Structural Characterization of Fish Egg Vitelline Envelope Proteins by Mass Spectrometry†

Costel C. Darie,‡ Martin L. Biniossek,§ Luca Jovine,‡ Eveline S. Litscher,‡ and Paul M. Wassarman*,‡

Brookdale Department of Molecular, Cell, and Developmental Biology, Mount Sinai School of Medicine, One Gustave L. Levy Place, New York, New York 10029-6574, and Institut fuer Molekulare Medizin und Zellforschung, Albert-Ludwigs Universitaet, Stefan-Meier-Strasse 19, 79104 Freiburg, Germany

Received February 27, 2004; Revised Manuscript Received April 6, 2004

ABSTRACT: The extracellular coat, or vitelline envelope (VE), of rainbow trout (*Oncorhynchus mykiss*) eggs consists of three proteins, called VEα (M_r ∼52 kDa), VEβ (M_r ∼48 kDa), and VEγ (M_r ∼44 kDa). Each of these proteins is related to mammalian egg zona pellucida (ZP) glycoproteins ZP1−3 and possesses an N-terminal signal sequence, a ZP domain, and a protease cleavage site near the C-terminus. VEα and VEβ also have a trefoil domain. All three proteins possess a relatively large number of cysteine residues (VEα, 18; VEβ, 18; VEγ, 12), of which 8 are present in the ZP domain and 6 are present in the trefoil domain of VEα and VEβ. Here, several types of mass spectrometry were employed, together with gel electrophoresis of chemical and enzymatic digests, to identify intramolecular disulfide linkages, as well as the N- and C-terminal amino acids of VEα, VEβ, and VEγ. Additionally, these methods were used to characterize two high molecular weight proteins (HMWPs; M_r > 110 kDa) of rainbow trout VEs that are heterodimers of individual VE proteins. These analyses have permitted assignment of disulfide linkages and identification of N- and C-terminal amino acids for the VE proteins and determination of the protein composition of two forms of HMWPs. These experiments provide important structural information about fish egg VE proteins and filaments and about structural relationships between extracellular coat proteins of mammalian and nonmammalian eggs.

The plasma membrane of vertebrate eggs is surrounded by an extracellular envelope, often referred to as a chorion, vitelline envelope (VE),[1] or zona pellucida (ZP) (*1*, *2*). In general, the coats surrounding fish eggs and mammalian eggs have been designated VE and ZP, respectively. These extracellular coats apparently function to restrict fertilization of eggs to sperm from the same species, to limit fertilization to a single sperm, and to protect developing embryos either outside or within the female reproductive tract. In some cases, these functions have been assigned to particular VE or ZP proteins, and the mechanisms of action have been investigated (*3*, *4*). In recent years, extensive molecular analyses of egg coat proteins from many animal species have revealed that these proteins are highly conserved and are related to each other. A major feature common to the proteins is the presence of a so-called ZP domain, a sequence of ∼260 amino acids containing 8 conserved cysteine (C) residues (*5*−*7*).

A great deal is known about the protein composition of fish egg VEs (*8*−*10*). For example, rainbow trout (*Oncorhynchus mykiss*) VEs consist of at least three proteins, called VEα, VEβ, and VEγ, that possess an N-terminal signal sequence, a proline−glutamine- (PQ-) rich region, and a ZP domain that is common to all ZP-like proteins (*5*, *6*, *11*−*13*). VEα and VEβ also have a trefoil domain, just upstream of the ZP domain (*14*−*17*). VE proteins contain a large number of C residues (VEα, 18; β, VE18; VEγ, 12), and the first 6 C residues of VEα and VEβ apparently are present as disulfides (C1−C4, C2−C5, C3−C6) in trefoil domains (*18*). Disulfide linkages for the remaining 12 C residues have not been reported as yet, although partial disulfide assignments for mouse ZP glycoproteins, ZP1−3, have been reported recently (*19*).

Here, we carried out structural characterization of rainbow trout egg VE proteins, VEα, VEβ, and VEγ, by using mass spectrometry measurements (MALDI-TOF-MS and ESI-Q-TOF-MS) and gel electrophoresis of chemical and enzymatic digests of the proteins. This experimental approach enabled us to identify the intramolecular disulfide linkages, N- and C-terminal amino acids, and the N-linked glycosylation site of rainbow trout VE proteins. In addition, monoclonal antibodies directed against fish VE proteins were used to determine the protein composition of high M_r components of rainbow trout VEs. The results obtained provide consider-

† Supported in part by a grant from the NIH (HD-35105 to P.M.W.). L.J. is a postdoctoral fellow supported by a long-term fellowship from the Human Frontier Science Program Organization.

* To whom correspondence should be addressed. Tel: 212-241-8616. Fax: 212-860-9279. E-mail: paul.wassarman@mssm.edu.

‡ Mount Sinai School of Medicine.

§ Albert-Ludwigs Universitaet.

[1] Abbreviations: VE, vitelline envelope; ZP, zona pellucida; ZP domain, zona pellucida domain; HMWP, high molecular weight protein; M_r, molecular weight; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; FA, formic acid; BNPS-skatole, BNPS-[2-(2-nitrophenylsulfenyl)-3-methyl-3-bromoindolene]; DTT, dithiothreitol; NR, nonreducing; R, reducing; M, oxidized methionine; *m*/*z*, mass/charge; MS, mass spectrometry; MALDI-TOF-MS, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; ESI-Q-TOF-MS, electrospray ionization quadrupole time-of-flight mass spectrometry; CID, collision-induced dissociation.

10.1021/bi0495937 CCC: $27.50
Published on Web 05/15/2004

able insight into both the specificity of assembly of ZP domain proteins in general and the interactions of VEα, VEβ, and VEγ in trout egg VE filaments.

MATERIALS AND METHODS

Reagents. Reagents were obtained from the following commercial sources: MALDI-TOF-MS standards, chymotrypsin, AspN, GluV8, pepsin, goat anti-mouse IgG−HRP, formic acid, BNPS-skatole (BNPS-[2-(2-nitrophenylsulfenyl)-3-methyl-3-bromoindolenine]), trifluoroacetic acid, and acetonitrile from Sigma-Aldrich (St. Louis, MO), trypsin from Roche Applied Science (Indianapolis, IN), peptide-*N*-glycosidase (PNGase F) from New England Biolabs (Beverly, MA), nitrocellulose membranes and enhanced chemiluminescence kits from Amersham Pharmacia Biotech (Piscataway, NJ), SDS−PAGE gels from Bio-Rad (Hercules, CA), and M_r markers from Bio-Rad, Invitrogen (Carlsbad, CA), and New England Biolabs.

Isolation of VE Proteins. The procedure described by Brivio et al. (*13*) was used to isolate VEs from rainbow trout (*O. mykiss*) eggs. To prepare enzyme digests of VE proteins, VEs were subjected to SDS−PAGE, and gel pieces containing individual VE proteins were excised and then digested with various proteases. To prepare chemical digests of VE proteins, VEs were subjected to preparative SDS−PAGE, gel pieces containing individual VE proteins were excised and electroeluted (Biotrap; Schleicher and Schuell, Keene, NH) in 25 mM Tris-HCl, pH 8.8, 192 mM glycine, and 0.025% (w/v) SDS for 12−16 h at room temperature (*20*), and individual VE proteins were subjected to chemical digestion.

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. (*21*) with certain modifications (*22*). Gel pieces containing 0.5−2 μg of protein were incubated with 60% (v/v) acetonitrile 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 C residues in the 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. Gel pieces were dehydrated, dried, and rehydrated twice. For assignment of disulfide linkages, the reduction and alkylation steps were not carried out; to prevent disulfide exchange, the pH was 6.5. 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 R and K residues), chymotrypsin (cleavage at the C-terminus of F, Y, W, M, and L residues), endoproteinase AspN (cleavage at the N-terminus of D and E residues), or endopeptidase GluV8 (cleavage at the C-terminus of E and D residues). For trypsin and chymotrypsin digestions, 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% acetonitrile/5% formic acid in 25 mM ammonium bicarbonate, pH 8.0, and shaking for 60−90 min at room temperature. Digestion by pepsin (cleavage at the C-terminus of F, Y, W, and L residues) was carried out in the same digestion buffer, adjusted with trifluoroacetic acid to pH 2.0. Solutions containing peptides of VE proteins were pooled, dried, and used for MALDI-TOF-MS in both linear and reflective modes and/or ESI-Q-TOF-MS.

Chemical Digestion of VE Proteins. Chemical digestions of individual VE proteins were carried out essentially according to published procedures (*23*, *24*).

To digest with formic acid (FA; cleaves the peptide bond between D and P residues), 15−30 μg of a VE protein (0.3−0.5 mg/mL) was diluted with 95% (v/v) FA, to a final concentration of 70% FA, and was vortexed and incubated in a fume hood in the dark for 48 h at room temperature. The solution was diluted 1:3 (v/v) with distilled water and dried completely. The dry sample was solubilized in 50 mM Tris-HCl/4 M urea, pH 8.0, and analyzed on 16% Tris−Tricine−PAGE gels.

To digest with BNPS-skatole (cleavage at the C-terminus of W residues), 15−30 μg of a VE protein (0.3−0.5 mg/mL) was diluted with 1% acetic acid to 0.1 mg/mL. BNPS-skatole (15−30 μL) (10 mg/mL dissolved in acetic acid) was then added to the samples at a final protein:BNPS-skatole ratio of 1.5:10, and samples were incubated in the dark for 30−60 min at room temperature. At the end of the incubation, samples were diluted 1:3 (v/v) with distilled water and were centrifuged at 10000 rpm for 5 min. Supernatants were collected and dried to 20−25% of the initial volume, diluted 1:5 with distilled water, and completely dried. The dry sample was solubilized in 50 mM Tris-HCl/4 M urea, pH 8.0, and analyzed either on 16% Tris−Tricine−PAGE gels or by SDS−PAGE.

Gel Electrophoresis and Western Blotting. VE proteins were separated by SDS−PAGE (*25*) under reducing (R) and nonreducing (NR) conditions. Tris−Tricine−PAGE was performed on 16% gels by using the procedure described by Schaegger et al. (*26*). For two-dimensional SDS−PAGE, samples were dissolved without R agents and separated on 10% gels. Lanes containing separated polypeptides were excised, placed in a solution containing 1% SDS (w/v) and 1% (v/v) β-mercaptoethanol, and then placed on top of another SDS−PAGE gel and subjected to electrophoresis (15−20 mA) in a second dimension. Gels were then stained with Coomassie blue, zinc, or silver (*27*). For immunoblotting, proteins were transferred onto nitrocellulose membrane at 80 V for 12 h at 4 °C and stained with colloidal gold. Membranes were incubated with primary monoclonal antibodies directed against *Salmo salar* zona radiata proteins (anti-VEα, MN-8C4; anti-VEβ, MN2B4; anti-VEγ, MN-7F2; Biosense Laboratories AS, Bergen, Norway), followed by goat anti-mouse IgG−horseradish peroxidase, and examined by enhanced chemiluminescence (ECL).

N-Terminal Sequence Analysis. Automated Edman degradation was carried out by using an Applied Biosystems 494 amino acid sequencer.

Matrix-Assisted Laser Desorption/Ionization (MALDI) Time-of-Flight (TOF) Mass Spectrometry (MS). MALDI-TOF-MS analysis of VE peptides was carried out on a TOF Voyager-DERP mass spectrometer (PerSeptive Biosystems, Foster City, CA) in linear mode and on a Reflex III TOF system (Bruker Daltonics, Leipzig, Germany) in reflective mode, both equipped with a nitrogen laser (337 nm). Spectra were obtained in either linear mode or reflective mode.

Samples were dissolved in 50% acetonitrile (linear mode) and cocrystallized with an α-cyano-4-hydroxycinnamic acid (HCCA) matrix on a stainless steel target. In reflective mode, the samples were prepared with a modified thin-layer method (*28*). The dry samples were dissolved with 2% trifluoroacetic acid. A mixture of HCCA and nitrocellulose was used as a matrix.

After drying, samples were analyzed undiluted or at suitable dilutions. Peptide data were collected after ionization 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 trypsin fragments. Peaks detected by MALDI-TOF-MS in linear mode corresponded to average mass/charge (*m*/*z*) peptides. The peaks detected by MALDI-TOF-MS in reflective mode corresponded to monoisotopic mass/charge (*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 linear mode, calibration differences up to 1 Da were observed, whereas calibration differences for measurements in reflective mode were generally 100 ppm. Interpretation of mass spectra was performed using Mascot (http://www.matrixscience.com) and FindMod, FindPep, and PeptideMass (http://www.expasy.ch), as well as manually.

Electrospray Ionization (ESI) Quadrupole (Q) Time-of-Flight (TOF) Mass Spectrometry (MS). Analyses were carried out with a Micromass-Q-TOF hybrid Q/TOF mass spectrometer 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 50 μm (inner diameter) fused-silica tubing to the liquid chromatography (LC) detector outlet. An LC Packings system, equipped with an Ultimate micropump 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% acetonitrile in 0.1% FA. The peaks detected by ESI-Q-TOF-MS corresponded to monoisotopic mass-to-charge (*m*/*z*) peptides. Processed files were subjected to a Mascot search (http://www.matrixscience.com).

Assignment of Disulfide Linkages to VE Proteins. Peptides generated in protease digests (carried out at pH 6.5 to minimize disulfide interchange) of gel slices containing individual VE proteins were analyzed by MALDI-TOF-MS in linear and reflective modes and by ESI-Q-TOF-MS. For MALDI-TOF-MS measurements in linear mode, samples were divided into nonreduced (NR conditions) and reduced (R conditions; 10 mM DTT, 40 min, 37 °C) portions. Spectra were compared for peak shifts due to the reduction step. Spectra that resulted from MALDI-TOF-MS in linear mode were then compared with those from MALDI-TOF-MS in reflective mode and with those from ESI-Q-TOF-MS. To identify disulfide linkages, masses of all theoretical combinations of C residue-containing peptides in a particular enzymatic digest were calculated. These values were then compared with the *m*/*z* of peaks from MALDI-TOF-MS spectra under NR and R conditions. Disappearance of a candidate peak under R conditions and appearance of at least one of the two C residue-containing peptides in the spectrum was taken as evidence for the presence of a disulfide.

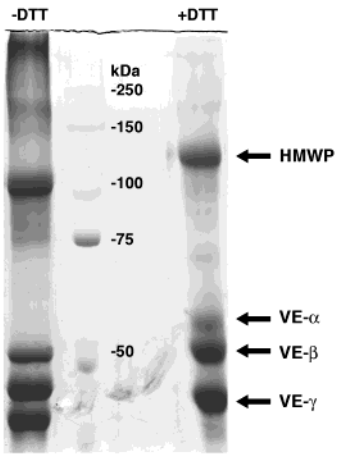

FIGURE 1: Coomassie-stained VE proteins after 10% SDS−PAGE under NR (−DTT) and R (+DTT) conditions. Four main bands are resolved: VEα, VE*β*, VE*γ*, and HMWP, with a higher electrophoretic mobility under NR conditions. The position of each VE protein and M_r standards are shown.

Polypeptide samples generated in chemical digests were divided into two batches and analyzed by Tris−Tricine−PAGE under NR and R conditions. Polypeptides were also separated by two-dimensional SDS−PAGE.

RESULTS

Electrophoretic Analysis of Rainbow Trout VE Proteins. Under NR conditions on SDS−PAGE, rainbow trout VEs contain four major proteins, three of which are called VEα (M_r ∼52 kDa), VE*β* (M_r ∼48 kDa), and VE*γ* (M_r ∼44 kDa). In addition, a fourth protein (M_r ∼110 kDa) is present that may be related to VE*γ* (*13*) and that is referred to here as HMWP (*h*igh *m*olecular *w*eight *p*rotein). As expected, and consistent with the idea that their C residues are only involved in intramolecular disulfides, under R conditions on SDS−PAGE the M_rs of these four proteins shift to ∼58 (VEα), ∼54 (VE*β*), ∼47 (VE*γ*), and ∼120−130 (HMWP) kDa (Figure 1). VE*β*, VE*γ*, and HMWP are stained by Coomassie blue to almost the same intensity, whereas VEα is stained less intensely, perhaps suggesting a less than stoichiometric ratio of VE proteins.

Overview of Mass Spectrometry Measurements. Each VE protein was analyzed individually by MALDI-TOF-MS in linear and reflective modes and identified in MSDB, NCB Inr, and/or SwissProt databases by using Mascot peptide mass fingerprinting (http://www.matrixscience.com). Sequence coverage was improved by using FindPept and FindMod (http://www.expasy.ch) programs. There is no reported sequence for HMWPs.

MALDI-TOF-MS in linear mode was used initially to assign the disulfides of VE proteins. Several criteria were imposed to ensure that disulfide linkages were assigned accurately. (1) Two C residues that constitute a disulfide cannot participate in other disulfides. For example, if a protein contains 4 disulfide-linked C residues and a disulfide is detected between C-3 and C-4, these residues cannot form disulfides with C-1 and/or C-2. (2) Every potential disulfide

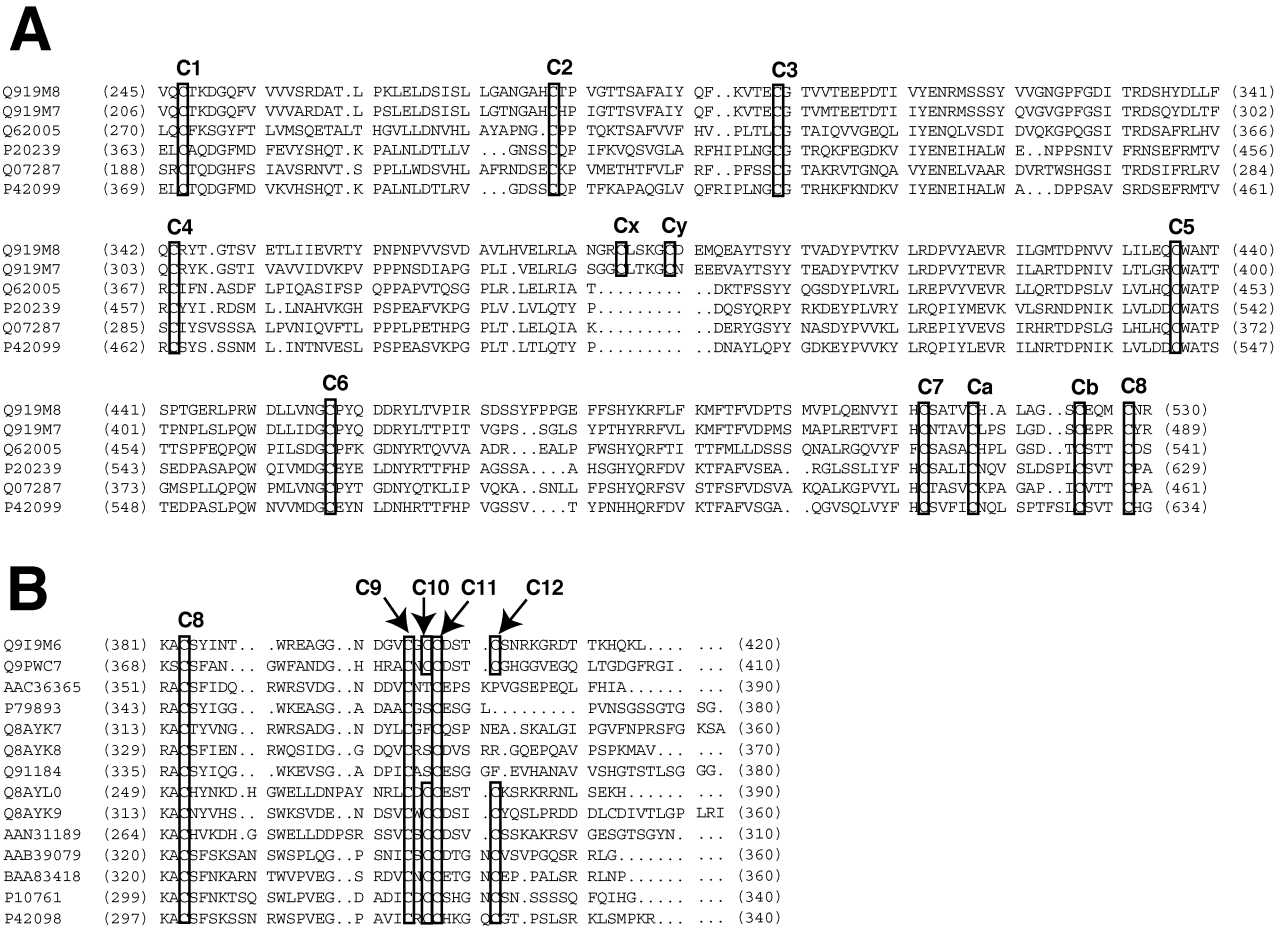

FIGURE 2: (A) Alignment of ZP1- and ZP2-like proteins of *O. mykiss* (VEα, Q919M8; VE*β*, Q919M7), *Mus musculus* (ZP1, Q62005; ZP2, P20239), and *Sus scrofa* (ZPB, Q07287; ZPA, P42099). The C residues conserved only within the ZP domain of all three ZP/VE proteins are designated C1 to C8. The C residues conserved only within the ZP domain of ZP1- and ZP2-like proteins are designated Ca and Cb. The C residues conserved only in fish ZP1-like proteins are designated Cx and Cy. (B) Alignment of the C-terminal region of ZP3-like proteins from different organisms showing that C10 and C12 are not conserved. The last C residue of the ZP domain (C8) is also shown. Proteins are *O. mykiss* VE*γ* (Q9I9M6), *Danio rerio* ZP3 (Q9PWC7), chorion proteic component (AAC36365), *Sparus aurata* chorion protein (P79893), *Oryzias latipes* ZPC5 (Q8AYK7), ZPC4 (Q8AYK8), choriogenin L (Q91184), ZPC1 (Q8AYL0), ZPC3 (Q8AYK9), ZPC2 (AAN31189), *Xenopus laevis* ZP3 (AAB39079), *Gallus gallus* ZP3 (BAA83418), *M. musculus* ZP3 (P10761), and *S. scrofa* ZPC (P42098).

detected by MALDI-TOF-MS under NR conditions must disappear under R conditions. (3) At least one *m*/*z* peak that corresponded to one of the two C-containing peptides participating in a disulfide must be detected under R conditions. (Note: Some peptides less than M_r ~800 Da were below the detection limits of the mass spectrometer.) (4) Disulfides detected by MALDI-TOF-MS in linear mode must be confirmed by using a different method (MALDI-TOF-MS in reflective mode, ESI-Q-TOF-MS, or chemical digestion and SDS−PAGE).

Disulfides were determined in the following manner. An intermolecular disulfide represents the sum of the individual unprotonated peptides minus 2 Da (C residue oxidation) plus 1 proton (peaks represent the protonated form of the peptides/disulfides in MALDI-TOF-MS). Intramolecular disulfides gain a net mass of 2 Da (1 disulfide) or 4 Da (2 disulfides) under R conditions. Here, a MALDI-TOF-MS peak is referred to as a protonated peptide/disulfide (*m*/*z*). An intermolecular disulfide detected by ESI-Q-TOF-MS represents the sum of the individual unprotonated peptides minus 2 Da (C residue oxidation), whereas a peptide with intramolecular disulfide loses a net mass of 2 Da (1 disulfide) or 4 Da (2 disulfides) after oxidation. The charged disulfides/peptides were calculated as [M + *n*H]/*n*, where M is the mass of unprotonated disulfide/peptide, H is a proton, and *n* is the number of protons. Also, the starting amino acid in every protein analyzed was M1.

*Cysteine Residues of VE*α*, VE*β*, and VE*γ. VEα and VE*β* each contain 18 C residues. The first 6 C residues are present in a conserved trefoil domain and, by analogy with other trefoil domain-containing proteins, are disulfide-linked C1−C5, C2−C4, and C3−C6, as reported previously (*18*). The next 12 C residues are located within the ZP domain of both VEα and VE*β*, are highly conserved (Figure 2A), and could be involved in 6 disulfides.

The ZP domain contains 8 C residues that are conserved in all ZP proteins, and these C residues are possibly involved in 4 disulfides. The C residues of VE proteins were designated C1 to C8 after the 8 conserved C residues of the ZP domain. The ZP domain of ZP1- or ZP2-like proteins contains, in addition to the 8 conserved C residues found in ZP3-like proteins, two extra C residues (conserved from fish

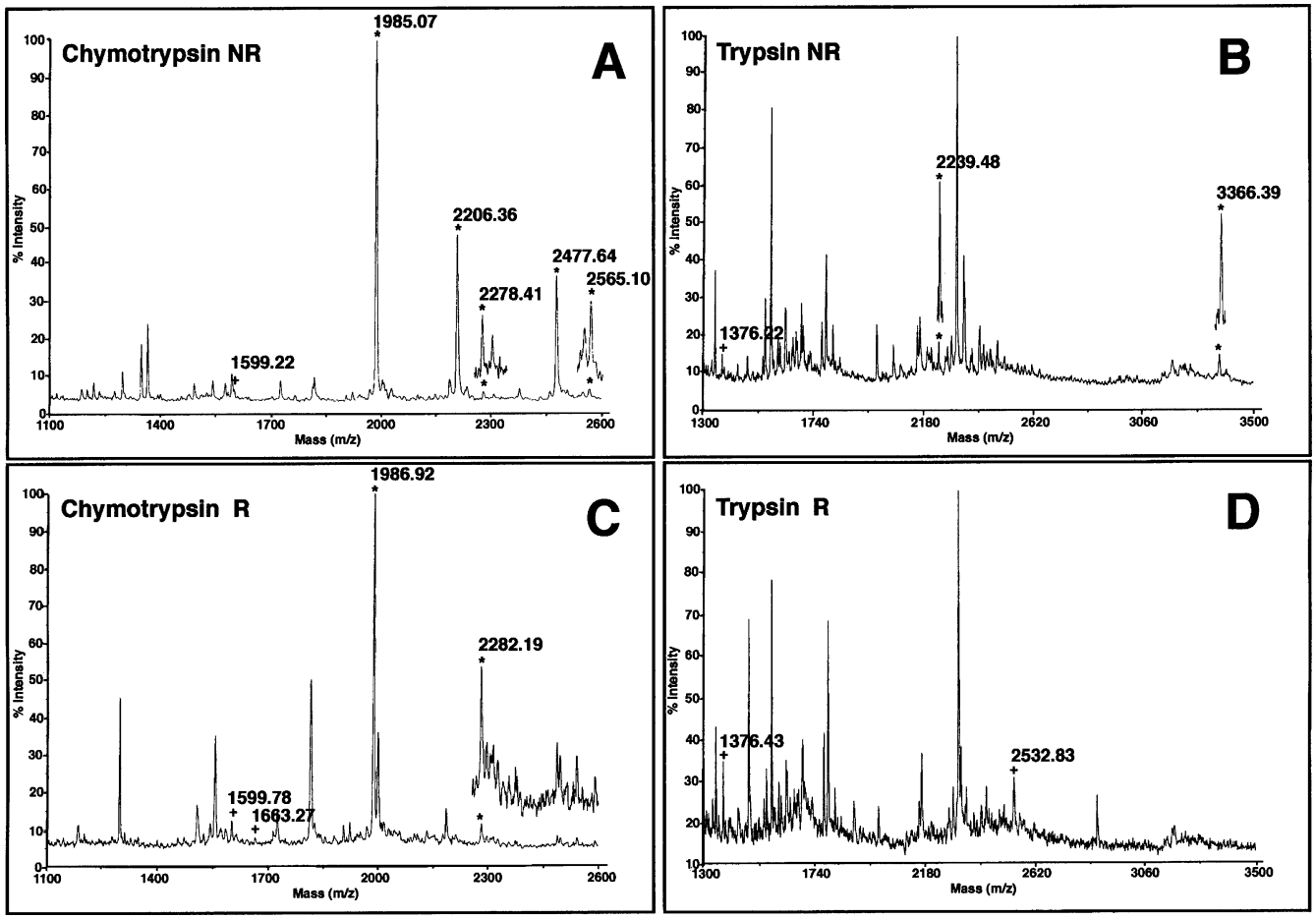

FIGURE 3: Assignment of the disulfide linkages of the VEβ by MALDI-TOF-MS (linear mode). The protein was digested with chymotrypsin (A, C) and trypsin (B, D) and measured under NR (A, B) and R (C, D) conditions. The disulfides are marked with an asterisk, and the C residue-containing peptides that resulted upon reduction are marked with a plus. The intramolecular disulfides (peaks of *m*/*z* 1985.07 and *m*/*z* 2278.41) gained 2 or 4 Da upon reduction. The peaks represent the average *m*/*z*. See text for details.

to mammals) designated Ca and Cb. Within the ZP domain of ZP1-like proteins from fish, there are two other C residues that have not been observed in mammals; these have been designated Cx and Cy. VEγ contains, in addition to the 8 C residues located within the ZP domain, 4 extra C residues (downstream of C8) designated C9 to C12. An alignment of these C residues, as well as the nomenclature used to assign C residues in rainbow trout VE proteins, is depicted in Figure 2.

Assignment of Disulfides to VEβ. The disulfides of VEβ were assigned by using enzymatic digests of the protein and MALDI-TOF-MS in linear mode under NR and R conditions. A series of peaks detected in MALDI-TOF-MS spectra under NR conditions that either disappeared or shifted toward a higher *m*/*z* under R conditions were considered the best candidates for disulfides. A peak of *m*/*z* 1985.07 found in chymotrypsin digests under NR conditions shifted toward a higher M_r (*m*/*z* 1986.92) under R conditions (Figure 3A,C). This corresponded to peptide ^{338}RLGSGGCLT-KGCNEE-EVAY^{356} (*m*/*z* calculated 1985.21, NR conditions; 1987.21, R conditions) that contains C344 (Cx) and C349 (Cy) disulfide-linked. This linkage was confirmed by analysis of trypsin digests under NR conditions (Figure 3B). A peak of *m*/*z* 3366.39 (calculated *m*/*z* 3366.65) disappeared under R conditions and corresponded to a disulfide linkage between peptides ^{339}LGSGGCLTK^{347} and ^{348}GCNEEEVAYTSY-YTEADYPVTK^{369} that contain C344 (Cx) and C349 (Cy), respectively. Under R conditions, a peak appeared that corresponded to a peptide containing C349 (Cy) with an *m*/*z* of 2532.83 (calculated *m*/*z* 2533.68) (Figure 3D). The C344 (Cx)−C349 (Cy) linkage was also confirmed using GluC (V8) and AspN digests (Table 2). Finally, additional evidence for this disulfide was provided by using ESI-Q-TOF-MS. Here, the peak of *m*/*z* 1985 observed by MALDI-TOF-MS was observed in a double charged state of *m*/*z* 992.55 (calculated *m*/*z* 992.45) that corresponded to the same peptide with the C residues disulfide-linked (Figure 4A). The CID produced a fragmentation spectrum in which b_{18}, b_{17}, and b_{16} ions were observed (Figure 4B). Also observed were a_{18} and a_{17} ions and some dehydrated ions (MH^+-H_2O, $y_{18}-H_2O$, $y_{16}-H_2O$, $b_{16}-H_2O$). These results suggest that C344 and C349 of VEβ are disulfide-linked (Cx and Cy).

A peak of *m*/*z* 2206.36 (calculated *m*/*z* 2205.43) detected in chymotrypsin digests under NR conditions disappeared under R conditions (Figure 3A, Table 2) and apparently corresponded to a disulfide between C396 (C5) and C417 (C6). Unfortunately, we were unable to detect a C residue-containing peptide under R conditions and, consequently, could not use this information.

A peak of *m*/*z* 2278.41 found in chymotrypsin digests (Figure 3A,C) gained 4 Da under R conditions (*m*/*z* 2282.19), suggesting the presence of two disulfide linkages. This peak corresponded to peptide ^{468}IHCNTAVCLPSLGDSCEP-RCY^{488} having a calculated *m*/*z* of 2278.62 (NR conditions)

Table 1: Disulfide Bridge Assignment in Fish VEα by MALDI-TOF-MS in Linear Mode[a]

Bridge	Observed/Calc. m/z	Sequence	Observed/Calc. m/z (P1 and P2)	Enzyme
C281-C300 C2-C3	3472.54/3472.84	275GANGAHCTPVGTTSAF290 296KVTECGTVVTEEPDTIVY313	ND/1491.62 1984.94/1984.22	Chymotrypsin
C281-C300 C2-C3	4114.23/4114.60	269DSISLLGANGAHCTPVGTTS AFAIYQFKVTECGTVVTEEP308	*	AspN
C384-C389 Cx-Cy	2983.63/2983.33	384CLSK387 388GCDEMQEAYTSYYTVADYPVTK409	ND/450.57 2536.47/2535.75	Trypsin
C384-C389 Cx-Cy	2153.99/2153.54 2156.51/2155.54	370DAVLHVELRLANGRCLSKGC389	*	AspN
C512-C517 C7-Ca	2974.51/2975.43 2976.56/2977.43	494TFVDPTSMVPLQENVYIHCS ATVCHAL520	*	Chymotrypsin

[a] The intramolecular bridges are marked with an asterisk. Peaks represent average *m*/*z*. P1 and P2 = peptides involved in disulfide bridges. Calc. = calculated. ND = nondetermined.

Table 2: Disulfide Bridge Assignment in Fish VE*β* by MALDI-TOF-MS in Linear Mode[a]

Bridge	Observed/Calc. m/z	Sequence	Observed/Calc. m/z (P1 and P2)	Enzyme
C208-C304 C1-C4	2239.48/2239.49	203SVTVQCTK210 295DSQYDLTFQCR305	ND/866.02 1376.43/1376.40	Trypsin
C208-C304 C1-C4	2477.64/2476.79	200YGKSVTVQCTKDGQF214 301TFQCRY306	1663.27/1661.87 ND/817.93	Chymotrypsin
C242-C261 C2-C3	2565.10/2564.91	236GTNGAHCHPIGTTSVF251 257KVTECGTVM265	1599.78/1599.76 968.96/968.17	Chymotrypsin
C242-C261 C2-C3	3630.5/3631.085	236GTNGAHCHPIGTTSVF251 257KVTECGTVMTEETDTIIY274	1599.78/1599.76 ND/2034.30	Chymotrypsin
C242-C261 C2-C3	2855.89/2856.20	236GTNGAHCHPIGTTSVF251 255QFKVTECGTV**M**265	1599.78/1599.76 ND/1259.48	Chymotrypsin
C344-C349 Cx-Cy	3366.39/3366.65	339LGSGGCLTK347 348GCNEEEVAYTSYYTEADYPVTK369	ND/835.99 2532.83/2533.68	Trypsin
C344-C349 Cx-Cy	1985.07/1985.21 1986.92/1987.21	338RLGSGGCLTKGCNEEEVAY356	*	Chymotrypsin
C344-C349 Cx-Cy	3919.75/3919.38 3922.59/3921.38	327DIAPGPLIVELRLGSGGCLT KGCNEEEVAYTSYYTEA363	*	AspN
C344-C349 Cx-Cy	1764.87/1764.98 1766.86/1766.98	337LRLGSGGCLTKGCNEEE353	*	GluC
C396-C417 C5-C6	2206.36/2205.43	394GRCW397 411DLLIDGCPYQDDRY424	ND/521.61 ND/1686.84	Chymotrypsin
C470-C475-C483-C487 C7-a-b-8	2278.41/2278.62 2282.19/2282.62	468IHCNTAVCLPSLGDSCEPRCY488	*	Chymotrypsin
C470-C475-C483-C487 C7-a-b-8	3942.24/3942.55 3946.31/3946.55	455DP**M**S**M**APLRETVFIHCNTAV CLPSLGDSCEPRCYR489	*	AspN
C483-C487 Cb-C8	1126.71/1127.25 1129.01/1129.25	481DSCEPRCYR489	*	AspN

[a] The intramolecular bridges are marked with an asterisk. The underlined M represents oxidated methionine. Peaks represent average *m*/*z*. P1 and P2 = peptides involved in disulfide bridges. Calc. = calculated. ND = nondetermined.

and *m*/*z* 2282.62 (R conditions) and contained C470 (C7), C475 (Ca), C483 (Cb), and C487 (C8) in disulfide linkages. These disulfides were also confirmed by ESI-Q-TOF-MS of chymotrypsin digests, where a triple charged (3+) peak of *m*/*z* 759.78 (calculated *m*/*z* 759.66) corresponded to the previously detected peptide ^{468}IHCNTAVCLPSLGDSCEP-RCY^{488}, with the C residues disulfide-linked (Figure 4A). However, no collision-induced dissociation spectra for these

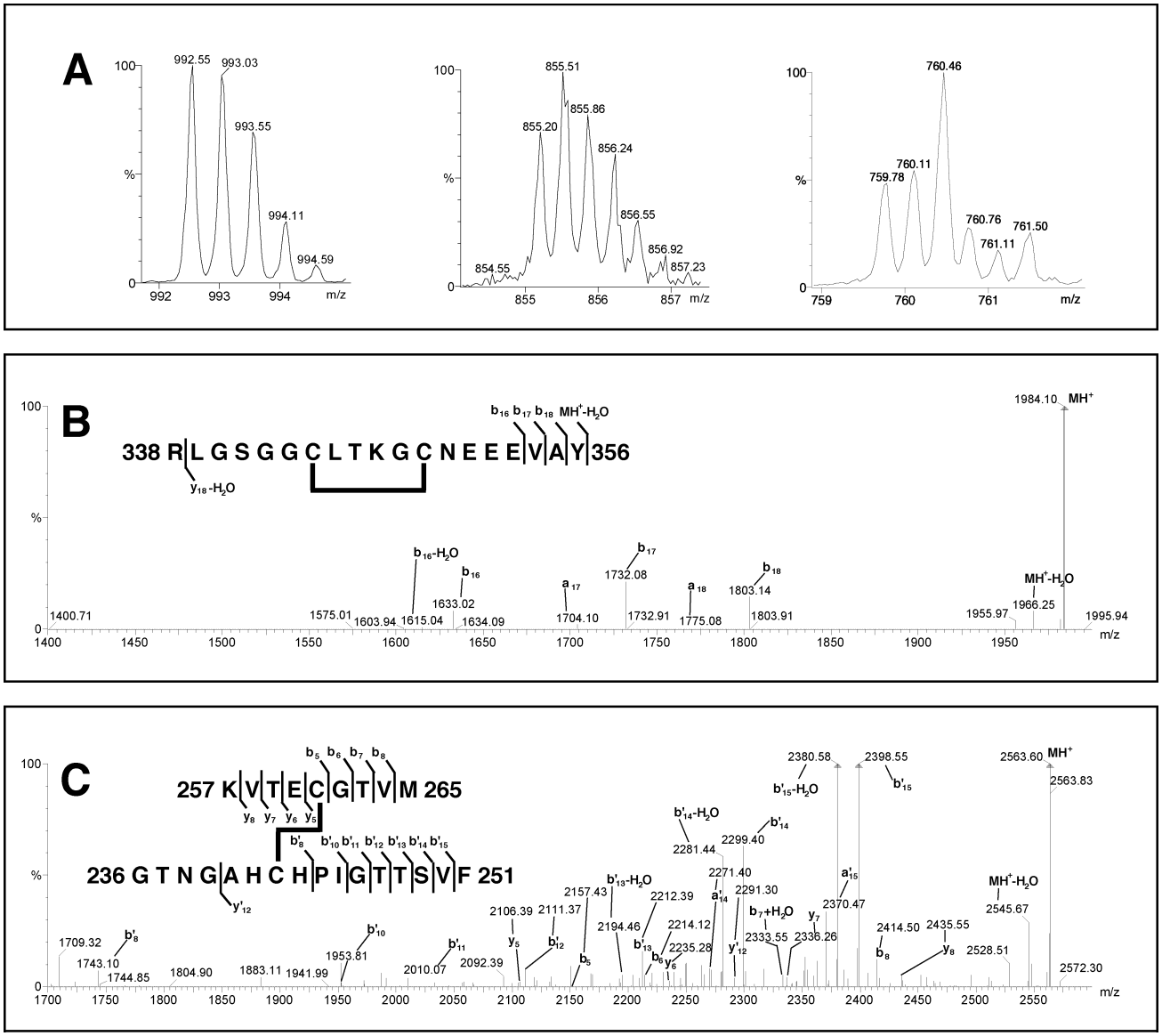

FIGURE 4: Disulfide linkage determination within VEβ by ESI-Q-TOF-MS. The peaks represent the monoisotopic *m*/*z*. (A) The *m*/*z* 992.55 ion represents the double charged state of the intramolecular disulfide-linked peptide shown in panel B. The *m*/*z* 855.20 ion represents the triple charged state of the intermolecularly linked peptides shown in panel C. The *m*/*z* 759.78 ion represents the triple charged state of the intramolecular disulfide-linked peptide ^{468}IHCNTAVCLPSLGDSCEPRCY^{488}. (B) CID spectrum of the 2+ charged ion *m*/*z* 992.55. The precursor ion of *m*/*z* 1984.10 and the resulting fragments correspond to peptide ^{338}RLGSGGCLTKGCNEEEVAY^{356} minus 2 Da, due to the intramolecular linkage. (C) CID spectrum of the 3+ charged ion *m*/*z* 855.20. The precursor ion of *m*/*z* 2563.60 and the fragment ions correspond to the peptides ^{257}KVTECGTVM^{265} and ^{236}GTNGAHCHPIGTTSVF^{251} disulfide-linked.

peaks were observed. Further evidence that these four C residues are disulfide-linked came from analysis of AspN digests (Table 2) under NR and R conditions. A peak of *m*/*z* 3942.24 (calculated *m*/*z* 3942.55) shifted under R conditions to *m*/*z* 3946.31 (calculated *m*/*z* 3946.55) that corresponded to peptide ^{455}DPMSMAPLRETVFIHCNTAVCLPSLGDSCEPRCYR^{489} that contains C470, C475, C483, and C487 disulfide-linked. Analysis of AspN digests (Table 2) revealed a peak of *m*/*z* 1126.71 (calculated *m*/*z* 1127.25) that gained 2 Da under R conditions (*m*/*z* 1129.01; calculated *m*/*z* 1129.25). This peak corresponded to peptide ^{481}DSCEPRCYR^{489} that contains C483 and C487 disulfide-linked (Cb and C8). In view of the latter result, it can be concluded that C470 and C475 are also disulfide-linked (C7 and Ca).

In a chymotrypsin digest of VEβ (Figure 3A,C), a peak of *m*/*z* 2477.64 (calculated *m*/*z* 2476.79) corresponded to a disulfide linkage between peptides ^{200}YGKSVTVQCTKDGQF^{214} and ^{301}TFQCRY^{306} containing C208 (C1) and C304 (C4). This peak disappeared under R conditions and gave rise to a peak of *m*/*z* 1663.27 (calculated *m*/*z* 1661.80) that contained C208 (C1). The presence of this disulfide was confirmed by using a trypsin digest of VEβ (Figure 3B,D). A peak of *m*/*z* 2239.48 (calculated *m*/*z* 2239.49) corresponded to a disulfide linkage between peptides ^{203}SVTVQCTK^{210} and ^{295}DSQYDLTFQCR^{305} containing C208 (C1) and C304 (C4), respectively, that disappeared under R conditions. The peak that corresponded to the peptide containing C304 (C4) was identified under R conditions (observed *m*/*z* 1376.43; calculated *m*/*z* 1376.40). These results suggest that C208 and C304 of VEβ are disulfide-linked (C1−C4).

A peak of *m*/*z* 2565.10 (calculated *m*/*z* 2564.91) detected in chymotrypsin digests under NR conditions (Figure 3A,C) corresponded to a disulfide linkage between peptides ^{236}GTNGAHCHPIGTTSVF^{251} and ^{257}KVTECGTVM^{265} that

contain C242 (C2) and C261 (C3), respectively. This peak disappeared under R conditions, and the peptide containing C242 (C2) appeared with a *m*/*z* of 1599.78 (calculated *m*/*z* 1599.70). In further experiments, peptide [236]GTNGAHCHPIGTTSVF[251] was involved in other C242 (C2) and C261 (C3) disulfides [*m*/*z* 3630.5 (calculated *m*/*z* 3631.08); *m*/*z* 2855.89 (calculated *m*/*z* 2856.20); Table 2], confirming that C242 (C2) and C261 (C3) are disulfide-linked. Additional proof was provided by ESI-Q-TOF-MS of chymotrypsin digests. Here, the peak of *m*/*z* 2565.10 (calculated *m*/*z* 2564.91) observed by MALDI-TOF-MS (Figure 3A) was also observed in a triple charged state of *m*/*z* 855.20 (calculated *m*/*z* 855.06) (Figure 4A). The fragmentation obtained by collision-induced dissociation (CID) produced a series of b, y, b′, and y′ ions (Figure 4C) that confirmed the MALDI-TOF-MS experiments. These results suggest that C242 and C261 of VEβ are disulfide-linked (C2−C3).

A summary of the assignment of disulfides for VEβ is presented in Figure 10B and Table 2.

*Assignment of Disulfides to VE*α. Alignments between VEα and VEβ (Figure 2) suggest that the C residues of these proteins are absolutely conserved. Therefore, it was anticipated that the disulfides in VEα would be identical to those in VEβ.

A peak of *m*/*z* 2983.63 (calculated *m*/*z* 2983.33) in trypsin digests of VEα under NR conditions corresponded to a disulfide linkage between peptides [384]CLSK[387] (C-384; Cx) and [388]GCDEMQEAYTSYYTVADYPVTK[409] (C-389; Cy) (Table 1). This peak disappeared under R conditions, yielding a peak that corresponds to a peptide containing C389 with a *m*/*z* of 2536.47 (calculated *m*/*z* 2535.75). Analysis of AspN digests of VEα under NR conditions gave a peak of *m*/*z* 2153.99 (calculated *m*/*z* 2153.54) that shifted 2 Da to *m*/*z* 2156.51 (calculated *m*/*z* 2155.54) under R conditions (Table 1). This peak corresponded to peptide [370]DAVLHVELRLANGRCLSKGC[389] that contains C384 and C389 disulfide-linked (Cx-Cy).

A peak of *m*/*z* 3472.54 (calculated *m*/*z* 3472.84) in chymotrypsin digests of VEα under NR conditions (Table 1) corresponded to a disulfide linkage between peptides [275]GANGAHCTPVGTTSAF[290] (C281; C2) and [296]KVTECGTVVTEEPDTIVY[313] (C300; C3). This peak disappeared under R conditions, yielding a peptide that contained C300 (C3) with a *m*/*z* of 1984.94 (calculated *m*/*z* 1984.22) (Table 1). These results suggest that C281 and C300 are disulfide-linked (C2−C3). In this context, it was noted that a peak corresponding to a C281 (C2) and C300 (C3) linkage was observed in AspN digests (Table 1), but no shift in mass occurred under R conditions.

A peak of *m*/*z* 2974.51 (calculated *m*/*z* 2975.43) in chymotrypsin digests of VEα under NR conditions shifted to *m*/*z* 2976.56 (calculated *m*/*z* 2977.43) under R conditions (Table 1). This peak corresponded to peptide [494]TFVDPTSMVPLQENVYIHCSATVCHAL[520] that contains C512 (C7) and C517 (Ca) disulfide-linked.

A summary of the assignment of disulfides for VEα is presented in Figure 10A and Table 1.

Assignment of Disulfides to VEγ. VEγ contains 12 C residues, 8 within the ZP domain (C1−C8), that could potentially form 6 disulfides. Two different methods were used to assign disulfides in VEγ: chemical (FA and BNPS-skatole) or enzymatic (trypsin and chymotrypsin) fragmentation of the protein.

There are two theoretical cleavage sites for FA in VEγ: between D82 and P83 and between D337 and P338 (Figure 5C). The first site is located before the first C residue (C108; C1) and the second between C299 (C6) and C365 (C7). Cleavage between D82 and P83 under NR conditions should produce the two major fragments (M_rs ∼6 and 38 kDa) seen in Figure 5A. Cleavage between D337 and P338 under NR conditions should produce an additional polypeptide with M_r >30 kDa, plus a fragment representing the C-terminus of the polypeptide (M_r ∼8 kDa), unless there is a disulfide between one of the first 6 and one of the last 6 C residues. Under R conditions, two additional peptides should appear if there is a disulfide. As seen in Figure 5A, in addition to the three major bands observed under NR conditions, two additional bands (M_rs ∼30 and 8 kDa) were observed under R conditions. These results suggest that there is a disulfide linkage between one of the first 6 and one of the last 6 C residues.

There are several cleavage sites (C-terminus of W residues) for BNPS-skatole in VEγ: before the first C residue (W25 and W77), between the C4 (C200) and C5 (C277) residues (W217 and W241), and between the C8 (C383) and C9 (C399) residues (W389). As seen in Figure 5A, BNPS-skatole digests of VEγ yielded six main bands following electrophoresis under both NR and R conditions. These results suggest that there are no disulfides between the first 4 C residues (C1, C2, C3, and C4), the next 4 C residues (C5, C6, C7, and C8), and the last 4 C residues (C9, C10, C11, and C12). This interpretation was confirmed by subjecting BNPS-skatole digests to SDS−PAGE in two dimensions: first under NR conditions on SDS−PAGE (first dimension) and then under R conditions (second dimension) (Figure 5B). No significant shift in the M_rs of the major bands was observed after the transition from NR to R conditions.

Overall, the results of these experiments with FA and BNPS-skatole digests suggest that there are disulfide linkages between the first 4 C residues and between the last 4 C residues. Furthermore, C5 is disulfide-linked to C7 or C8, and C6 is disulfide-linked to C7 or C8. There are no disulfide linkages between C7 and C8 or between C5 and C6.

To pursue further the positions of disulfides in VEγ, MALDI-TOF-MS (linear mode) analyses were performed on enzymatic digests of the protein under NR and R conditions. In chymotrypsin digests under NR conditions, a peak of *m*/*z* 2337.76 (calculated *m*/*z* 2338.70) was found that corresponded to a disulfide linkage between peptides [136]GDCPM[140] (C138; C2) and [157]QSCGSQLR*M*TTNSLIY[172] (C159; C3). Under R conditions, this peak disappeared, and a peak of *m*/*z* 1817.93 (calculated *m*/*z* 1819.05) appeared that corresponded to the peptide containing C159 (C3) (Figure 6A,C). In the same spectrum (Figure 6A,C), under NR conditions, a peak of *m*/*z* 3700.50 (calculated *m*/*z* 3701.20) was observed that also disappeared under R conditions. This peak corresponded to peptide [157]QSCGSQLRMTTNSLIY[172] that contains C159 (C3) seen under R conditions and that is disulfide-linked to peptide [136]GDCPMTGFDNINQVLIF[152], which contains C138 (C2). These results suggest that C138 (C2) and C159 (C3) are disulfide-linked.

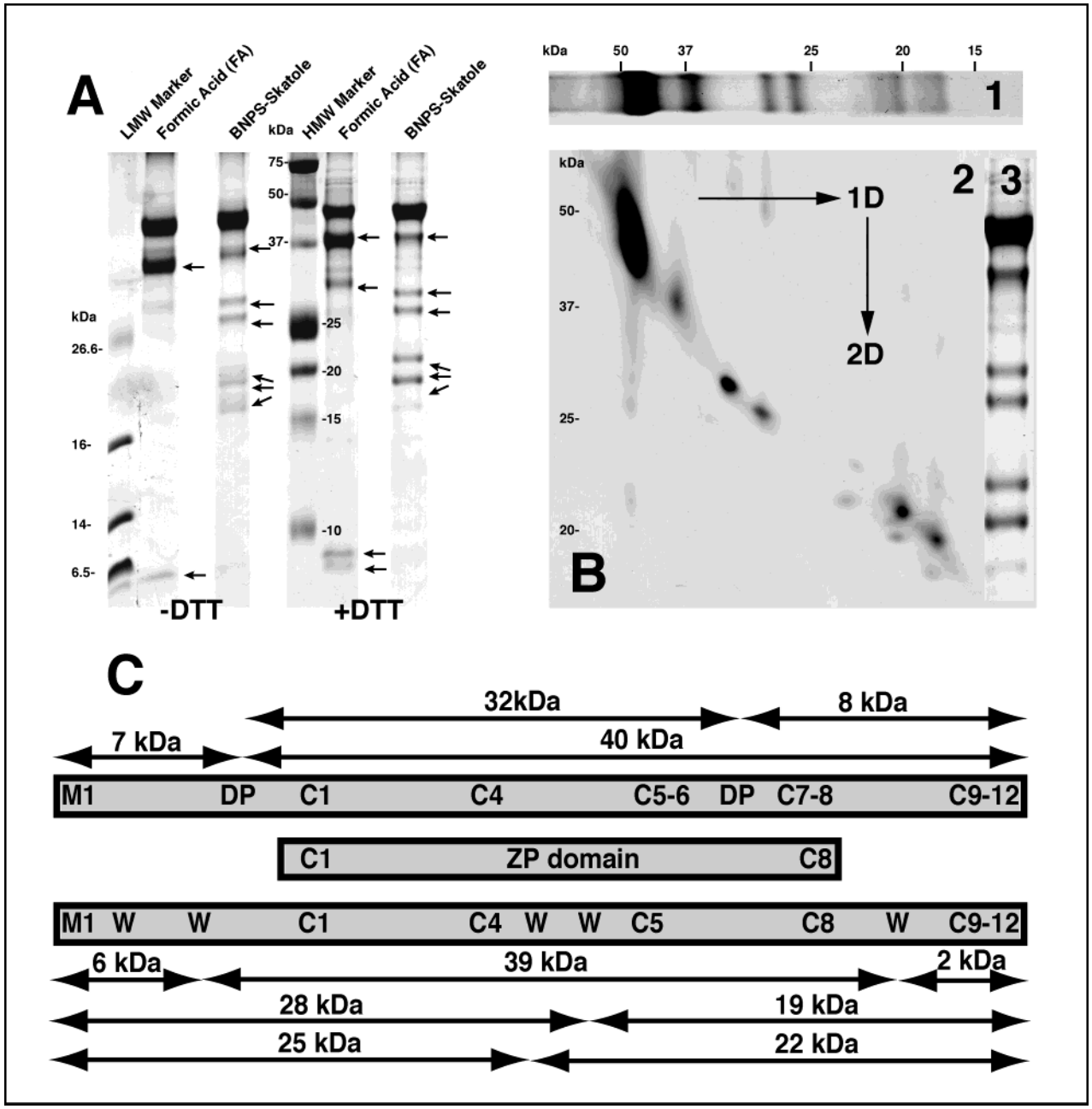

FIGURE 5: Analysis of disulfide linkages in VEγ digested by chemical methods. (A) 16% Tris−Tricine−PAGE of VEγ digested with either FA or BNPS-skatole. Digestion products were separated under NR (−DTT) and R (+DTT) conditions. Low M_r (LMW) markers are on the left, and high M_r (HMW) markers are on the right. The main digestion products are indicated by arrows. The gel is Coomassie-stained. (B) Two-dimensional (2D) PAGE of VEγ digested with BNPS-skatole. The digestion products were separated on 10% SDS−PAGE under NR conditions (1). One gel lane containing the separated products was excised, reduced, and oriented horizontally for the second dimension on 10% SDS−PAGE under R conditions (2). A BNPS-skatole digest under R conditions is shown for comparison (3). The first dimension (1D) was transiently zinc-stained, and the second dimension (2D) was silver-stained. Protein markers are on the top for 1D gels and on the left for 2D gels. Note that gels 1, 2, and 3 resulted from different experiments, due to the transient zinc staining on 1D, and the arrangement of the protein spots on the 1D (equivalent to the protein lanes of the 2D) are approximated. (C) Schematic representation of theoretical digestions of VEγ by FA and BNPS. The starting amino acid in the protein is marked M1. The C residues are designated C1−C12; C residues C1−C8 are conserved within the ZP domain. Top: The FA cleavage site is DP, one of them situated before C1 and another between C6 and C7. The main theoretical cleavage products and their M_rs are indicated. Bottom: The main BNPS cleavage sites, noted as W, are situated before C1, between C4 and C5, and between C8 and C9. The main theoretical cleavage products and their M_rs are shown.

A peak of *m*/*z* 3319.68 (calculated *m*/*z* 3319.79) present in trypsin digests under NR conditions disappeared under R conditions and corresponded to a disulfide linkage between C108 (C1) and C200 (C4) (Figure 6B). Under R conditions, one of the peptides, ^{106}AECRENMVHVEAK^{118} (Figure 6D), was observed as a peak of *m*/*z* 1516.81 (calculated *m*/*z* 1516.72). Similarly, a peak of *m*/*z* 3879.27 (calculated *m*/*z* 3879.35) (Table 3) present in trypsin digests under NR conditions disappeared under R conditions and corresponded to a C1−C4 disulfide. Under R conditions this peak disappeared, and one of the peptides, ^{191}TNDAMINIEC-HYPR^{204}, was observed as a peak of *m*/*z* 1693.82 (calculated *m*/*z* 1693.88) (Table 3). The peak corresponding to this particular disulfide was also observed in ESI-Q-TOF-MS of

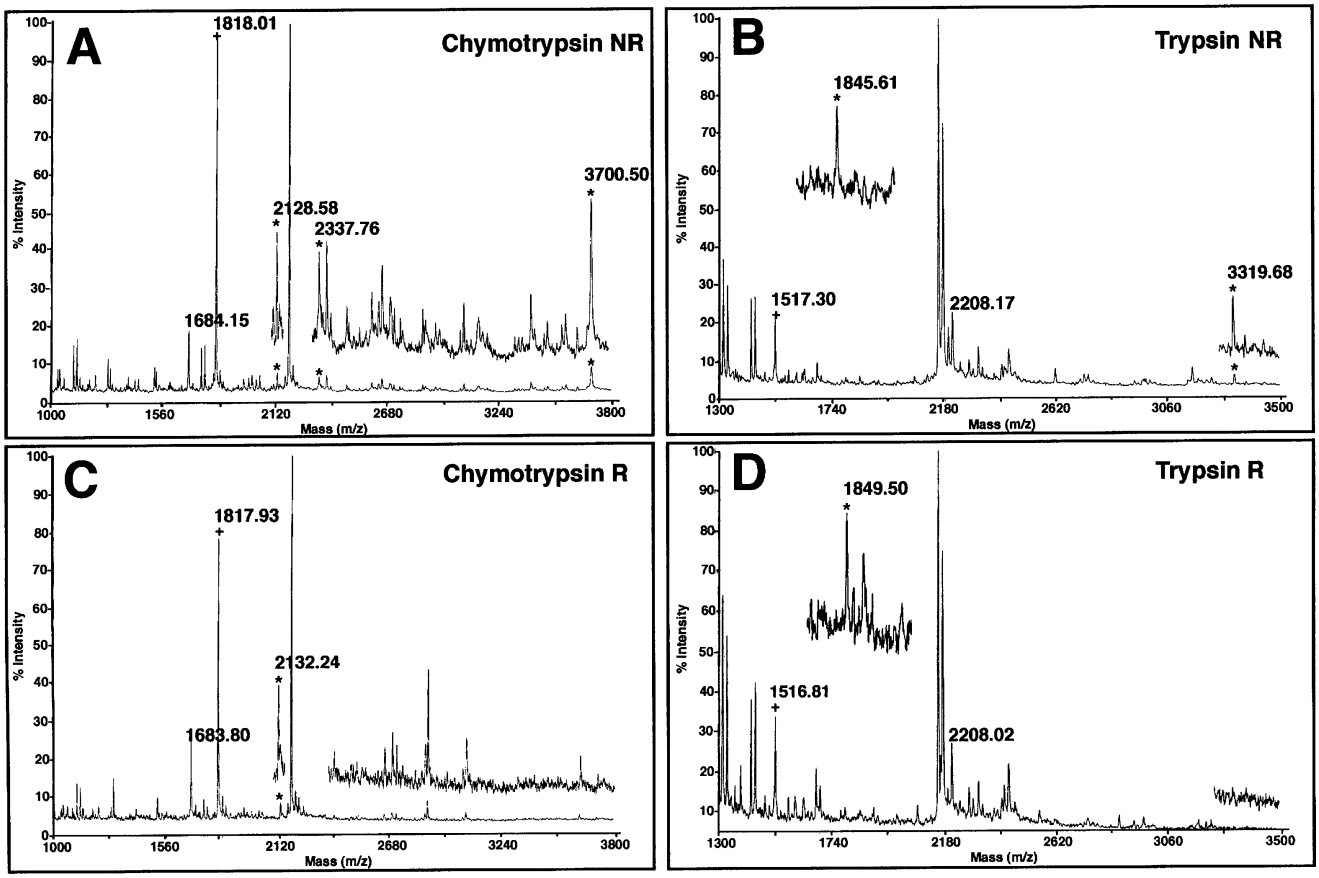

FIGURE 6: Assignment of the disulfide linkages of VEγ by MALDI-TOF-MS (linear mode). The protein was digested with either chymotrypsin (A, C) or trypsin (B, D) and measured under NR (A, B) and R (C, D) conditions. Disulfides are marked with an asterisk, and the C residue-containing peptides that resulted upon reduction or as a laser-induced homolytic cleavage of the disulfides are marked with a plus. The intramolecular disulfides (peaks *m*/*z* 2128.58 and *m*/*z* 1845.61) gained 4 Da upon reduction. The peaks of *m*/*z* 1845.61 and *m*/*z* 1849.50 (Figure 6B,D, inset) are from a different experiment. The chymotrypsin digest peaks of *m*/*z* 1684.15 (Figure 6A) and *m*/*z* 1683.80 (Figure 6C) and the trypsin digest peaks of *m*/*z* 2208.17 (Figure 6B) and *m*/*z* 2208.02 (Figure 6D) represent the N-terminus of VEγ. The peaks represent the average *m*/*z*. See text for details.

trypsin digests, where the disulfide-linked peptides ^{191}TNDAMINIECHYPR^{204} and ^{89}IIGSSEARTPVAANSVRAECR^{109} were observed in a 5+ (observed *m*/*z* 776.28, calculated *m*/*z* 776.17; Figure 7A) and 4+ charged state (observed *m*/*z* 970.11, calculated *m*/*z* 969.96; not shown). However, no fragmentation for these peaks was observed. This suggests that C108 (C1) and C200 (C4) are disulfide-linked (Table 3).

In other experiments with chymotrypsin digests under NR conditions, peaks of *m*/*z* 3465.16 (calculated *m*/*z* 3464.87) and *m*/*z* 3481.57 (calculated *m*/*z* 3480.87) were observed (Table 3). Both peaks corresponded to peptide ^{294}IENHGCLIDAKMTGSHSQF^{312} containing C299 (C6) disulfide-linked to peptide ^{379}EYKACSYINTW^{389} (calculated *m*/*z* 1378.54) containing C383 (C8). The peak of *m*/*z* 3481.57 represented the M-oxidized product of the peak at *m*/*z* 3465.16 (Table 3). Under R conditions both peaks disappeared, and a peptide with a *m*/*z* of 1378.72 containing C383 (C8) appeared (Table 3). The C6−C8 linkage was confirmed by trypsin digests in which a peak of *m*/*z* 4532.41 (calculated *m*/*z* 4531.13) disappeared under R conditions and corresponded to peptide ^{382}ACSYINTWR^{390} containing C383 (C8) disulfide linked to C299 (C6)-containing peptide ^{291}YAFIENHGCLIDAKMTGSHSQFMPRSADYK^{320}. Under R conditions, a peak corresponding to a peptide containing C383 (C8) appeared with a *m*/*z* of 1114.16 (calculated *m*/*z* 1114.26) (Table 3). Together with results obtained using chemical methods, these results suggest that C299 (C6) and C383 (C8), as well as C277 (C5) and C365 (C7), are disulfide-linked.

A peak of *m*/*z* 2128.58 in chymotrypsin digests under NR conditions (Figure 6A,C) gained 4 Da under R conditions (*m*/*z* 2132.24), suggesting the presence of two disulfides. The peak corresponded to the peptide ^{390}REAGGNDGVCGGCCDSTCSNRK^{410} containing C399 (C9), C401 (C10), C402 (C11), and C406 (C12) (NR conditions, calculated *m*/*z* 2129.30; R conditions, calculated *m*/*z* 2133.30). The disulfides were confirmed in trypsin digests (Figure 6B,D) in which a peak of *m*/*z* 1845.61 (calculated *m*/*z* 1844.94) under NR conditions became a peak of *m*/*z* 1849.50 (calculated *m*/*z* 1848.94) under R conditions. This *m*/*z* 1845.61 disulfide detected by MALDI-TOF-MS was also confirmed by ESI-Q-TOF-MS of trypsin digests, where a double charged peak of *m*/*z* 922.48 (calculated *m*/*z* 922.32) (Figure 7B) and a triple charged peak of *m*/*z* 615.37 (calculated *m*/*z* 615.20; data not shown) corresponded to the peptide ^{391}EAGGNDGVCGGCCDSTCSNR^{409} with the C residues disulfide-linked. Therefore, the last 4 C residues (C9, C10, C11, and C12) are disulfide-linked.

Analyses of BNPS-skatole digests confirmed that there are disulfide linkages between C1−C4, C5−C8, and C9−C12. Similarly, analyses of FA digests confirmed that there are disulfide linkages between C5 and C7 and between C6 and C8. Since C10 and C12 of VEγ are not strictly conserved

Table 3: Disulfide Bridge Assignment in Fish VEγ by MALDI-TOF-MS in Linear Mode[a]

Bridge	Observed/Calc. m/z	Sequence	Observed/Calc. m/z (P1 and P2)	Enzyme
C108-C200 C1-C4	3319.68/3319.79	106AECRENMVHVEAK118 191TNDAMINIECHYPRK205	1516.81/1516.72 ND/1806.06	Trypsin
C108-C200 C1-C4	3879.27/3879.35	89IIGSSEARTPVAANSVRAECR109 191TNDAMINIECHYPR204	ND/2188.46 1693.82/1693.88	Trypsin
C138-C159 C2-C3	3179.50/3180.50	136GDCPMTGFDNINQVL150 157QSCGSQLRMTTNSL170	ND/1640.82 1542.69/1542.72	Chymotrypsin
C138-C159 C2-C3	2337.76/2338.70	136GDCPM 140 157QSCGSQLRMTTNSLIY172	ND/522.61 1817.93/1819.05	Chymotrypsin
C138-C159 C2-C3	2940.99/2940.33	134TLGDCPM140 144DNINQVLIFESPLQSCGSQL163	ND/736.87 2207.26/2206.47	Chymotrypsin
C138-C159 C2-C3	3700.50/3701.20	136GDCPMTGFDNINQVLIF152 157QSCGSQLRMTTNSLIY172	ND/1885.15 1817.93/1819.05	Chymotrypsin
C299-C383 C6-C8	3465.16/3464.87	294IENHGCLIDAKMTGSHSQF312 379EYKACSYINTW389	ND/2089.34 1378.72/1378.54	Chymotrypsin
C299-C383 C6-C8	3481.57/3480.87	294IENHGCLIDAKMTGSHSQF312 379EYKACSYINTW389	ND/2105.34 1378.72/1378.54	Chymotrypsin
C299-C383 C6-C8	4532.41/4531.13	291YAFIENHGCLIDAKMTGSHSQFMP RSADYK320 382ACSYINTWR390	ND/3419.87 1114.16/1114.26	Trypsin
C399-C401-C402-C406 C9-10-11-12	1845.61/1844.94 1849.50/1848.94	391EAGGNDGVCGCCDSTCSNR409	*	Trypsin
C399-C401-C402-C406 C9-10-11-12	2128.58/2129.30 2132.24/2133.30	390REAGGNDGVCGCCDSTCSNRK410	*	Chymotrypsin

[a] The intramolecular bridges are marked with an asterisk. The underlined M represents oxidated methionine. Peaks represent average *m*/*z*. P1 and P2 = peptides involved in disulfide bridges. Calc. = calculated. ND = nondetermined.

(Figure 2B), it can be concluded that C9 is disulfide-linked to C11 and C10 to C12.

A summary of the assignment of disulfides for VEγ is presented in Figure 10C and Table 3.

Determination of the N-Terminus of VE Proteins. VE proteins are secreted proteins and contain a signal sequence at the N-terminus that directs them to the endoplasmic reticulum and is removed from the mature proteins. Using the Signal P program (http://www.cbs.dtu.dk/services/SignalP/), cleavage sites for signal peptidase are predicted to be before Q23, Q21, and Q23 in VEα, VEβ, and VEγ, respectively.

MALDI-TOF-MS (linear mode) analyses of chymotrypsin digests of VEγ revealed that Q23 is the N-terminal amino acid of the protein; the amino acid is present as pyroglutamic acid (loss of 17 Da), a common form of N-terminal Q residues. Under NR and R conditions (Figure 6A,C) peaks of *m*/*z* 1684.15 and *m*/*z* 1683.80, respectively, were observed and corresponded to peptide ^{23}QNWPPFSKPVQQPF^{36} having a calculated *m*/*z* of 1683.91 (1699.93 − 17) and Q23 in a pyrrolidone form. The same N-terminus was found in spectra of trypsin digests (Figure 6B,D). Peaks of *m*/*z* 2208.17 and *m*/*z* 2208.02 (calculated *m*/*z* 2207.50) corresponded to peptide ^{23}QNWPPFSKPVQQPFRPNR^{40}, with Q23 as pyroglutamic acid.

This was confirmed by using MALDI-TOF-MS in reflective mode in which a peak represents the monoisotopic, protonated form of the peptide (as compared with MALDI-TOF-MS in linear mode in which a peak represents the average, protonated form of the peptide). Peaks of *m*/*z* 1682.71 (calculated *m*/*z* 1682.83) and *m*/*z* 2206.11 (calculated *m*/*z* 2206.15) of the chymotrypsin and trypsin digests, respectively, corresponded to the same peptides detected by MALDI-TOF-MS in linear mode (data not shown).

However, the Q23 N-terminus of VEγ was detected not only as pyroglutamic acid but also as unmodified glutamine. In ESI-Q-TOF-MS experiments with trypsin digests of VEγ, both N-termini were detected. Triple charged peaks of *m*/*z* 736.22 (calculated *m*/*z* 736.05) and *m*/*z* 741.55 (calculated *m*/*z* 741.71) corresponded to peptide ^{23}QNWPPFSKPVQQPFRPNR^{40}, with Q23 in pyroglutamic acid and glutamine form (Figure 7C). However, although no fragmentation spectra were observed, these results were also confirmed by MALDI-TOF-MS (linear mode) experiments (not shown).

The N-terminal amino acid of VEβ was detected by MALDI-TOF-MS (linear mode) analysis of pepsin digests. A peak of *m*/*z* 1569.59 corresponded to peptide ^{21}QIYLEKPGWPPIQ^{33} (calculated *m*/*z* 1569.84) in which Q21 was not modified. Due to a lack of reproducibility, this assignment remains tentative. We failed to detect an N-terminus for VEα due to the presence of a PQ-rich stretch at the very N-terminal region of the protein which hindered its enzymatic digestion.

Determination of the C-Terminus of VE Proteins. Nascent VE proteins contain a propeptide at the C-terminus that is removed by a proprotein convertase (a furin-like protease)

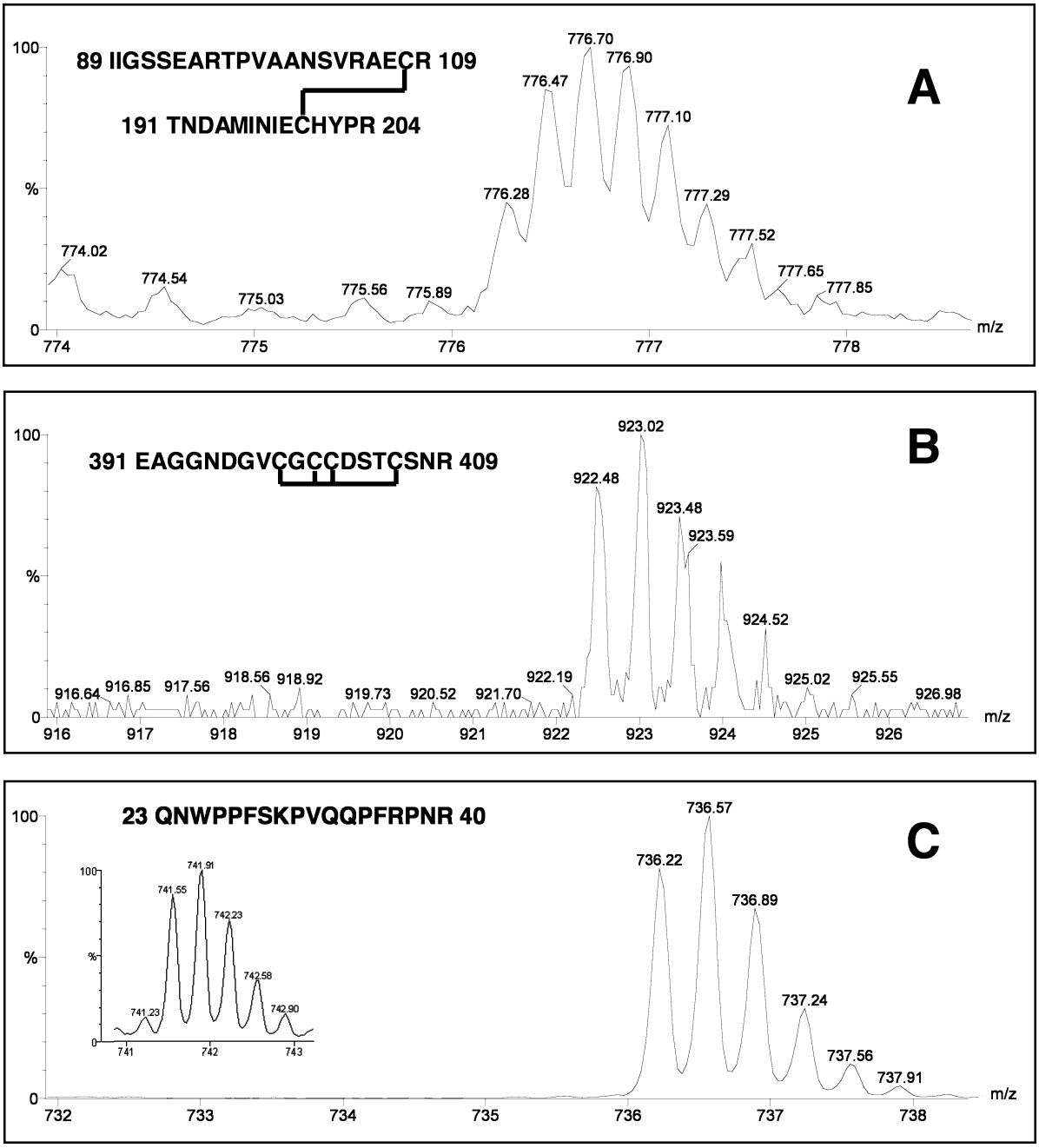

FIGURE 7: Determination of disulfide linkages and N-terminus of VEγ by ESI-Q-TOF-MS (trypsin digests). The peaks represent the monoisotopic *m*/*z*. (A) The intermolecular disulfide-linked peptides ^{89}IIGSSEARTPVAANSVRAECR^{109} and ^{191}TNDAMINIECHYPR^{204} were observed as *m*/*z* 776.28 in the 5+ charged state and as *m*/*z* 970.11 in the 4+ charged state (not shown). (B) An intramolecularly disulfide-linked peptide ^{391}EAGGNDGVCGGCCDSTCSNR^{409} was detected as a double charged peak of *m*/*z* 922.48 and a triple charged peak of *m*/*z* 615.37 (not shown). (C) The N-terminus of VEγ was observed as triple charged peaks of *m*/*z* 736.22 and *m*/*z* 741.55 and corresponded to peptide ^{23}QNWPPFSKPVQQPFRPNR^{40}, with Q23 as pyroglutamic acid (*m*/*z* 736.22) and as unmodified glutamine (*m*/*z* 741.55).

acting on the consensus sequence RKXR and is not present in the mature (assembled) proteins (*29*).

The C-terminus of VEγ was determined by MALDI-TOF-MS (linear mode) analysis of chymotrypsin digests since there is no cleavage site for this protease at RKXR. Peaks of *m*/*z* 2132.34 (calculated *m*/*z* 2133.30) and *m*/*z* 2128.58 (calculated *m*/*z* 2129.30) were observed under R and NR conditions, respectively, and corresponded to the peptide ^{390}REAGGNDGVCGCCDSTCSNRK^{410} (Figure 6A,C). K410, the C-terminal amino acid, lies two amino acids upstream of the furin-like cleavage site.

To determine the C-terminus of VEβ, MALDI-TOF-MS (linear mode) analysis of chymotrypsin and AspN digests was used, yielding two different results. For chymotrypsin digests, the peptide ^{468}IHCNTAVCLPSLLGDSCEPRCY^{488} was observed under both NR (*m*/*z* 2278.41; calculated *m*/*z* 2278.62) and R (*m*/*z* 2282.19; calculated *m*/*z* 2282.62) conditions (Figure 3A,C). This result was confirmed using reflective MALDI-TOF-MS, by which a monoisotopic peak of *m*/*z* 2509.04 (calculated *m*/*z* 2509.07) was identified and corresponded to the same peptide, but with all C residues reacted with iodoacetamide (data not shown). Lastly, the

Table 4: MALDI-TOF-MS (Linear Mode) Analysis of HMWP-1 Digested by Trypsin[a]

Protein	Obs. m/z	Calc. m/z	Peptide	Position	Modifications
VE-α	1106.43	1106.26	DGQFVVVVSR	250-259	
VE-α	1317.2	1317.53	VLRDPVYAEVR	410-420	
VE-α	1454.54	1454.59	DSHYDLLFQCR	334-344	CysCAM
VE-α	1581.57	1581.8	YTGTSVETLIIEVR	345-358	
VE-α	1626.5	1625.95	MCFYGKTVTVQCTK	236-249	MSO
VE-α	2051.91	2052.21	SDSSYFPPGEFFSHYKR	471-487	
VE-α	2220.38	2220.53	TYPNPNPVVSVDAVLHVELR	359-378	
VE-α	2558.9	2558.92	VLCGLSGINAAQCQAISCCFDGR	213-235	4xCysCAM
VE-α	2594.29	2594.94	WDLLVNGCPYQDDRYLTVPIR	450-470	CysCAM
VE-α	2712.52	2713.21	MCFYGKTVTVQCTKDGQFVVVVSR	236-259	MSO
VE-α	3282.47	3282.8	ILGMTDPNVVLILEQCWANTSPTGE RLPR	421-449	CysCAM
VE-β	1090.13	1090.26	DGQFVVVVAR	211-220	
VE-β	1199.25	1199.39	TDPNIVLTLGR	385-395	
VE-β	1347.58	1347.55	VLRDPVYTEVR	370-380	
VE-β	1433.17	1433.53	DSQYDLTFQCR	295-305	CysCAM
VE-β	2131.23	2131.66	FVLKMFTFVDPMSMAPLR	446-463	
VE-β	2311.81	2311.59	YLTTPITVGPSSGLSYPTHYR	424-444	
VE-β	2467.38	2467.78	YLTTPITVGPSSGLSYPTHYRR	424-445	
VE-β	2651.79	2651.96	VQCGLPDITAAHCDAINCCFDGR	174-196	4xCysCAM
VE-β	2663.63	2663.99	ETVFIHCNTAVCLPSLGDSCEPR	464-486	3xCysCAM
VE-β	3335.01	3334.7	CWATTTPNPLSLPQWDLLIDGCPY QDDR	396-423	2xCysCAM
VE-β	4572.16	4572.33	MFTFVDPMSMAPLRETVFIHCNTA VCLPSLGDSCEPRCYR	450-489	2xMSO
VE-γ	1173.1	1173.35	LYFQVEAFR	321-329	
VE-γ	1312.82	1313.47	LMTADWQYER	236-245	
VE-γ	1329.37	1329.46	LMTADWQYER	236-245	MSO
VE-γ	1422.47	1422.63	YAEELLYFSMR	225-235	
VE-γ	1438.24	1438.63	YAEELLYFSMR	225-235	MSO
VE-γ	1516.6	1516.72	AECRENMVHVEAK	106-118	
VE-γ	1651.56	1651.87	YAFIENHGCLIDAK	291-304	CysCAM
VE-γ	1737.47	1737.95	SADYKLYFQVEAFR	316-329	
VE-γ	1751.36	1750.94	TNDAMINIECHYPR	191-204	CysCAM;MSO
VE-γ	2077.03	2077.15	EAGGNDGVCGCCDSTCSNR	391-409	4xCysCAM
VE-γ	2154.43	2154.42	QPPQQPQQPQQPPYQKPR	41-58	pyroglutamate
VE-γ	2171.13	2171.42	QPPQQPQQPQQPPYQKPR	41-58	
VE-γ	2245.32	2245.51	IIGSSEARTPVAANSVRAECR	89-109	CysCAM
VE-γ	2413.32	2412.74	TPVAANSVRAECRENMVHVEAK	97-118	

[a] All three VE proteins were detected. Peaks represent average *m*/*z*. CysCAM = carbamidomethylcysteine; MSO = methionine sulfoxide.

same peptide with the C-terminus as Y488 was also confirmed by ESI-Q-TOF-MS of chymotrypsin digests, where a triple charged (3+) peak of *m*/*z* 759.78 (calculated *m*/*z* 759.66) corresponded to the previous detected peptide ^{468}IHCNTAVCLPSLGDSCEPRCY^{488}, with the C residues disulfide-linked (Figure 4A).

For AspN digests, a peak of *m*/*z* 1129.01 (calculated *m*/*z* 1129.25) that corresponded to peptide ^{481}DSCEPRCYR^{489} was observed under R conditions with a C-terminal R residue (Table 2). Its oxidized form, with a *m*/*z* of 1126.71 (calculated *m*/*z* 1127.25), was also detected under NR conditions (Table 2). The C-terminus of VE*β* was also detected as peaks of *m*/*z* 3942.24 (calculated *m*/*z* 3942.55) and *m*/*z* 3946.31 (calculated *m*/*z* 3946.55), under NR and R conditions, respectively, and corresponded to peptide ^{455}DPMSMAPLRETVFIHCNTAVCLPSLGDSCEPRCYR^{489} with a C-terminal R residue (Table 2). It is noteworthy that similar analyses of VE*β*-containing HMWP-1 (see below) yielded a peak of *m*/*z* 4572.16 (calculated *m*/*z* 4572.33) that corresponded to peptide ^{450}MFTFVDPMSMAPLRETVFIHCNTAVCLPSLGDSCEPRCYR^{489} with a C-terminal R residue (Table 4). Taken together, these data suggest that the C-terminus of VE*β* is probably R489 (chymotrypsin can also cleave between Y488 and R489) but could also be Y488.

VEα possesses two potential furin-like cleavage sites, ^{529}RQRR^{533} and ^{541}KKTK^{544}, but the alignments of VE*β* and VEα suggest that the first site should be cleaved (Figure 2A). Unfortunately, we have no evidence for the nature of the C-terminal amino acid from MALDI-TOF-MS analyses of protease-digested VEα.

Determination of Glycosylation Sites in VE Proteins. Previously, it was reported that of the four rainbow trout VE proteins, only VE*γ* and HMWP are N-glycosylated (*13*).

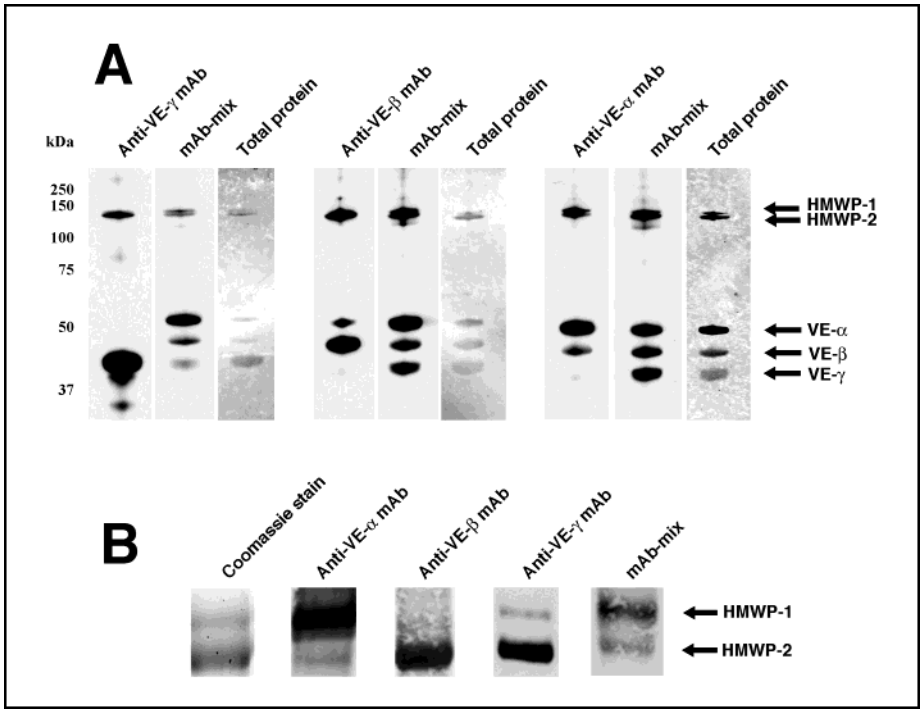

FIGURE 8: Western blot of VE proteins probed with monoclonal antibodies. (A) VE proteins were separated on a 10% SDS−PAGE minigel and transferred to a nitrocellulose membrane. VE proteins were first probed with one of three monoclonal antibodies against *S. salar* zona radiata proteins: anti-VEγ (MN-7F2), anti-VEβ (MN-2B4), and anti-VEα (MN-8C4). The membrane was then stripped and reprobed with a 1:1:1 mixture of all three antibodies. The total protein on the membrane, stained with colloidal gold, is also shown. M_r markers are on the left, and VE proteins are on the right. (B) VE proteins were separated on a 10% SDS−PAGE maxigel for better separation in the high M_r range. One lane was Coomassie-stained, and the other lanes were transferred to a nitrocellulose membrane. HMWPs (on the right) were then probed with each of the three antibodies and with a 1:1:1 mixture of antibodies.

Indeed, following treatment of VE proteins under denaturing conditions with *N*-glycanase (48 h at 37 °C), only VEγ and HMWP underwent a small reduction (~4 kDa) in M_r; the M_rs of VEα and VEβ remained unaltered (data not shown). VEγ possesses a single consensus sequence (NXS/T) for N-linked glycosylation (^{227}NVS^{229}), and amino acid sequencing of a peptide from VEγ containing this sequence suggests that N227 is indeed glycosylated. As discussed below, HMWP contains VEγ, thus accounting for the glycosylation.

VE Proteins Assemble into Heterodimers. HMWP is visualized as a single band of M_r ~120−130 kDa following SDS−PAGE under R conditions. However, with longer separation times on SDS−PAGE under NR conditions, two bands, called HMWP-1 and HMWP-2, are observed (Figure 8B). HMWP-2 stains intensely with Coomassie blue, whereas HMWP-1 stains weakly.

To determine the M_rs of these proteins more accurately, their electrophoretic mobility was examined using native and deglycosylated forms of the proteins, different M_r markers, and different percentage gels. The average M_rs for HMWP-1 and HMWP-2 are (under R conditions) ~118 and ~111 kDa, respectively. The M_r of each protein was reduced by ~3 kDa after enzymatic deglycosylation.

To determine the relationship between HMWPs and other VE proteins, monoclonal antibodies (mAb) directed against the highly related *S. salar* zona radiata proteins were employed. Anti-VEγ (MN-7F2) immunoreacted with VEγ and HMWPs but not with VEβ or VEα (Figure 8A). Anti-VEβ (MN2B4) immunoreacted with VEβ and HMWPs, to a much lesser extent with VEα, and not at all with VEγ. Anti-VEα (MN-8C4) immunoreacted with VEα and HMWPs, to a lesser extent with VEβ, and not at all with VEγ. These observations strongly suggest that the HMWPs consist of the lower M_r VEα, VEβ, and VEγ proteins.

To obtain further information about the composition of HMWP-1 and HMWP-2, the proteins were separated more fully from each other by SDS−PAGE, and the same antibodies were used in immunoblots against them (Figure 8B). Anti-VEα reacted intensely with HMWP-1 and produced a faint band with HMWP-2. Anti-VEβ reacted with HMWP-2 but not at all with HMWP-1. Anti-VEγ reacted with HMWP-1 and HMWP-2: very intensely with the latter and faintly with the former. From these immunoblots it was concluded that VEγ is present in both HMWP-1 and HMWP-2, VEα is present only in HMWP-1, and VEβ is present only in HMWP-2. These data are supported by SDS−PAGE experiments in which the M_r of HMWPs were determined. The determined M_r of deglycosylated HMWP-1 (115 kDa) is close to the calculated mass of a deglycosylated VEα−VEγ heterodimer (102 kDa), and the M_r detected for deglycosylated HMWP-2 (108 kDa) is close to the calculated mass of a VEβ−VEγ heterodimer (97 kDa). Collectively, these data suggest that HMWP-1 is a VEα−VEγ heterodimer and HMWP-2, the more abundant species, is a VEβ−VEγ heterodimer.

To obtain more information about the composition of HMWPs, MALDI-TOF-MS in linear mode was carried out on proteolytic digests of these proteins. The MALDI-TOF-MS spectrum of a trypsin digest of HMWPs is shown in Figure 9. In the spectrum of HMWP-1 (Figure 9A), aside from the expected peptides for VEα and VEγ, peaks corresponding to peptides from VEβ also were observed. The latter finding conflicts with results of immunoblotting experiments described above and suggests that a small

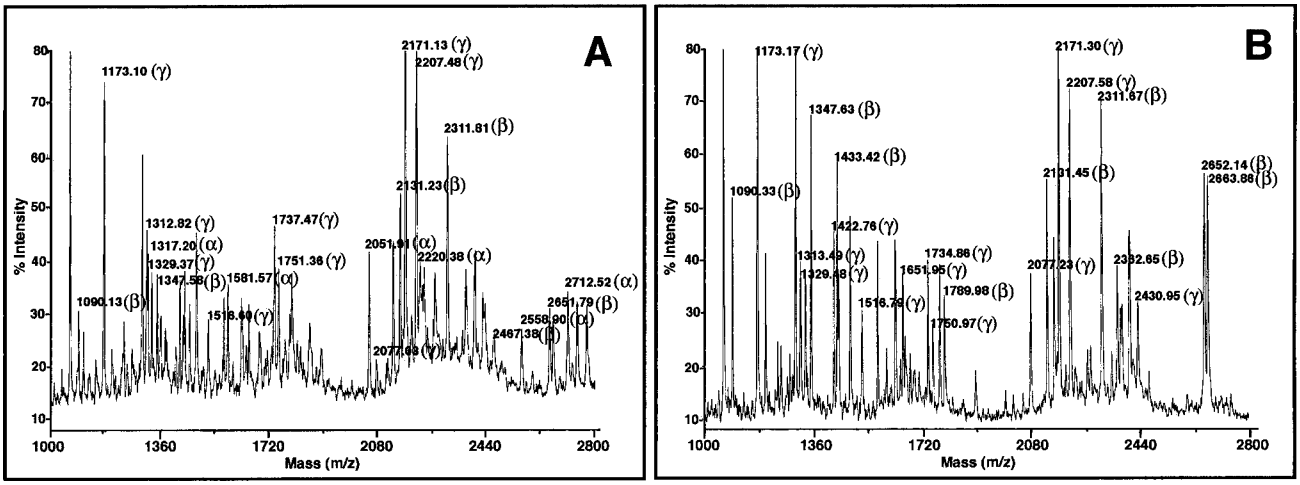

FIGURE 9: MALDI-TOF-MS in linear mode of trypsin-digested HMWP-1 (A) and HMWP-2 (B), measured under R conditions. The average *m/z* peaks marked α, *β*, and *γ* represent peptide fragments of VEα, VE*β*, and VE*γ*. Note that in (A) all VE proteins were detected, while in (B) only VE*β* and VE*γ* were found.

amount of a VEα−VE*β* or VE*β*−VE*β* dimer may be present in HMWP-1. The presence of VE*β* was confirmed by MALDI-TOF-MS analysis of AspN digests of HMWP-1 (data not shown). On the other hand, spectra of HMWP-2 yielded peaks corresponding only to peptides of VE*β* and VE*γ*, as expected for a VE*β*−VE*γ* heterodimer (Figure 9B) and confirmed by MALDI-TOF-MS analysis of AspN digests. A list of VEα, VE*β*, and VE*γ* peptides detected in trypsin digests of HMWP-1 and HMWP-2 is presented in Tables 4 and 5. A similar list for AspN digests of HMWP-1 and HMWP-2 analyzed by MALDI-TOF-MS in linear mode is presented in Tables 1A and 1B (see Supporting Information).

DISCUSSION

The protein composition of fish VEs has been described for carp (*30*, *31*), winter flounder (*32*), goldfish (*30*), medaka (*33*−*36*), cod (*37*), sea bream (*38*, *39*), sea bass (*40*), rainbow trout (*11*, *13*, *40*, *41*), and zebrafish (*42*−*45*), among others. Overall, fish VE proteins are a highly conserved group of proteins that are related to each other, as well as to their amphibian, avian, and mammalian egg coat counterparts. For example, rainbow trout VEα and VE*β* are related to mammalian ZP1 and ZP2 and VE*γ* is homologous with ZP3 (*40*).

Since it is well documented that disulfide linkages restrict the number of possible conformations for a specific polypeptide and are a major force involved in stabilizing polypeptide three-dimensional structure (*46*, *47*), we determined the positions of disulfides in rainbow trout egg VE proteins using primarily MALDI-TOF-MS (linear mode) under NR and R conditions (*48*−*51*). C-containing peptides that resulted from laser-dependent homolytic fragmentation of disulfides (*52*−*55*) also were used as additional evidence for positions of disulfide linkages. Furthermore, MALDI-TOF-MS (linear mode) information was supplemented with data from one- and two-dimensional SDS−PAGE, MALDI-TOF-MS (reflective mode), or ESI-Q-TOF-MS.

The ZP domain is an ∼260 amino acid region of ZP proteins (*5*) that is found in a large number of other kinds of proteins from a wide range of organisms, from urochordates to mammals (*6*, *7*). The functions of ZP domain proteins vary tremendously, from serving as structural components of mammalian egg coats, appendicularian mucous houses, and nematode dauer larvae to serving as mechanotransducers in flies and receptors in mammals. It has been reported that the ZP domain is responsible for polymerization of proteins into higher order structures, such as fibrils, filaments, and matrices (*7*).

Disulfides of VE Proteins. The ZP domain of ZP and VE proteins contains 8 conserved C residues (C1−C8; Figure 2) (*5*−*7*). By combining chemical digestion and two-dimensional SDS−PAGE with enzymatic digestion and both MALDI-TOF-MS (linear mode) and ESI-Q-TOF-MS for VE*γ*, disulfide linkages could be assigned to its ZP domain: C1−C4, C2−C3, C5−C7, and C6−C8 (Figure 10C). These results are consistent with the positions of disulfides in mouse ZP3 (*19*) and partial results reported for pig ZPC (*56*) and suggest that these disulfides, as well as the overall protein structure, are conserved from fish to mammals. These results are also supported by sequence alignments of ZP3-like proteins that reveal that when C2 is missing from a protein, C3 is absent as well (e.g., ZP glycoprotein-3 from zebrafish; NCBI no. 18859585). By comparing sequence alignments of ZP3-like proteins from different organisms (Figure 2B), it can be concluded that the last 4 C residues, downstream of the ZP domain, are likely to be disulfide-linked, C9−C11 and C10−C12, since C10 and C12 are not absolutely conserved residues.

Disulfide linkages of VE*β* and VEα were determined by MALDI-TOF-MS in linear mode and confirmed by MALDI-TOF-MS in reflective mode or ESI-Q-TOF-MS. In addition to 8 conserved C residues in the ZP domain of VE*β* and VEα are 2 other C residues, Ca and Cb, that are found in ZP1- and ZP2-like proteins of fish and mammals. Two other C residues, Cx and Cy, are also present in fish VEα and VE*β* but are not present in mammalian ZP proteins. For VE*β*, disulfide linkages could be assigned C1−C4, C2−C3, Cx−Cy, C5−C6, C7−Ca, and Cb−C8 (Figure 10B). For VEα, C2−C3, Cx−Cy, and C7−Ca could be assigned from MALDI-TOF-MS spectra, and the other three disulfide linkages were inferred (C1−C4, C5−C6, Cb−C8) from sequence similarities between VEα and VE*β* (Figure 10A). It is of interest to note that Cx and Cy, residues specific to

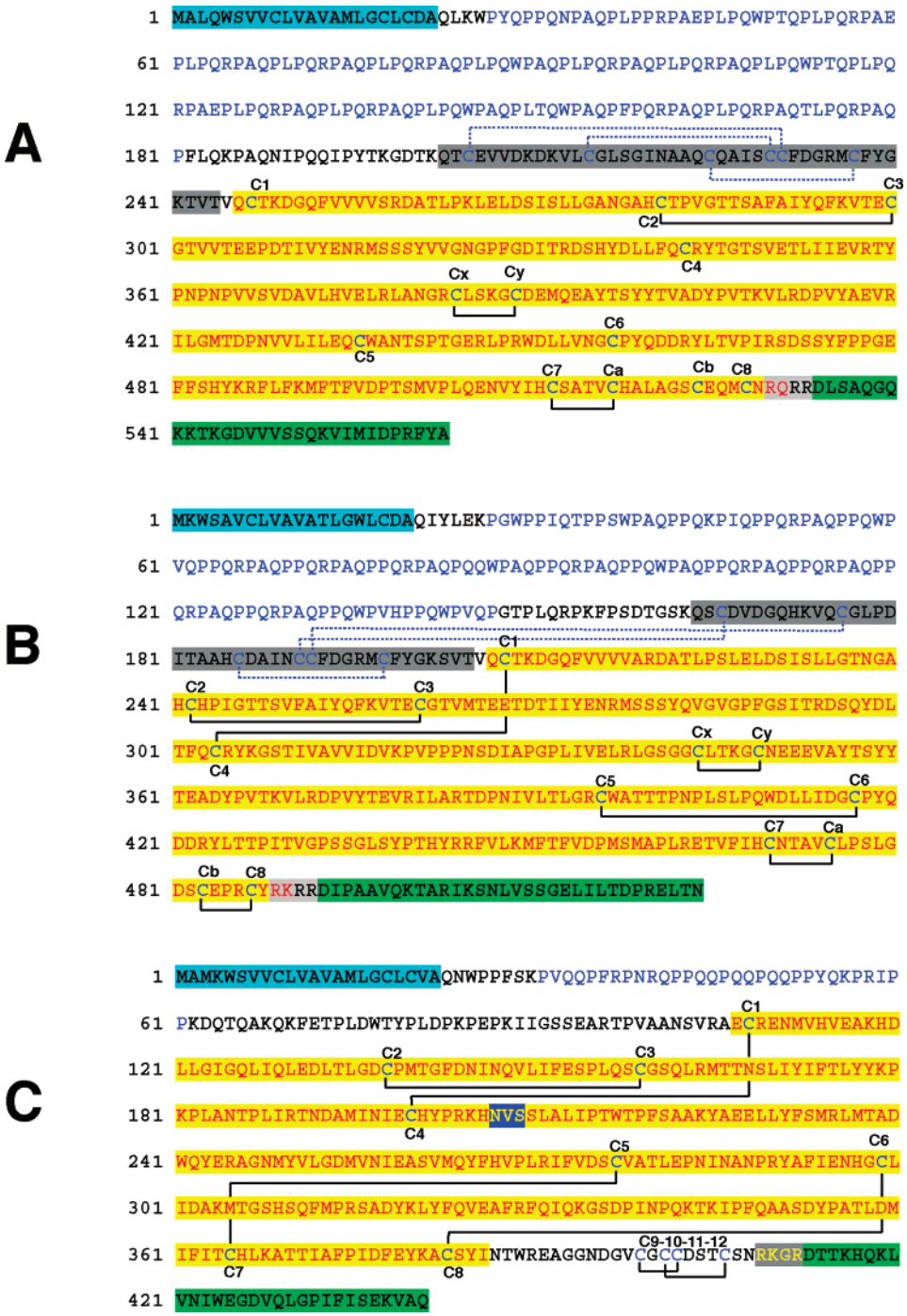

FIGURE 10: Summary for VEα (A), VEβ (B), and VEγ (C). Each protein contains a signal peptide (highlighted in turquoise), a PQ-rich region at the N-terminus (blue letters), a ZP domain containing eight conserved C residues (red letters and highlighted in yellow), and a hydrophobic C-terminal tail (highlighted in dark green) that is removed by a furin-like protease(s). The furin-like cleavage site is shown in gray and lies before the C-terminal tail. Note that VEα contains a second potential cleavage site, ^{541}KKTK^{544}. The trefoil domain of VEα and VEβ, situated upstream of ZP domain, is represented by black letters, highlighted in gray. C residues contained within the trefoil domain are linked C1−C5, C2−C4, and C3−C6 (*18*). Other C residues are designated according to Figure 2. The disulfide linkages determined for each protein are shown. The ^{227}NVS^{229} glycosylation site determined for VEγ is represented by yellow letters, highlighted in blue.

fish ZP1-like proteins and not present in mammals, are disulfide-linked. Furthermore, disulfides C1−C4 and C2−C3 are found in VEα, VEβ, and VEγ (reported here) and in mouse ZP2 and ZP3 proteins (*19*), suggesting that the structure of the N-terminal portion of the ZP domain (containing C1−C4) is highly conserved in fish and mammals. However, reports differ concerning disulfide assignments for the second half of the ZP domain of ZP1/ZP2-like proteins. For example, for mouse, C residues are linked C5−C6 (ZP1) and C7−Ca and C8−Cb (ZP2) (*19*), whereas for pig ZPB, a ZP1-like protein, the disulfides appear to be C5−C7, C6−C8, and Ca−Cb (*56*). A possibility for the differences is that the disulfide linkages determined for pig ZPB could result from disulfide exchange that is frequently observed at alkaline pH (pH 9.0 was used for lysylendopeptidase digestion of ZPB).

In terms of functions and disulfide linkages, there are two different types of ZP domain-containing proteins, ZP3-like and ZP2/ZP1-like proteins. Both have the same C1−C4 and C2−C3 disulfides but differ in the second part of the ZP

Table 5: MALDI-TOF-MS (Linear Mode) Analysis of HMWP-2 Digested by Trypsin[a]

Protein	Obs. m/z	Calc. m/z	Peptide	Position	Modifications
VE-β	1090.33	1090.26	DGQFVVVVAR	211-220	
VE-β	1199.6	1199.39	TDPNIVLTLGR	385-395	
VE-β	1347.63	1347.55	VLRDPVYTEVR	370-380	
VE-β	1433.42	1433.53	DSQYDLTFQCR	295-305	CysCAM
VE-β	1675.97	1676.01	MFTFVDPMSMAPLR	450-463	2xMSO
VE-β	1773.94	1774.01	MSSSYQVGVGPFGSITR	278-294	
VE-β	1789.98	1790.01	MSSSYQVGVGPFGSITR	278-294	MSO
VE-β	2131.45	2131.66	FVLKMFTFVDPMSMAPLR	446-463	
VE-β	2311.67	2311.59	YLTTPITVGPSSGLSYPTHYR	424-444	
VE-β	2362.65	2362.58	VTECGTVMTEETDTIIYENR	258-277	Cys
VE-β	2378.22	2378.58	VTECGTVMTEETDTIIYENR	258-277	CysCAM;MSO
VE-β	2467.38	2467.78	YLTTPITVGPSSGLSYPTHYRR	424-445	
VE-β	2652.14	2651.96	VQCGLPDITAAHCDAINCCFDGR	174-196	4xCysCAM
VE-β	2663.63	2663.99	ETVFIHCNTAVCLPSLGDSCEPR	464-486	3xCysCAM
VE-β	3210.71	3210.71	VQCGLPDITAAHCDAINCCFDGR	174-202	CysCAM
			MCFYGK		
VE-γ	915.18	915.03	TPVAANSVR	97-105	
VE-γ	1171.03	1171.31	ACSYINTWR	382-390	CysCAM
VE-γ	1173.17	1173.35	LYFQVEAFR	321-329	
VE-γ	1279.45	1279.47	MTGSHSQFMPR	305-315	
VE-γ	1313.49	1313.47	LMTADWQYER	236-245	
VE-γ	1329.48	1329.46	LMTADWQYER	236-245	MSO
VE-γ	1422.76	1422.63	YAEELLYFSMR	225-235	
VE-γ	1438.46	1438.63	YAEELLYFSMR	225-235	MSO
VE-γ	1516.79	1516.72	AECRENMVHVEAK	106-118	
VE-γ	1651.95	1651.87	YAFIENHGCLIDAK	291-304	CysCAM
VE-γ	1734.86	1734.94	TNDAMINIECHYPR	191-204	CysCAM
VE-γ	1750.97	1750.94	TNDAMINIECHYPR	191-204	CysCAM;MSO
VE-γ	2077.23	2077.15	EAGGNDGVCGCCDSTCSNR	391-409	4xCysCAM
VE-γ	2171.3	2171.42	QPPQQPQQPQQPPYQKPR	41-58	
VE-γ	2207.58	2207.53	QNWPPFSKPVQQPFRPNR	23-40	pyroglutamate
VE-γ	2430.95	2430.76	QKFETPLDWTYPLDPKPEPK	67-88	

[a] Only VEβ and VEγ were detected. Peaks represent average *m/z*. CysCAM = carbamidomethylcysteine; MSO = methionine sulfoxide.

domain. For VEβ the C residues are linked C5−C6, C7−Ca, and C8−Cb (Figure 10B), whereas for VEγ they are linked C5−C7 and C6−C8 (Figure 10C). These different disulfide arrangements could explain why the presence of both ZP2 and ZP3 is required for assembly of the mouse ZP (*57*−*59*) and why assembly probably starts with formation of ZP2−ZP3 heterodimers (*60*). On the other hand, it does not explain why uromodulin, a ZP2-like protein, forms filaments in the absence of a ZP3-like protein (*7*). Differences in the disulfide linkages of ZP1/ZP2/VEβ/VEα could reflect that they are initially present as in pig ZPB, but during polymerization disulfide exchange takes place, resulting in disulfide linkages as reported for the fish and mouse proteins. This could account for the finding that in some ZP1/ZP2-related proteins when C6 is missing, C8 is also missing (NCBI no. 1216409), and when C5 is missing, C7 is missing as well (NCBI no. 1314322 and no. 7301839). It would also account for why uromodulin and, perhaps, some other ZP1/ZP2-related proteins (cuticlin, NCBI no. 2648041; vomeroglandin, NCBI no. 7209584; ebnerin, NCBI no. 1363286) can form homomeric filaments without the presence of a ZP3-like protein.

Perhaps, ZP1/ZP2-like proteins are structurally versatile proteins. Their standard disulfide linkages are C5−C7, C6−C8, and Ca−Cb, but during assembly of homopolymers some of the protein undergoes disulfide exchange to C5−C6, C7−Ca, and C8−Cb, and its conformation is altered. The same protein, present in two conformations, is now able to assemble into filaments. This could explain why ZP1/ZP2-related proteins, such as uromodulin, can form homopolymers. Alternatively, ZP2-like proteins such as α- and β-tectorin (gi, 1915909; gi, 7339532) (*61*, *62*) that form heteropolymers are already specialized, with one of the two disulfide linkages mentioned above. Deletion of α-tectorin would suppress expression of β-tectorin and, consequently, formation of α/β-tectorin heteropolymers (*63*). Similarly, for ZP assembly, two proteins are required; ZP2 with disulfide linkages C5−C6, C7−Ca, and C8−Cb and ZP3 with disulfide linkages C5−C7 and C6−C8.

Maturation of VE Proteins. Fish VEα, VEβ, and VEγ are secreted proteins. During processing, VE proteins undergo removal of their N-terminal signal sequence that directs them to the endoplasmic reticulum. To date, the N-terminal amino acid has not been reported for any fish VE protein. Here, we found that the N-terminal amino acid of VEγ, after removal of the signal sequence, is Q23 as pyroglutamic acid. In addition to pyroglutamic acid, unmodified Q23 was also found at the N-terminus of VEγ. Preliminary data suggest that the N-terminus of VEβ is also an unmodified Q21. In this context, for mouse ZP1 and ZP3 the N-terminal amino

acid is Q as pyroglutamic acid (Q21 for ZP1 and Q23 for ZP3), whereas for ZP2 it is V35 (*19*).

Brivio et al. (*13*) reported that, for fish VE proteins, only VEγ and HMWP have N-linked oligosaccharides. We found that VEγ has only one glycosylated residue, N227, and that the same residue is glycosylated in HMWPs, which contain VEγ. In an alignment of VEγ and mammalian ZP3 proteins it was observed that this particular N-linked glycosylation site of VEγ is highly conserved in mammals as well. Indeed, Boja et al. (*19*) detected that the corresponding glycosylation site of mouse ZP3, N146, is occupied.

After removal of the N-terminal signal sequence and after N-linked glycosylation, the final step in maturation of VE proteins is removal of the C-terminal tail. This processing step is carried out by furin-like enzymes (proprotein convertases) at a consensus cleavage site, possibly followed by trimming of some basic amino acids by carboxypeptidase H (*64*). While in mammals the cleavage site of ZP-like proteins is R-X-K/R-R (*7*), it seems that in fish it is R-K-X-R (*29*). Reports differ insofar as the C-terminal amino acid of ZP-like and VE-like proteins. The C-termini of mature mouse ZP proteins are R546 for ZP1 (^{545}RRRR^{548}) (*19*), S633 (*19*) or R635 (*65*) for ZP2 (^{632}RSKR^{635}), and N351 (*19*) or R353 (*65*) for ZP3 (^{350}RNRR^{353}). The C-termini of pig ZPB and ZPC are A462 (^{462}ARRRR^{466}) and S332 (^{332}SRKLS^{336}), respectively (*56*). Quail ZPC ends with F369 (^{368}RFRR^{371}) (*66*) and *Xenopus* gp43 with Q370 (^{367}RWNQKR^{372}) (*64*). The C-termini of medaka ZI proteins are G557 (^{555}RKGR^{558}) for ZI-1 and ZI-2 and T393 (^{391}RKTR^{394}) for ZI-3 (*29*). Our studies suggest that the C-terminus of VEγ is K410 (^{409}RKGR^{412}) and for VEβ is R489 and, perhaps, Y488 (^{488}YRKRR^{492}). Collectively, these results suggest that cleavage of the C-terminal tail takes place at a proprotein convertase cleavage site with the more general sequence (K/R)X$_n$(K/R), with $n = 0$, 2, 4, or 6 (*67*). Apparently, the only difference is trimming of the cleaved end of these proteins that is carried out by carboxypeptidase H, an event that may occur to different extents depending on various other factors.

Assembly of VE Proteins. Assembly of the fish VE takes place in the ovary; however, synthesis of VE proteins can occur in the liver as well as the ovary (*10*). Rainbow trout VE proteins α, β, and γ are expressed in livers of both male and female fish, although mRNA of VEγ is also detected in the ovary (*11*). It has been reported that, in addition to VEα, VEβ, and VEγ of rainbow trout, there is a fourth VE protein (HMWP) with a M_r in the range of ~110−129 kDa that shares antigenic determinants with VEγ (*13*, *41*). Relatively high M_r proteins have also been found in VEs from frogs (120 kDa) (*68*), fish (170 kDa) (*69*), chickens (180 kDa) (*70*), and other vertebrates. Here, we found that HMWPs (M_r ~120−130 kDa) from VEs of rainbow trout are dimers consisting of combinations of VEα, VEβ, and VEγ. These covalently linked species may reflect a low degree of cross-linking of N-terminal extensions of VE proteins by transglutaminase, as reported for other fish (*8*). On the basis of results of Coomassie staining and immunoblots, the most abundant heterodimers are VEβ−VEγ and VEα−VEγ. The least abundant dimer, detected only by mass spectrometry, could be a VEα−VEβ heterodimer or a VEβ−VEβ homodimer. In all cases, the oligosaccharide component of these heterodimers is contributed by VEγ. In most (*33*, *29*, *40*) but perhaps not all (*71*) instances, it would appear that heterodimerization of VE proteins takes place in the ovary, not in the liver, of fish. Results reported here suggest that, much as mouse ZP2 and ZP3 form heterodimers that polymerize into ZP filaments (*60*, *72*), VE heterodimers probably serve as the building blocks of fish VE filaments.

Final Comments. Results presented here provide some clues as to why some ZP domain proteins (e.g., uromodulin and cuticlin) can form homopolymers, whereas others (e.g., ZP3) cannot. The former ZP1/ZP2-like proteins may undergo disulfide exchange in the C-terminal half of their ZP domains that permits interactions necessary for polymerization of the proteins into filaments. Due to the lower number of C residues, the disulfides of proteins such as ZP3 may not be exchanged, and this could necessitate the presence of a ZP1/ZP2-like protein for assembly into heterodimers and filaments. Our finding that VEβ−VEγ and VEα−VEγ are the principal heterodimers in trout egg VE HMWPs is consistent with this hypothesis.

ACKNOWLEDGMENT

We thank Norman Soule and the staff of the Cold Spring Harbor Fish Hatchery (Cold Spring Harbor, NY) for providing us with generous hospitality and buckets of rainbow trout eggs. We also thank Franco Cotelli (University of Milan, Italy) for excellent advice about trout egg VE preparations, Mary Ann Gawinowicz and Yelena Milgrom (Protein Core Facility, Columbia University, New York) for mass spectrometry measurements and invaluable discussions, Anders Goksoyr (Biosense Laboratories AS, Bergen, Norway) for a gift of monoclonal antibodies against fish egg VE proteins, and Joerg O. Thumfart and Heidi Braeuner (University of Freiburg, Germany) and Alisa G. Woods for measurements, advice, and discussion.

SUPPORTING INFORMATION AVAILABLE

Two tables showing MALDI-TOF-MS analysis of HMWP-1 and HMWP-2 digested with AspN. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Dumont, J. N., and Brummett, A. R. (1985) Egg envelopes in vertebrates, in *Developmental Biology: A Comprehensive Synthesis* (Browder, E. R., Ed.), Vol. 1, pp 235−288, Plenum Press, New York.
2. Laale, H. W. (1980) The perivitelline space and egg envelopes of bony fishes: a review, *Copeia* 210−226.
3. Hardy, D. M., Ed. (2002) Fertilization, in *Fertilization*, 427 pp, Academic Press, San Diego, CA.
4. Wassarman, P. M., Ed. (2003) Fertilization, *ChemTracts-Biochem. Mol. Biol. 16*, 117−204.
5. Bork, P., and Sander, C. (1992) A large domain common to sperm receptors (Zp2 and Zp3) and TGF-beta type III receptor, *FEBS Lett. 300*, 237−240.
6. Wassarman, P. M., Jovine, L., and Litscher, E. S. (2001) A profile of fertilization in mammals, *Nat. Cell Biol. 3*, E59−E64.
7. Jovine, L., Qi, H., Williams, Z., Litscher, E., and Wassarman, P. M. (2002) The ZP domain is a conserved module for polymerization of extracellular proteins, *Nat. Cell Biol. 4*, 457−461.
8. 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.
9. Sugiyama, H., and Iuchi, I. (2000) The third egg envelope subunit in fish: cDNA cloning and analysis, and gene expression, *Recent Res. Dev. Comp. Biochem. Physiol. 1*, 139−161.

10. Conner, S. J., and Hughes, D. C. (2003) Analysis of fish ZP1/ZPB homologous genes—evidence for both genome duplication and species-specific amplification models of evolution, *Reproduction 126*, 347−352.
11. Hyllner, S. J., Oppen-Berntsen, D. O., Helvik, J. V., Walther, B. T., and Haux, C. (1991) Oestradiol-17 beta induces the major vitelline envelope proteins in both sexes in teleosts, *J. Endocrinol. 131*, 229−236.
12. Hyllner, S. J., and Haux, C. (1992) Immunochemical detection of the major vitelline envelope proteins in the plasma and oocytes of the maturing female rainbow trout, Oncorhynchus mykiss, *J. Endocrinol. 135*, 303−309.
13. Brivio, M. F., Bassi, R., and Cotelli, F. (1991) Identification and characterization of the major components of the Oncorhynchus mykiss egg chorion, *Mol. Reprod. Dev. 28*, 85−93.
14. Thim, L. (1989) A new family of growth factor-like peptides. "Trefoil" disulphide loop structures as a common feature in breast cancer associated peptide (pS2), pancreatic spasmolytic polypeptide (PSP), and frog skin peptides (spasmolysins), *FEBS Lett. 250*, 85−90.
15. Carr, M. D. (1992) ^{1}H NMR-based determination of the secondary structure of porcine pancreatic spasmolytic polypeptide: one of a new family of "trefoil" motif containing cell growth factors, *Biochemistry 31*, 1998−2004.
16. Bork, P. (1993) A trefoil domain in the major rabbit zona pellucida protein, *Protein Sci. 2*, 669−670.
17. Wong, W. M., Poulsom, R., and Wright, N. A. (1999) Trefoil peptides, *Gut 44*, 890−895.
18. Carr, M. D., Bauer, C. J., Gradwell, M. J., and Feeney, J. (1994) Solution structure of a trefoil-motif-containing cell growth factor, porcine spasmolytic protein, *Proc. Natl. Acad. Sci., U.S.A. 91*, 2206−2210.
19. Boja, E. S., Hoodbhoy, T., Fales, H. M., and Dean, J. (2003) Structural characterization of native mouse zona pellucida proteins using mass spectrometry, *J. Biol. Chem. 278*, 34189−34202.
20. Hunkapiller, M. W., Lujan, E., Ostrander, F., and Hood, L. E. (1983) Isolation of microgram quantities of proteins from polyacrylamide gels for amino acid sequence analysis, *Methods Enzymol. 91*, 227−236.
21. Hellman, U., Wernstedt, C., Gonez, J., and Heldin, C. H. (1995) Improvement of an "In-Gel" digestion procedure for the micropreparation of internal protein fragments for amino acid sequencing, *Anal. Biochem. 224*, 451−455.
22. Shevchenko, A., Wilm, M., Vorm, O., Jensen, O. N., Podtelejnikov, A. V., Neubauer, G., Mortensen, P., and Mann, M. (1996) A strategy for identifying gel-separated proteins in sequence databases by MS alone, *Biochem. Soc. Trans. 24*, 893−896.
23. Fontana, A. (1986) Fragmentation of polypeptides by chemical methods, in *Methods in Protein Chemistry: A Handbook* (Darbre, A., Ed.) pp 68−120, John Wiley & Sons, New York.
24. Smith, B. (1994) Chemical cleavage of proteins, in *Methods in Molecular Biology* (Walker, J., Ed.) Vol. 32, pp 297−309, Humana Press, Totowa, NJ.
25. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature 227*, 680−685.
26. Schaegger, H., and von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa, *Anal. Biochem. 166*, 368−379.
27. Heukeshoven, J., and Dernick, R. (1988) Improved silver staining procedure for fast staining in PhastSystem Development Unit. I. Staining of sodium dodecyl sulfate gels, *Electrophoresis 9*, 28−32.
28. Kussmann, M., Nordhoff, E., Rahbek-Nielsen, H., Haebel, S., Rossel-Larsen, M., Jakobsen, L., Gobom, J., Mirgorodskaya, E., Kroll-Kristensen, A., Palm, L., and Roepstorff, P. (1997) Matrix-assisted laser desorption/ionization mass spectrometry sample preparation techniques designed for various peptide and protein analytes, *J. Mass Spectrom. 32*, 593−601.
29. Sugiyama, H., Murata, K., Iuchi, I., Nomura, K., and Yamagami, K. (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, *J. Biochem. (Tokyo) 125*, 469−475.
30. Chang, Y. S., Wang, S. C., Tsao, C. C., and Huang, F. L. (1996) Molecular cloning, structural analysis, and expression of carp ZP3 gene, *Mol. Reprod. Dev. 44*, 295−304.
31. Chang, Y. S., Hsu, C. C., Wang, S. C., Tsao, C. C., and Huang, F. L. (1997) Molecular cloning, structural analysis, and expression of carp ZP2 gene, *Mol. Reprod. Dev. 46*, 258−267.
32. Lyons, C. E., Payette, K. L., Price, J. L., and Huang, R. C. (1993) Expression and structural analysis of a teleost homologue of a mammalian zona pellucida gene, *J. Biol. Chem. 268*, 21351−21358.
33. 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., U.S.A. 94*, 2050−2055.
34. 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.
35. Sugiyama, H., Yasumasu, S., Murata, K., Iuchi, I., and Yamagami, K. (1998) The third egg envelope subunit in fish: cDNA cloning and analysis, and gene expression, *Dev. Growth Differ. 40*, 35−45.
36. Kanamori, A. (2000) Systematic identification of genes expressed during early oogenesis in medaka, *Mol. Reprod. Dev. 55*, 31−36.
37. Oppen-Berntsen, D. O., Helvik, J. V., and Walther, B. T. (1990) The major structural proteins of cod (Gadus morhua) eggshells and protein cross-linking during teleost egg hardening, *Dev. Biol. 137*, 258−265.
38. Hyllner, S. J., Fernandez-Palacios Barber, H., Larsson, D. G., and Haux, C. (1995) Amino acid composition and endocrine control of vitelline envelope proteins in European sea bass (Dicentrarchus labrax) and gilthead sea bream (Sparus aurata), *Mol. Reprod. Dev. 41*, 339−347.
39. Del Giacco, L., Vanoni, C., Bonsignorio, D., Duga, S., Mosconi, G., Santucci, A., and Cotelli, F. (1998) Identification and spatial distribution of the mRNA encoding the gp49 component of the gilthead sea bream, Sparus aurata, egg envelope, *Mol. Reprod. Dev. 49*, 58−69.
40. Hyllner, S. J., Westerlund, L., Olsson, P. E., and Schopen, A. (2001) Cloning of rainbow trout egg envelope proteins: members of a unique group of structural proteins, *Biol. Reprod. 64*, 805−811.
41. Iuchi, I., Masuda, K., and Yamagami, K. (1991) Change in component proteins of the egg envelope (chorion) of rainbow trout during hardening, *Dev. Growth Differ. 33*, 85−92.
42. Wang, H., and Gong, Z. (1999) Characterization of two zebrafish cDNA clones encoding egg envelope proteins ZP2 and ZP3, *Biochim. Biophys. Acta 1446*, 156−160.
43. Bonsignorio, D., Perego, L., Del Giacco, L., and Cotelli, F. (1996) Structure and macromolecular composition of the zebrafish egg chorion, *Zygote 4*, 101−108.
44. Del Giacco, L., Diani, S., and Cotelli, F. (2000) Identification and spatial distribution of the mRNA encoding an egg envelope component of the Cyprinid zebrafish, Danio rerio, homologous to the mammalian ZP3 (ZPC), *Dev. Genes Evol. 210*, 41−46.
45. Mold, D. E., Kim, I. F., Tsai, C. M., Lee, D., Chang, C. Y., and Huang, R. C. (2001) Cluster of genes encoding the major egg envelope protein of zebrafish, *Mol. Reprod. Dev. 58*, 4−14.
46. Fischer, W. H., Greenwald, J., Park, M., Craig, A. G., Choe, S., and Vale, W. (1999) The disulfide bond arrangement in the extracellular domain of the activin type II receptor, *J. Protein Chem. 18*, 437−446.
47. Wedemeyer, W. J., Welker, E., Narayan, M., and Scheraga, H. A. (2000) Disulfide bonds and protein folding, *Biochemistry 39*, 4207−4216.
48. Crimmins, D. L., Saylor, M., Rush, J., and Thoma, R. S. (1995) Facile, in situ matrix-assisted laser desorption ionization-mass spectrometry analysis and assignment of disulfide pairings in heteropeptide molecules, *Anal. Biochem. 226*, 355−361.
49. Moritz, R. L., Hall, N. E., Connolly, L. M., and Simpson, R. J. (2001) Determination of the disulfide structure and N-glycosylation sites of the extracellular domain of the human signal transducer gp130, *J. Biol. Chem. 276*, 8244−8253.
50. Pitt, J. J., Da Silva, E., and Gorman, J. J. (2000) Determination of the disulfide bond arrangement of Newcastle disease virus hemagglutinin neuraminidase. Correlation with a beta-sheet propeller structural fold predicted for paramyxoviridae attachment proteins, *J. Biol. Chem. 275*, 6469−6478.

51. Tie, J. K., Mutucumarana, V. P., Straight, D. L., Carrick, K. L., Pope, R. M., and Stafford, D. W. (2003) Determination of disulfide bond assignment of human vitamin K-dependent gamma-glutamyl carboxylase by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *J. Biol. Chem. 278*, 45468−45475.
52. Haniu, M., Arakawa, T., Bures, E. J., Young, Y., Hui, J. O., Rohde, M. F., Welcher, A. A., and Horan, T. (1998) Human leptin receptor. Determination of disulfide structure and N-glycosylation sites of the extracellular domain, *J. Biol. Chem. 273*, 28691−28699.
53. Gorman, J. J., Wallis, T. P., and Pitt, J. J. (2002) Protein disulfide bond determination by mass spectrometry, *Mass Spectrom. Rev. 21*, 183−216.
54. Jones, M. D., Patterson, S. D., and Lu, H. S. (1998) Determination of disulfide bonds in highly bridged disulfide-linked peptides by matrix-assisted laser desorption/ionization mass spectrometry with postsource decay, *Anal. Chem. 70*, 136−143.
55. Patterson, S. D., and Katta, V. (1994) Prompt fragmentation of disulfide-linked peptides during matrix-assisted laser desorption ionization mass spectrometry, *Anal. Chem. 66*, 3727−3732.
56. Yonezawa, N., and Nakano, M. (2003) Identification of the carboxyl termini of porcine zona pellucida glycoproteins ZPB and ZPC, *Biochem. Biophys. Res. Commun. 307*, 877−882.
57. Rankin, T. L., O'Brien, M., Lee, E., Wigglesworth, K., Eppig, J., and Dean, J. (2001) Defective zonae pellucidae in Zp2-null mice disrupt folliculogenesis, fertility and development, *Development 128*, 1119−1126.
58. Rankin, T., Familiari, M., Lee, E., Dwyer, N., Blanchette-Mackie, J., Darago, J., and Dean, J. (1996) Mice homozygous for an insertional mutation in the ZP3 gene lack a zona pellucida and are infertile, *Development 122*, 2903−2910.
59. Liu, C., Litscher, E. S., Mortillo, S., Sakai, Y., Kinloch, R. A., Stewart, C. L., and Wassarman, P. M. (1996) Targeted disruption of the mZP3 gene results in production of eggs lacking a zona pellucida and infertility in female mice, *Proc. Natl. Acad. Sci., U.S.A. 93*, 5431−5436.
60. Greve, J. M., and Wassarman, P. M. (1985) Mouse egg extracellular coat is a matrix of interconnected filaments possessing a structural repeat, *J. Mol. Biol. 181*, 253−264.
61. Goodyear, R. J., and Richardson, G. P. (2002) Extracellular matrices associated with the apical surfaces of sensory epithelia in the inner ear: molecular and structural diversity, *J. Neurobiol. 53*, 212−227.
62. Legan, P. K., Rau, A., Keen, J. N., and Richardson, G. P. (1997) The mouse tectorins. Modular matrix proteins of the inner ear homologous to components of the sperm-egg adhesion system, *J. Biol. Chem. 272*, 8791−8801.
63. Legan, P. K., Lukashkina, V. A., Goodyear, R. J., Kossi, M., Russell, I. J., and Richardson, G. P. (2000) A targeted deletion in alpha-tectorin reveals that the tectorial membrane is required for the gain and timing of cochlear feedback, *Neuron 28*, 273−285.
64. Kubo, H., Matsushita, M., Kotani, M., Kawasaki, H., Saido, T. C., Kawashima, S., Katagiri, C., and Suzuki, A. (1999) Molecular basis for oviductin-mediated processing from gp43 to gp41, the predominant glycoproteins of Xenopus egg envelopes, *Dev. Genet. 25*, 123−129.
65. Litscher, E. S., Qi, H., and Wassarman, P. M. (1999) Mouse zona pellucida glycoproteins mZP2 and mZP3 undergo carboxy-terminal proteolytic processing in growing oocytes, *Biochemistry 38*, 12280−12287.
66. Sasanami, T., Pan, J., Doi, Y., Hisada, M., Kohsaka, T., and Toriyama, M. (2002) Secretion of egg envelope protein ZPC after C-terminal proteolytic processing in quail granulosa cells, *Eur. J. Biochem. 269*, 2223−2231.
67. Siegfried, G., Khatib, A. M., Benjannet, S., Chretien, M., and Seidah, N. G. (2003) The proteolytic processing of pro-platelet-derived growth factor-A at RRKR(86) by members of the proprotein convertase family is functionally correlated to platelet-derived growth factor-A-induced functions and tumorigenicity, *Cancer Res. 63*, 1458−1463.
68. Barisone, G. A., Albertali, I. E., Sanchez, M., and Cabada, M. O. (2003) The envelopes of amphibian oocytes: physiological modifications in Bufo arenarum, *Reprod. Biol. Endocrinol. 1*, 18.
69. Scapigliati, G., Meloni, S., and Mazzini, M. (1999) A monoclonal antibody against chorion proteins of the sea bass Dicentrarchus labrax (Linnaeus, 1758): studies of chorion precursors and applicability in immunoassays, *Biol. Reprod. 60*, 783−789.
70. Waclawek, M., Foisner, R., Nimpf, J., and Schneider, W. J. (1998) The chicken homologue of zona pellucida protein-3 is synthesized by granulosa cells, *Biol. Reprod. 59*, 1230−1239.
71. Fujita, T., Shimizu, M., Hiramatsu, N., Fukada, H., and Hara, A. (2002) Purification of serum precursor proteins to vitelline envelope (choriogenins) in masu salmon, Oncorhynchus masou, *Comp. Biochem. Physiol. B: Biochem. Mol. Biol. 132*, 599−610.
72. Wassarman, P. M., and Mortillo, S. (1991) Structure of the mouse egg extracellular coat, the zona pellucida, *Int. Rev. Cytol. 130*, 85−110.

BI0495937