C. M., Lee, D., Chang, C. Y., and Huang, R. C. (2001) Mol. Reprod. Dev. 58, 4–14
15. Chang, Y. S., Wang, S. C., Tsao, C. C., and Huang, F. L. (1996) Mol. Reprod. Dev. 44, 295–304
16. Chang, Y. S., Hsu, C. C., Wang, S. C., Tsao, C. C., and Huang, F. L. (1997) Mol. Reprod. Dev. 46, 258–267
17. Hyllner, S. J., and Haux, C. (1992) J. Endocrinol. 135, 303–309
18. Hamazaki, T. S., Nagahama, Y., Iuchi, I., and Yamagami, K. (1988) Biol. Chem. 133, 101–110
19. Lyons, C. E., Payette, K. L., Price, J. L., and Huang, R. C. (1993) J. Biol. Chem. 268, 21351–21358
20. Murata, K., Sugiya, H., Yanagihara, K., Iuchi, S., Yamagami, K., Krasnov, J., Stavros, T., and Giancotti, F. (1996) J. Biol. Chem. 271, 295–304
21. Del Gaudio, L., Vanoni, C., Bonsignorio, D., Duga, S., Mosconi, G., Santucci, A., and Cotelli, F. (1998) Mol. Reprod. Dev. 49, 58–69
22. Sasanami, T., Nagahama, Y., Iuchi, I., Enami, J., Yamagami, K., and Yamagami, K. (1999) J. Biochem. (Tokyo) 125, 469–475
23. Sasanami, T., Pan, J., Doi, Y., Hisada, M., Kohsaka, T., and Toriyama, M. (2002) Biochim. Biophys. Acta 1562, 7459–7478

Processing of Fish VE Proteins

Why rainbow trout VE proproteins lack a TM domain. These and other issues will be the focus of future studies of VE formation in fish.

VOLUME 280 • NUMBER 45 • NOVEMBER 11, 2005
JOURNAL OF BIOLOGICAL CHEMISTRY 37598