Comparative Structures of the Apopolysialoglycoproteins from Unfertilized and Fertilized Eggs of Salmonid Fishes†

Ken Kitajima,‡ Hiroyuki Sorimachi,‡ Sadako Inoue,*,§ and Yasuo Inoue*,‡

Department of Biophysics and Biochemistry, Faculty of Science, University of Tokyo, Hongo-7, Bunkyo-ku, Tokyo 113, Japan, and School of Pharmaceutical Sciences, Showa University, Hatanodai-1, Shinagawa-ku, Tokyo 142, Japan

Received February 26, 1988; Revised Manuscript Received May 19, 1988

ABSTRACT: The complete amino acid sequence of the major polysialoglycoproteins (PSGPs) from two genera of salmonid fish eggs, *Salvelinus* and *Oncorhynchus*, has been determined. The occurrence of tandem repeats of a genus-specific dodeca- and tridecapeptide was found for the apoPSGP of *Salvelinus leucomaenis pluvius* (Slp) and *Oncorhynchus masou ishikawai* (Omi), respectively, their amino acid sequences being highly homologous with that of rainbow trout [*Salmo gairdneri* (Sg)] apoPSGP (* denotes the glycosylation site; $\langle N \rangle = \sim 25$):

H-PSGP(Slp): (Asp-Asp-Ala-Thr*-Ser*-Glu-Ala-Ala-Thr*-Gly-Pro-Ser-)$_N$
H-PSGP(Omi): (Asp-Asp-Ala-Thr*-Ser*-Glu-Ala-Ala-Thr*-Gly-Pro-Ser-Ser-)$_N$
H-PSGP(Sg): (Asp-Asp-Ala-Thr*-Ser*-Glu-Ala-Ala-Thr*-Gly-Pro-Ser-Gly-)$_N$

Within 5–7 min following fertilization H-PSGP is converted to the low-molecular-mass PSGP (L-PSGP) by a specific protease (PSGPase). We have purified L-PSGP from the fertilized eggs of *S. leucomaenis pluvius* and *Oncorhynchus keta* (chum salmon) and compared it with rainbow trout egg L-PSGP(Sg) by analysis of their amino acid sequence:

L-PSGP(Slp): Asp-Ala-Thr*-Ser*-Glu-Ala-Ala-Thr*-Gly-Pro-Ser-Asp
L-PSGP(Ok): Asp-Asp-Ala-Thr*-Ser*-Glu-Ala-Ala-Thr*-Gly-Pro-Ser-Ser
L-PSGP(Sg): Asp-Asp-Ala-Thr*-Ser*-Glu-Ala-Ala-Thr*-Gly-Pro-Ser-Gly

The data support the conclusion that H-PSGP is degraded in vivo 5–7 min after fertilization to L-PSGP by proteolytic cleavage at the position two residues C-terminally to the Pro residue, i.e., -Pro-Ser-Xaa↓Asp- (Xaa = either Gly, Ser, or Asp) by the action of PSGPase.

Polysialoglycoprotein (PSGP)[1] was discovered in the unfertilized eggs of rainbow trout by Inoue and Iwasaki (1978). Since then homologous high-molecular-mass PSGPs (H-PSGP; molecular mass ~200 kDa) have also been purified from other salmonid fish eggs (Shimamura et al., 1983, 1984; Iwasaki et al., 1985; Iwasaki & Inoue, 1985), and thus the widespread distribution of PSGP was demonstrated by examination of eight different species from three genera of Salmonidae. We initiated the structural studies of H-PSGP from different species with initial emphasis on their carbohydrate moieties as these accounted for the major portion of the molecules' mass and vast abundance of sialic acid content (Shimamura et al., 1983, 1985; Iwasaki & Inoue, 1985; Inoue & Matsumura, 1979, 1980; Inoue & Iwasaki, 1980; Inoue et al., 1981, 1982; Iwasaki et al., 1984a,b, 1987; Nomoto et al., 1982; Kitajima et al., 1984; Nadano et al., 1986). Our previous studies revealed that (1) H-PSGP is a major component of cortical alveoli (Inoue & Inoue, 1986; Inoue et al., 1987), (2) it undergoes depolymerization (200 to 9 kDa) upon fertilization (Inoue & Inoue, 1986), and (3) the H-PSGP molecule is made up of about 25 tandem repeats of a glycotridecapeptide, and within 5–7 min postfertilization H-PSGP is converted to the repeating unit by a specific protease (Inoue & Inoue, 1986; Kitajima et al., 1986).

This report details the amino acid sequence of major H- and L-PSGP from Salmonidae fish eggs of two other genera, *Salvelinus* and *Oncorhynchus*, and compares it with that of rainbow trout PSGP (genus *Salmo*). Although apoPSGPs are shown to be homologous, the definite intergenus variation is demonstrated. The knowledge about structures of L-PSGP provides an insight into the specificity of a unique protease involved in the conversion of H-PSGP to L-PSGP.

Materials and Methods

Fish Eggs. Both fertilized and unfertilized ovulated eggs were supplied by courtesy of the Gunma Prefectural Fisheries Experimental Stations at Kawaba (*Salvelinus leucomaenis pluvius*), Hakoshima (*Oncorhynchus masou ishikawai*), and the Hokkaido Salmon Hatchery, Chitose Branch (*Onco-*

† This work was supported in part by Grant-in-Aids 63490018 (to S.I.) and 63470135 (to Y.I.) for Scientific Research from the Ministry of Education, Science, and Culture of Japan.

‡ University of Tokyo.
§ Showa University.

[1] Abbreviations: PSGP, polysialoglycoprotein; H-PSGP, high-molecular-mass PSGP isolated from unfertilized eggs; L-PSGP, low-molecular-mass PSGP obtained from fertilized eggs; st-H-PSGP and st-L-PSGP, *Arthrobacter* sialidase treated H- and L-PSGP, respectively; PSGP(Slp), PSGP(Omi), PSGP(Ok), and PSGP(Sg), PSGP isolated from *Salvelinus leucomaenis pluvius*, *Oncorhynchus masou ishikawai*, *Oncorhynchus keta*, and *Salmo gairdneri*, respectively; Slp-V8-A and Slp-V8-B, *S. leucomaenis pluvius* st-L-PSGP derived V8 proteolytic glycopeptide fractions; Ok-V8-A and Ok-V8-B, glycopeptide fragments obtained by V8 protease digestion of *O. keta* st-L-PSGP after Smith degradation; V8-GP, major glycopeptide fraction obtained by V8 protease digestion of st-H-PSGP when fractionated on Sephadex G-50 column [e.g., V8-GP(Slp) and V8-GP(Omi) represent the corresponding glycopeptides derived from the unfertilized eggs of *S. leucomaenis pluvius* and *O. masou ishikawai*, respectively]; Sia, sialic acid(s); apoPSGP, the core protein of PSGP; PSGPase, polysialoglycoproteinase, an enzyme responsible for degradation of H-PSGP into L-PSGP; TLC, thin-layer chromatography; dansyl, 5-(dimethylamino)naphthalene-1-sulfonyl; PTH, phenylthiohydantoin; TBA, thiobarbituric acid; Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride.

0006-2960/88/0427-7141$01.50/0

rhynchus keta). Eggs were washed with 0.8% NaCl and stored at −30 °C until use.

Preparation of H-PSGP from the Unfertilized Eggs of S. leucomaenis pluvius, O. masou ishikawai, and O. keta and Sialidase Digestion of H-PSGP. H-PSGP samples were prepared from the unfertilized eggs as described previously (Inoue & Matsumura, 1979), and 160 mg (11.2 μmol) of each of the H-PSGPs was digested with *Arthrobacter ureafaciens* sialidase (Nakarai Chemicals, Kyoto) in 50 mM sodium acetate buffer (pH 5.5) for 5 days at 37 °C [the enzyme (total 10 units) was added at 12-h intervals]. The hydrolysate was separately applied to a Sephacryl S-200 column (2.3 × 163 cm) and eluted with 0.1 M NaCl. The elution patterns were monitored by absorbance at 215 nm, by the phenol–sulfuric acid method (Dubois et al., 1956) (at 490 nm) for neutral sugar, and by the thiobarbituric acid method (Aminoff, 1961) (TBA assay; at 549 nm) with modification (Uchida et al., 1977) for sialic acid. For the TBA assay aliquots were treated with 0.1 N trifluoroacetic acid for 10 min at 100 °C. Sialidase-treated H-PSGP (st-H-PSGP) thus obtained was desalted by dialysis.

Staphylococcus aureus V8 Protease Digestion of st-H-PSGP. st-H-PSGP (7.9 mg = 2.4 μmol each) derived from *S. leucomaenis pluvius* and *O. masou ishikawai* unfertilized eggs, dissolved in 800 μL of 50 mM NH_4HCO_3, pH 7.8, was digested for 20 h at 37 °C with 420 μg (or 240 units) of *S. aureus* V8 protease (Miles Laboratories). The digest was purified by applying to a column (1.6 × 150 cm) of Sephadex G-50 and eluting with 0.1 M NaCl. The elution profile was determined by absorbance at 210 nm and by neutral sugar assay. Glycopeptides were desalted by passage through a Sephadex G-25 column (1.6 × 108 cm; eluate, 5% ethanol).

Amino Acid and Carbohydrate Analysis. Amino acid and carbohydrate compositions of glycoproteins and glycopeptides were determined as described previously [e.g., Kitajima et al. (1986)].

N-Terminal Analysis. NH_2-terminal residues of glycoproteins and glycopeptides (2 nmol each) were identified by the dansyl method of Gray (1972), using TLC on polyamide thin-layer sheets (Woods & Wang, 1967).

Amino Acid Sequence Determination. Glycopeptides were sequenced by gas-phase Edman degradation with the aid of an Applied Biosystems Model 470A protein sequencer. The phenylthiohydantoin derivatives of amino acids were identified by high-pressure liquid chromatography using a Spectra-Physics SP8750 and quantified by a Spectra-Physics SP4270 integrator.

Pronase Digestion of S. leucomaenis pluvius and O. masou ishikawai st-H-PSGP. st-H-PSGP (3.3 mg = 1 μmol each) was dissolved in 500 μL of 0.1 M Tris-HCl buffer (pH 8.0)/10 mM $CaCl_2$ and digested with pronase P (Kaken Kagaku, Tokyo); the mixture was incubated for a total of 6 days at 37 °C under toluene. Two hundred forty micrograms of the enzyme (total 720 μg) was added stepwise at 48-h intervals. Following acidification with trichloroacetic acid to a final concentration of 5%, the digest was centrifuged for 10 min at 2500 rpm at 0 °C. The supernatant fraction was applied, after neutralization with an equivalent amount of NaOH, on a Bio-Gel P-4 column (1.3 × 140 cm). The column was eluted with 0.1 M NaCl and fractions were monitored by absorbance at 210 nm.

Preparation of L-PSGP from the Fertilized Eggs of S. leucomaenis pluvius and O. keta and Sialidase Digestion of L-PSGP(Slp) and L-PSGP(Ok). Isolation and purification of L-PSGP from the fertilized eggs (30 min postinsemination) of the respective fish species followed closely those described previously (Inoue & Inoue, 1986). In order to facilitate the structural analysis of the protein core, L-PSGP (40 mg = 2.8 μmol each), dissolved in 50 mM sodium acetate buffer (pH 5.5), was subjected to exhaustive digestion with *A. ureafaciens* sialidase by stepwise addition of a total of 6.6 units under toluene for 5.5 days at 37 °C. The digest was applied to a Sephadex G-50 column (1.6 × 150 cm) and eluted with 0.1 M NaCl. Fractions were monitored by absorbance at 215 nm, by the phenol–sulfuric acid method, and by the thiobarbituric acid method for the residual sialidase-resistant sialic acid residues present in the sialidase-treated L-PSGP. The glycoprotein material was pooled, desalted, and designated as st-L-PSGP and more specifically st-L-PSGP(Slp) or st-L-PSGP(Ok).

Smith Degradation of st-L-PSGP(Slp) and st-L-PSGP(Ok). The reaction described by Spiro (1966) was used with some modification. st-L-PSGP (3.3 mg = 1 μmol each) was dissolved in 3.2 mL of 50 mM sodium acetate containing 12.5 mM sodium periodate, pH 4.7. The solution was kept in the dark at 4 °C for 60 h during which the consumption of periodate was determined by using violet reagent (Avigad, 1969). The oxidized st-L-PSGP was reduced with 1.7 mL of 0.5 M $NaBH_4$/0.5 M sodium borate buffer (pH 8.0) for 12 h at 0 °C in the dark, acidified with a few drops of 1 M CH_3COOH, and desalted on a Sephadex G-25 column (vide supra). After drying, the sample was subjected to mild acid hydrolysis with 2 mL of 0.05 N H_2SO_4 (25 mM) for 1 h at 80 °C. The solution was neutralized with an equivalent amount of NaOH and passed through a column (1.3 × 140 cm) of Bio-Gel P-4. These treatments afforded an extensively deglycosylated st-L-PSGP in which three residues of the formerly 3-substituted proximal GalNAc remained undegraded.

S. aureus V8 Protease Digestion of the Smith Degradation Product of st-L-PSGP. Partially deglycosylated st-L-PSGP (800 μg = 420 nmol each) was digested with V8 protease as described above.

Ion-Exchange Chromatography of V8 Protease Digests of the Partially Deglycosylated st-L-PSGP Derived from the Fertilized Eggs of S. leucomaenis pluvius and O. keta (Pacific Salmon). (i) S. leucomaenis pluvius st-L-PSGP Derived V8 Proteolytic Products. The V8 protease digest of the partially deglycosylated st-L-PSGP (800 μg) was applied to a Dowex 1 × 2 column (0.7 × 8.5 cm; preequilibrated with 20 mM Tris-HCl, pH 7.0). The column was washed with 15 mL of 20 mM Tris-HCl (pH 7.0) and then eluted with 120 mL of 0–0.4 M NaCl linear gradient in the same buffer to yield two distinctive major peaks designated Slp-V8-A and *Slp*-V8-B.

(ii) O. keta st-L-PSGP Derived V8 Proteolytic Products. The V8 protease digest of the Pacific salmon egg st-L-PSGP Smith degradation product was likewise subjected to DEAE-Sephadex A-25 chromatography. The pass-through fraction and the fraction eluted with 0.75 M NaCl were collected and designated as Ok-V8-A and Ok-V8-B, respectively. Both of these fractions were found to be homogeneous.

Results

st-H-PSGP Derived from the Unfertilized Eggs of O. masou ishikawai and S. leucomaenis pluvius, st-H-PSGP(Omi) and st-H-PSGP(Slp). Exhaustive digestion of H-PSGP(Omi) and H-PSGP(Slp) with *A. ureafaciens* sialidase produced extensively desialylated H-PSGPs, free sialic acid, and a small amount of 2-keto-3-deoxy-D-*glycero*-D-*galacto*-nononic acid (KDN) capped oligosialic acids (Nadano et al., 1986). When the sialidase digests of H-PSGP(Omi) and H-PSGP(Slp) were chromatographed on Sephacryl S-200, the corresponding st-

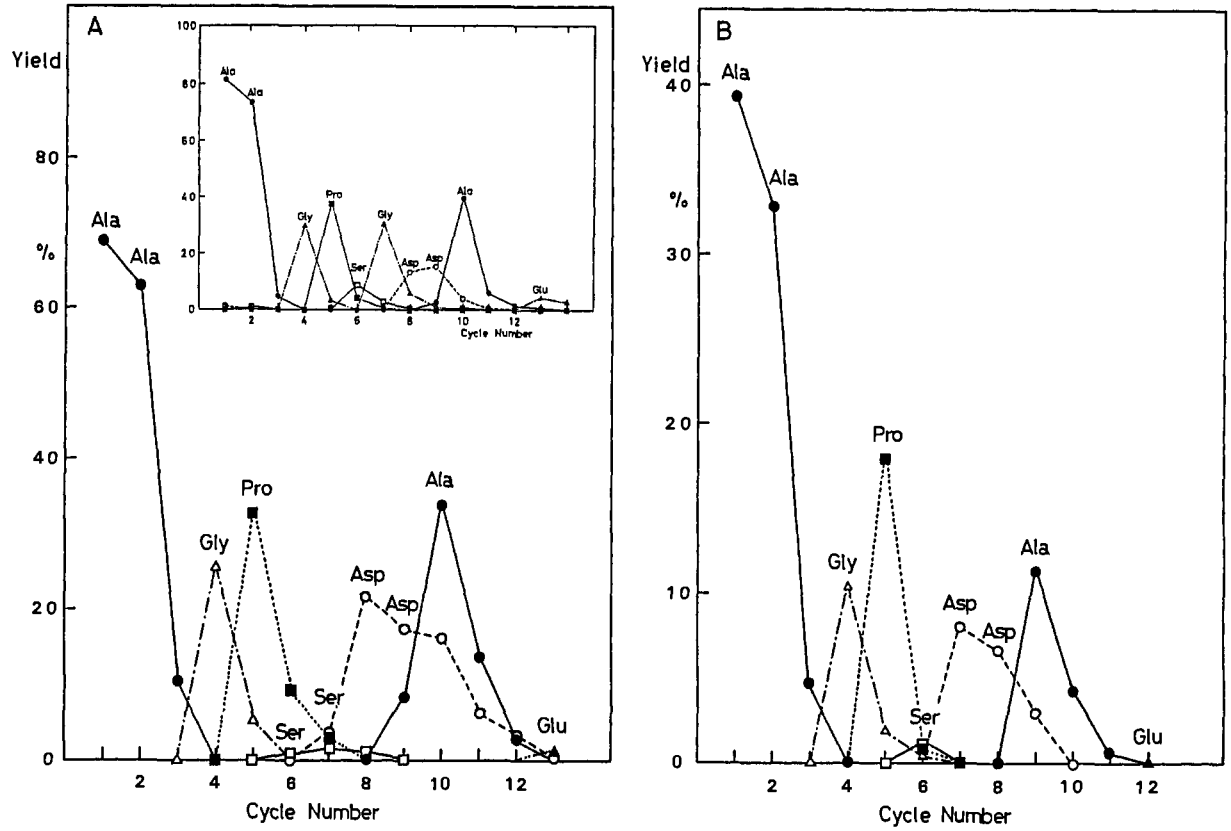


FIGURE 1: Automated Edman degradation of the V8 proteolytic fragments (A) V8-GP(Omi) and (B) V8-GP(Slp). Edman degradation yields are expressed as percentages relative to the standard PTH derivatives determined by high-pressure liquid chromatography. Inset: Corresponding data for V8-GP(Sg) (Kitajima et al., 1986).

H-PSGP(Omi) and st-H-PSGP(Slp) in which 94% of the total NeuGc was liberated were eluted shortly after the void volume (data[2] not shown). Amino acid and carbohydrate analysis of the st-H-PSGP(Omi) and st-H-PSGP(Slp) showed that the composition was nearly identical with that of the original H-PSGP(Omi) and H-PSGP(Slp) material, except for the expected loss of sialic acid (data[2] not shown).

Digestion of st-H-PSGP(Omi) and st-H-PSGP(Slp) with S. aureus V8 Protease and Amino Acid Sequencing of the Proteolytic Glycopeptides V8-GP(Omi) and V8-GP(Slp). Samples of st-H-PSGP(Omi) and st-H-PSGP(Slp) were digested with V8 protease, and the digests were chromatographed on the Sephadex G-50 column. The elution profiles[2] showed one major broad carbohydrate-containing peak centering at K_{av} = 0.44. A small peak appearing ahead of the major broad peak was devoid of carbohydrate and identified as V8 protease used. Late-eluting fractions of the major peak, on the basis of amino acid composition analysis, sequence data, and comparison with the corresponding results for rainbow trout egg PSGP (Kitajima et al., 1986), were shown to contain the COOH-terminal fragment of the st-H-PSGP as the major component. The major portion of the carbohydrate-containing peak was found to be the major proteolytic product and designated as V8-GP. Recovery, on the basis of the neutral sugar content, amounted to more than 90% of that in the starting material, st-H-PSGP.

The amino acid composition of V8-GP(Omi) and V8-GP(Slp) is essentially identical with that of st-H-PSGP(Omi) and st-H-PSGP(Slp), respectively, and the NH_2-terminal amino acid was identified as Ala (data[2] not shown). Taken together, the results strongly indicate that both *O. masou ishikawai* and *S. leucomaenis pluvius* PSGP consist of "repetitive" units along the peptide backbone as was found for rainbow trout egg PSGP (Kitajima et al., 1986).

V8-GP(Omi) and V8-GP(Slp) were thus subjected to automated amino acid sequence analysis, and the results are shown in Figure 1. These sequence data allowed us to determine the *dodecapeptide* and *tridecapeptide* core structures for the major, if not all, components in V8-GP(Slp) and V8-GP(Omi), respectively:

	1	2	3	4	5	6	7	8	9	10	11	12	13
V8-GP(Slp):	Ala	Ala	Xaa	Gly	Pro	Ser	Asp	Asp	Ala	Xaa	Xaa	Glu	
V8-GP(Omi):	Ala	Ala	Xaa	Gly	Pro	Ser	Ser	Asp	Asp	Ala	Xaa	Xaa	Glu

At cycles marked Xaa, no phenylthiohydantoin amino acids (PTH-aa) were identified, but comparison of the amino acid composition with the amino acids identified during the automated Edman degradation indicates that the residues at these positions are likely to be either Ser or Thr with O-linked glycosyl chain attached. In order to allocate the glycosylated Ser and Thr residues to the proper positions, we have conducted Pronae digestion of st-H-PSGP(Omi) and st-H-PSGP(Slp) and analyzed the Pronase-derived glycopeptide products. As in the case with st-H-PSGP(rainbow trout) [Inoue & Inoue, 1986; Kitajima et al., 1986; see also Shimamura et al. (1985) for glycopeptides derived from Pacific salmon egg H-PSGP], the glycopeptides obtained were essentially assorted into the two distinct groups Thr*-Ser*-Glu and Thr*-Gly-Pro-Ser (data not shown). Consequently, the following structures are determined for V8-GP(Slp) and V8-GP(Omi):

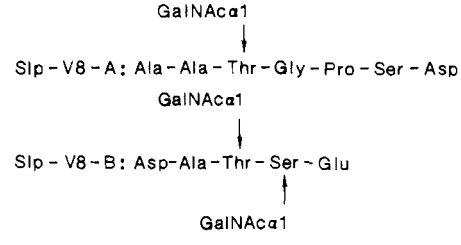
	1	2	3	4	5	6	7	8	9	10	11	12	13
V8-GP(Slp):	Ala	Ala	Thr*	Gly	Pro	Ser	Asp	Asp	Ala	Thr*	Ser*	Glu	
V8-GP(Omi):	Ala	Ala	Thr*	Gly	Pro	Ser	Ser	Asp	Asp	Ala-	Thr*	Ser*	Glu

[2] These data were submitted to the scrutiny of the reviewers and will be furnished to the interested reader by writing directly to the authors.

Sialidase Digestion of L-PSGP from Fertilized (30-min) Eggs of O. keta (Pacific Salmon) and S. leucomaenis pluvius. The sialidase digests of L-PSGP from the fertilized eggs of Pacific salmon and *S. leucomaenis pluvius* were chromatographed on Sephadex G-50 in 0.1 M NaCl in order to separate st-L-PSGP(Ok) and st-L-PSGP(Slp) from NeuGc and KDN–sialyloligomers (data[2] not shown). Both st-L-PSGP(Ok) and st-L-PSGP(Slp) had a K_{av} of 0.47, almost identical with that of st-L-PSGP(Sg). The average molecular weight of st-L-PSGP(Ok) was estimated to be about 3300 from the K_{av} value. The analytical data for st-L-PSGP(Ok) and st-L-PSGP(Slp) again showed that the composition was substantially the same as that of the original L-PSGP(Ok) and L-PSGP(Slp), respectively, obviously indicating the peptide moieties remain unchanged (more than 90% of sialic acid was removed from L-PSGP material) (data[2] not shown). There was a difference in Ser content of the L-PSGP from the two fish species as was found for the H-PSGP from *O. masou ishikawai* and *S. leucomaenis pluvius*.

Smith Degradation of st-L-PSGP(Ok) and st-L-PSGP(Slp). Preceding the analysis of L-PSGP peptide structure, the glycopeptides were deglycosylated by Smith degradation in order to eliminate extensive heterogeneity due to structural variation of the carbohydrate moieties. The Smith degradation products of Pacific salmon st-L-PSGP and st-L-PSGP(Slp) gave only one sharp peak when chromatographed on Bio-Gel P-4 (data not shown), indicating that the reaction proceeded uniformly to give a single product. The chemical compositions of Smith degradation products of st-L-PSGP(Ok) and st-L-PSGP(Slp) are Asp/Thr/Ser/Glu/Pro/Gly/Ala/GalNAc = 2.0/1.8/2.6/1.0/1.0/1.1/2.8/3.5 and Asp/Thr/Ser/Glu/Pro/Gly/Ala/GalNAc = 1.9/1.8/2.1/1.0/1.1/1.1/3.4/3.0. The data show that, with the exception of the three proximal GalNAc residues directly O-linked to one Ser and two Thr residues, almost all other carbohydrate residues were removed. This is in complete agreement with the expectation that other than the proximal GalNAc residues substituted at position 3 complete removal of the monosaccharide residues could be achieved if one considers the five different types of oligosaccharide structures established for PSGP (Iwasaki & Inoue, 1985; Inoue & Inoue, 1986; Kitajima et al., 1986; Nadano et al., 1986; Inoue et al., 1987; Iwasaki et al., 1987). Smith degradation used in this study was also shown to cleave most of the sugars from st-L-PSGP, leaving the peptide backbone intact.

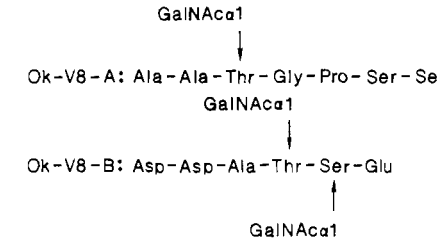
Glycopeptides Fragments Formed by V8 Proteolysis of the Mostly Deglycosylated st-L-PSGP(Ok) and st-L-PSGP(Slp). (*i*) *S. leucomaenis pluvius st-L-PSGP Derived V8 Proteolytic Products.* The extensively deglycosylated st-L-PSGP(Slp) was treated with V8 protease, and the digest was fractionated by chromatography on Dowex 1 × 2 to yield two major distinct peaks, designated Slp-V8-A (yield 90%) and Slp-V8-B (yield 63%), and the amino acid compositions are Asp/Thr/Ser/Pro/Gly/Ala/GalNAc = 1.1/1.0/1.0/1.0/1.0/2.1/1.2 for Slp-V8-A and Asp/Thr/Ser/Glu/Ala/GalNAc = 1.0/1.0/1.0/1.2/1.0/2.3 for Slp-V8-B. The amino acid sequence of these glycopeptides was determined:



These results demonstrate that L-PSGP of *S. leucomaenis pluvius* is predominantly composed of the following glycododecapeptide:

L-PSGP(Slp):
Asp-Ala-Thr*-Ser*-Glu-Ala-Ala-Thr*-Gly-Pro-Ser-Asp

(*ii*) *O. keta st-L-PSGP Derived V8 Proteolytic Products.* The chemical compositions of Ok-V8-A (yield 61%) and Ok-V8-B (yield 81%) (see Materials and Methods) are Thr/Ser/Pro/Gly/Ala/GalNAc = 1.0/1.8/1.0/1.2/2.0/1.1 for Ok-V8-A and Asp/Thr/Ser/Glu/Gly/Ala/GalNAc = 2.1/1.1/1.2/1.0/0.26/1.3/2.2 for Ok-V8-B. These glycopeptides were found to be homogeneous enough for structural analysis, so they were subjected to automated Edman degradation without further purification. The entire amino acid sequences of Ok-V8-A and Ok-V8-B glycopeptides determined are shown:



These results are consistent with the conclusion that *O. keta* L-PSGP is composed of the glycotridecapeptide whose structure is given:

L-PSGP(Ok): Asp-Asp-Ala-Thr*-Ser*-Glu-Ala-Ala-Thr*-Gly-Pro-Ser-Ser

Discussion

The ubiquitous presence of PSGP in the eggs of eight fish species of Salmonidae studied (Iwasaki et al., 1985) suggests that PSGP may constitute a new glycoprotein family although their function in vivo is only beginning to be understood. It seemed important to determine how homologous the structure is seen for PSGP molecules from a variety of salmonid fish eggs. We have therefore studied the chemical structure of H-PSGP and L-PSGP from three different genera of Salmonidae fish, i.e., *Oncorhynchus*, *Salvelinus*, and *Salmo*.

H-PSGP molecules have been confirmed to be made up of a tandem repeat of a glycododeca- or glycotridecapeptide unit as previously shown for the rainbow trout egg H-PSGP (Kitajima et al., 1986). Amino acid sequences of the repeating units were readily obtained from st-H-PSGP digest by using staphylococcal V8 protease (specific for the C-terminal side of Glu residues). The resulting V8 proteolytic glycopeptides from different fish species were found to have striking homology to the reported sequence of rainbow trout PSGP (Kitajima et al., 1986) to allow alignments along the peptide chain. Fragmentary glycopeptides derived from cleavages with Pronase were able to provide the requisite allocation of all the glycosylated Ser and Thr residues and allow the unequivocal assignment of a continuous sequence of apoH-PSGP. The amino acid sequences thus determined for major H-PSGP isolated from the unfertilized eggs of three different genera, *Oncorhynchus*, *Salmo*, and *Salvelinus*, of salmonid fish are

	1 2 3 4 5 6 7 8 9 10 11 12 13
	S. leucomaenis pluvius
H-PSGP(Slp):	(-Asp-Asp-Ala-Thr*-Ser*-Glu-Ala-Ala-Thr*-Gly-Pro-Ser- -)$_N$
	S. gairdneri
H-PSGP(Sg):	(-Asp-Asp-Ala-Thr*-Ser*-Glu-Ala-Ala-Thr*-Gly-Pro-Ser-Gly-)$_N$
	O. masou ishikawai
H-PSGP(Omi):	(-Asp-Asp-Ala-Thr*-Ser*-Glu-Ala-Ala-Thr*-Gly-Pro-Ser-Ser-)$_N$
	O. keta
H-PSGP(Ok):	(-Asp-Asp-Ala-Thr*-Ser*-Glu-Ala-Ala-Thr*-Gly-Pro-Ser-Ser-)$_N$

Different apoH-PSGP molecules were selectively expressed in different fish species. Aside from such differences noted, the similarity in the amino acid sequences of H-PSGP from Salmonidae fish species is striking, and the residues other than at positions 13 in the residue numbering for the repeating units are conserved between species. The amino acid sequence data obtained here and in the previous study (Inoue & Inoue, 1986; Kitajima et al., 1986) suggest that H-PSGPs from Salmonidae fish eggs belong to a family of structurally related apoproteins that evolved from a common ancestral gene and apoH-PSGP has evolved through a series of gene duplications.

In a number of molecular variants of H-PSGP studied, only specific hydroxyamino acid residues marked by asterisks (*) are glycosylated. The dodeca- or tridecapeptide repeat unit is presumably important for how the particular serine and all the threonine residues become substituted with oligo(poly)-sialylglycans. In all the H-PSGP studied, the glycosylated Ser residues are preceded by the Asp-Asp-Ala-Thr tetrapeptide sequence and followed by the Glu-Ala-Ala-Thr sequence. The Ser residues unglycosylated are those preceded at the position adjacent to and two residues N-terminally by a proline residue. At present it is unclear how these particular Ser residues in H-PSGP are never substituted with sialylglycan chains.

Comparison of the sequences of L-PSGP from the fertilized eggs of a variety of species indicates that the sequences and residues implicated here as being involved in recognition and cleavage of H-PSGP by polysialoglycoproteinase (PSGPase) upon fertilization seem to be -Pro-Ser-*Xaa*$\downarrow$*Asp*-Asp- (Xaa = Gly or Ser) and -Pro-Ser-*Asp*$\downarrow$*Asp*. Work is currently under way to isolate PSGPase and to define its catalytic specificity more precisely.

Acknowledgments

We thank Professor M. Tomita, Showa University, who generously made available an Applied Biophysics Model 470A protein sequencer, for his help and advice. We thank Dr. F. Hayashi, the former head of the Gunma Prefectural Fisheries Experimental Station at Kawaba, for his cooperation. We are indebted to Mariko Iwasaki for providing the *S. leucomaenis pluvius* H-PSGP used in this study.

References

Aminoff, D. (1961) *Biochem. J. 81*, 384–392.
Avigad, G. (1969) *Carbohydr. Res. 11*, 119–123.
Dubois, M., Gilles, K. A., Hamiltom, J. K., Rebers, P. A., & Smith, F. (1956) *Anal. Chem. 28*, 350–356.
Gray, W. R. (1972) *Methods Enzymol. 25*, 121–138.
Inoue, S., & Iwasaki, M. (1978) *Biochem. Biophys. Res. Commun. 83*, 1018–1023.
Inoue, S., & Matsumura, G. (1979) *Carbohydr. Res. 74*, 361–368.
Inoue, S., & Iwasaki, M. (1980) *Biochem. Biophys. Res. Commun. 93*, 162–165.
Inoue, S., & Matsumura, G. (1980) *FEBS Lett. 121*, 33–36.
Inoue, S., & Inoue, Y. (1986) *J. Biol. Chem. 261*, 5256–5261.
Inoue, S., Iwasaki, M., & Matsumura, G. (1981) *Biochem. Biophys. Res. Commun. 102*, 1295–1301.
Inoue, S., Matsumura, G., & Inoue, Y. (1982) *Anal. Biochem. 125*, 118–124.
Inoue, S., Kitajima, K., Inoue, Y., & Kudo, S. (1987) *Dev. Biol. 123*, 442–454.
Iwasaki, M., & Inoue, S. (1985) *Glycoconjugate J. 2*, 209–228.
Iwasaki, M., Inoue, S., Kitajima, K., Nomoto, H., & Inoue, Y. (1984a) *Biochemistry 23*, 305–310.
Iwasaki, M., Nomoto, H., Kitajima, K., Inoue, S., & Inoue, Y. (1984b) *Biochem. Int. 8*, 573–579.
Iwasaki, M., Inoue, S., Tazawa, I., & Inoue, Y. (1985) *Proceedings of the International Symposium on Glycoconjugates, 8th* (Davidson, E. A., Williams, J. C., & Di Ferrante, N. M., Eds) pp 51–52, Praeger, New York.
Iwasaki, M., Inoue, S., Nadano, D., & Inoue, Y. (1987) *Biochemistry 26*, 1452–1457.
Kitajima, K., Nomoto, H., Inoue, Y., Iwasaki, M., & Inoue, S. (1984) *Biochemistry 23*, 310–316.
Kitajima, K., Inoue, Y., & Inoue, S. (1986) *J. Biol. Chem. 261*, 5262–5269.
Nadano, D., Iwasaki, M., Endo, S., Kitajima, K., Inoue, S., & Inoue, Y. (1986) *J. Biol. Chem. 261*, 11550–11557.
Nomoto, H., Iwasaki, M., Endo, T., Inoue, S., Inoue, Y., & Matsumura, G. (1982) *Arch. Biochem. Biophys. 218*, 335–341.
Shimamura, M., Endo, T., Inoue, Y., & Inoue, S. (1983) *Biochemistry 22*, 959–963.
Shimamura, M., Endo, T., Inoue, Y., Inoue, S., & Kambara, H. (1984) *Biochemistry 23*, 317–322.
Shimamura, M., Inoue, Y., & Inoue, S. (1985) *Biochemistry 24*, 5470–5480.
Spiro, R. G. (1966) *Methods Enzymol. 8*, 26–52.
Uchida, Y., Tsukada, Y., & Sugimori, T. (1977) *J. Biochem. (Tokyo) 82*, 1425–1433.
Woods, K. R., & Wang, K.-T. (1967) *Biochim. Biophys. Acta 133*, 369–370.