Porcine Dentin Sialophosphoprotein

LENGTH POLYMORPHISMS, GLYCOSYLATION, PHOSPHORYLATION, AND STABILITY*

Received for publication, January 24, 2008, and in revised form, March 13, 2008. Published, JBC Papers in Press, March 20, 2008, DOI 10.1074/jbc.M800633200

Yasu Yamakoshi†, Yuhe Lu†, Jan C.-C. Hu†, Jung-Wook Kim†, Takanori Iwata‡, Kazuyuki Kobayashi‡, Takatoshi Nagano‡‡, Fumiko Yamakoshi‡, Yuanyuan Hu‡, Makoto Fukae‡, and James P. Simmer††

From the Departments of †Biological and Materials Sciences and ‡Orthodontics and Pediatric Dentistry, University of Michigan School of Dentistry, Ann Arbor, Michigan 48108, the †Department of Pediatric Dentistry & Dental Research Institute, School of Dentistry, Seoul National University, Seoul, Korea 110-768, the ‡Institute of Advanced Biomedical Engineering and Science, Tokyo Women’s Medical University, Tokyo 162-8666, Japan, and the Departments of **Periodontics and Endodontics and ††Biochemistry, School of Dental Medicine, Tsurumi University, Yokohama 230-8501, Japan

Dentin sialophosphoprotein (DSPP) is critical for proper mineralization of tooth dentin, and mutations in DSPP cause inherited dentin defects. Dentin phosphoprotein (DPP) is the C-terminal cleavage product of DSPP that binds collagen and induces intrafibrillar mineralization. We isolated DPP from individual pigs and determined that its N-terminal and C-terminal domains are glycosylated and that DPP averages 155 phosphates per molecule. Porcine DPP is unstable at low pH and high temperatures, and complexing with collagen improves its stability. Surprisingly, we observed DPP size variations on SDS-PAGE for DPP isolated from individual pigs. These variations are not caused by differences in proteolytic processing or degrees of phosphorylation or glycosylation, but rather to allelic variations in Dspp. Characterization of the DPP coding region identified 4 allelic variants. Among the 4 alleles, 27 sequence variations were identified, including 16 length polymorphisms ranging from 3 to 63 nucleotides. None of the length variations shifted the reading frame, and all localized to the highly redundant region of the DPP code. The 4 alleles encode DPP domains having 551, 575, 589, or 594 amino acids and completely explain the DPP size variations. DPP length variations are polymorphic and are not associated with dentin defects.

About 85% of the organic material in tooth dentin is type I collagen. The non-collagenous proteins in dentin are predominantly cleavage products of dentin sialophosphoprotein (DSPP). DSPP is processed by proteases into numerous parts. In pig, the three major cleavage products are dentin sialoprotein (DSP) (1), dentin glycoprotein (DGP) (2), and dentin phosphoprotein (DPP). The linear order of the structural domains in porcine DSPP is DSP-DGP-DPP. In the rat, the designation DSP is used to include all of the non-DPP region of DSPP (DSP-DPP). The initial proteolytic cleavage of DSPP is in a conserved context and releases DPP so that DSP proteins isolated from various mammals share the same N-terminal sequence: Asp-Asp-Pro-Asn (DDPN) (3, 4). The early work on rat DSP and DPP was performed without knowing that DSP and DPP are initially part of the same protein and must be secreted in equal amounts. Based upon the results of these early studies, it became generally accepted that DPP is ten times more abundant than DSP in dentin extracts (5–7). As DPP and DSP are expressed in a 1:1 ratio (8), to get to a 10:1 ratio DSP must be rapidly degraded, and/or there must be alternative forms of DSP that were not previously recognized, which, when counted, bring the ratio back to 1:1. In fact, high molecular weight forms of DSP have been discovered in rat dentin (9). Porcine DSP is a highly glycosylated proteoglycan that forms covalent dimers (1). These glycosylations protect some parts of the protein from detection by DSP polyclonal antibodies, so it seems likely that DSP cleavage products may be more abundant than was originally thought.

Defects in human DSPP cause various types of inherited dentin defects, including dentin dysplasia type II and dentinogenesis imperfecta types II and III (10, 11). Dspp knock-out mice exhibit dentin defects reminiscent of those observed in human dentinogenesis imperfecta (12). The DSPP mutations reported so far are primarily along intron-exon borders near the 5′-end of the gene (13–19). This region encodes DSP, the N-terminal DSPP cleavage (20). The DSP coding region in the last exon of DSPP (exon 5) has a highly redundant region that is prone to artifacts during reverse transcription and polymerase chain reactions (21). Technical difficulties have so far thwarted mutational analyses of the DPP code in kindreds with inherited dentin defects. The only sequence variation reported to be disease-causing in the DPP region was a compound insertion (6 codons)...
Porcine Dentin Sialophosphoprotein

and a deletion (12 codons) in the repetitive region (22). These insertions and deletions, however, are apparently normal sequence variations (polymorphisms) that do not cause dental disease in kindred. Alignment of 4 human DSPP sequences in GenBankTM identified 10 length variations in the DPP redundant region, ranging in length from 3 to 153 nucleotides and all maintaining the reading frame (10). An independent analysis of the kindred with the compound insertion/deletion found similar (but not identical) deletions and insertions in the DPP redundant region in normal controls and identified a missense mutation (p.P17S) in the DSP code that functional studies showed caused retention of the DSPP protein in the endoplasmic reticulum (11). Separately, a p.P17S missense alteration was found in the affected members of another family with inherited dentin defects and was thought to be causing the disease (19). These findings suggest that there may be extensive sequence length variations in the human DPP redundant region and that these length variations are compatible with normal function.

DPP protein has never been extensively characterized, and it is still not known if DPP is glycosylated (23). In this study, we isolated porcine DPP, determined its abundance relative to DSP, characterized its post-translational modifications, and determined the molecular basis for variability in the size of DPP protein in dentin extracts.

EXPERIMENTAL PROCEDURES

All experimental procedures involving the use of animals were reviewed and approved by the Institutional Animal Care and Use Program at the University of Michigan.

Preparation of Dentin Powder—Tooth germs of permanent second molars were surgically extracted with a hammer and chisel from the maxillae and mandibles of six-month-old pigs, within minutes of each animal’s termination at the Michigan State University Meat Laboratory (East Lansing, MI). Typically two maxillary and two mandibular second molars were obtained from each animal. The developmental stage of the molars was advanced in crown formation, although prior to the onset of root formation. The soft tissue was removed with forceps, and secretary and maturation stage enamel was scraped off with a curette. The remaining hard tissue was reduced to powder using a jaw crusher (Retsch, Newtown, PA).

Extraction of Proteins from Dentin Powder—Dentin powder (1.7–4.4 gm each) obtained from 22 pigs, individually, or combined from ~8 pigs (40 gm) for large scale preparations, was suspended with 50 mM Tris-HCl, 4 M guanidine buffer, pH 7.4 containing Protease Inhibitor Mixture Set III (1 mM AEBSF, 0.8 μM aprotinin, 50 μM bestatin, 15 μM E-64, 20 μM leupeptin, and 10 μM pepstatin) (Calbiochem, San Diego, CA) and 1 mM 1,10-phenanthroline (Sigma). Each sample was intensively agitated with an orbital shaker for 24 h at 4 °C. The insoluble material was pelleted by centrifugation (15,900 × g), and the guanidine extraction repeated for two more days. The insoluble guanidine extracts were dialyzed against 16 liters of 0.5 M acetic acid (HAc) containing 5 mM benzamidine (Sigma), 1 mM phenylmethylsulfonyl fluoride (Sigma), and 1 mM 1,10-phenanthroline. Each day the calcium concentration in the reservoir was measured using the Calcium Reagent Set (Pointe Scientific, Canton, MI), and the HAc/protease inhibitor solution was replaced. After 5 days, the calcium ion concentration of the HAc reservoir fell below 0.2 mM, indicating that the tooth mineral had fully dissolved. The dialysis bag contents were centrifuged, and the pellet was extracted with 0.5 M acetic acid, 2 M NaCl (AN), which dissolved dentin phosphoprotein (DPP) and DSP proteoglycan products. The AN supernatant was fractionated to purify DPP and high molecular weight DSP-containing proteoglycan.

Purification of Porcine DPP—AN extracts (~10–28 mg each) were dissolved in 0.05% trifluoroacetic acid and fractionated by reversed phase-high-performance liquid chromatography (RP-HPLC) using a Discovery C-18 column (1.0 cm × 25 cm, Sigma-Aldrich/Supelco) run at a flow rate of 1.0 ml/min and monitored at 220 nm (Buffer A: 0.05% trifluoroacetic acid; Buffer B: 80% acetonitrile, Buffer A).

Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)—SDS-PAGE was performed using Novex 4–20% Tris-glycine or NuPAGE 3–8% Tris-acetate gels (Invitrogen). Samples were dissolved in Laemmli sample buffer (Bio-Rad), and electrophoresis was carried out using a current of 20 mA for 65 min or 150 V for 1 h, respectively. The gels were stained with Simply Blue Safe Stain (Invitrogen) or Stains-all (Sigma). The apparent molecular weights of protein bands were estimated by comparison with SeeBlue® Plus2 Pre-Stained Standard (Invitrogen).

Release of N- and O-Linked Glycan Chains by Glycopeptidase A and O-Glycanase Digestion—Acid phosphatase-treated DPP (0.2 mg each) in 0.1 M citrate phosphate buffer (pH 5.0) or in 0.25 M sodium phosphate buffer (pH 5.0) was incubated with 1 milliunit of glycopeptidase A (Seikagaku America, East Falmouth, MA) or a mixture of 2.5 milliunits of O-glycosidase and 5 milliunits of sialidase Au (QA-Bio, LLC, Palm Desert, CA) containing the Protease Inhibitor Mixture Set II for 48 h or 2 h at 37 °C, respectively. At the end of the incubation period, aliquots were analyzed by SDS-PAGE. The six purified pronase-digested DPP peptides were digested with glycopeptidase A, precipitated with three volumes of ice-cold ethanol, and pelleted by centrifugation for 10 min at 10,000 × g. Glycans in the supernatant were evaporated and stored at −80 °C.

Quantitative Determination of Phosphoserine in DPP—DPP from 22 individual pigs (0.5–0.7 mg each) was partially hydrolyzed with 0.2 ml of 6 N HCl for 3 h at 110 °C. After the reaction, the sample was evaporated and dissolved in 1 ml of deionized (DI) water, and 0.25 ml was fractionated on a TSK-gel SAX column (4.6 mm × 15 cm, TOSOH, Tokyo, JPN). The column was equilibrated with 40 mM potassium phosphate buffer (pH 4.0), and eluted with the same buffer at a flow rate of 1.5 ml/min at room temperature. Phosphoserine was detected by monitoring absorbance at 210 nm, and authentic phosphoserine, phosphothreonine, and phosphotyrosine (Sigma) were used as references. The mol phosphates released was divided by the starting weight × 100 g/mol (100-kDa molecular mass) to calculate phosphate per molecule.

Quantitative Determination of DPP Phosphorylation by Acid Phosphatase Digestion—DPP (150 μg each) in 150 ml of 10 mM sodium acetate, 50 mM EDTA buffer (pH 5.8) from individual pigs was incubated with 0.2 unit of acid phosphatase (white potato) (Sigma) containing the Protease Inhibitor Mixture Set.
II for 48 h at 37 °C. Free phosphorus in aliquots (20 ml) of acid phosphatase digests was measured by colorimetric assay using the Inorganic Phosphorus Reagent Set (Pointe Scientific, Canton, MI). The number of phosphates per molecule was calculated assuming a molecular mass of 100 kDa (Table 1).

**TABLE 1**

| Pig | IP | P-Ser | Ave | Pig | IP | P-Ser | Ave |
|-----|----|-------|-----|-----|----|-------|-----|
| A   | 217| 199   | 208 | L   | 254| 211   | 233 |
| B   | 227| 184   | 206 | M   | 217| 180   | 199 |
| C   | 218| 198   | 208 | N   | 229| 184   | 207 |
| D   | 202| 189   | 196 | O   | 215| 193   | 204 |
| E   | 234| 192   | 213 | P   | 235| 163   | 209 |
| F   | 246| 203   | 223 | Q   | 206| 175   | 191 |
| G   | 255| 196   | 226 | R   | 262| 202   | 232 |
| H   | 203| 181   | 192 | S   | 252| 228   | 240 |
| I   | 272| 177   | 225 | T   | 251| 237   | 244 |
| J   | 231| 156   | 194 | U   | 216| 185   | 201 |
| K   | 249| 167   | 208 | V   | 219| 194   | 206 |
|     |    |       |     | Total|     | 2558 | 2172 | 2365 |
|     | Ave| 216  | 186 | 209 | Ave| 233  | 197 |

**Extraction of DPP from Gels**—Three DPP protein bands from large scale dentin preparations were excised from a polyacrylamide gel after Stains-all staining. Each gel slice was transferred to a D-Tube Dialyzer (midi size) (Novagen/EMD Chemicals, Inc., Gibbstown, NJ) and electroeluted with 25 mM Tris, 250 mM Tricine, 0.025% SDS buffer (pH 8.5) at 150 V for 3 h. The eluate (DPP) was precipitated with 20% trichloroacetic acid. The pellet was incubated in acetone overnight at −20 °C and centrifuged at 4 °C for 30 min at 14,000 × g. The supernatant was decanted, the pellet dried under a hood, and characterized by Edman sequencing.

**Purification of DPP Glycopeptide by Pronase Digestion**—DPP (25 mg) was digested with acid phosphatase, dialyzed against water for 2 days, and lyophilized. Acid phosphatase-treated DPP (18 mg) was dissolved with 1 ml of 50 mM Tris-HCl buffer, pH 8.0. This solution was incubated with 0.1 mg of pronase (Calbiochem) for 15 h at 37 °C. The digest was fractionated by size exclusion chromatography using a Sephadex G-15 column (1.6 cm × 100 cm, GE Healthcare Life Sciences) equilibrated with 50 mM Tris-HCl buffer, pH 8.0. DPP peptides were eluted with the same buffer at a flow rate of 0.2 ml/min at 4 °C with absorbance monitored at 220 nm, and analyzed for glycosylation using a phenol-sulfuric acid assay at 490 nm. The fraction of peptides containing DPP glycopeptide was eluted in the first peak, and this fraction was separated by hydrophilic interaction liquid chromatography (HILIC) using a polyhydroxyethyl A column (4.6-mm inner diameter × 20 cm, The Nest Group, Inc., Southborough, MA) equilibrated with 15 mM triethylamine phosphate (TEAP) in 95% acetonitrile (pH 3.0). Peptides were eluted with a linear acetonitrile gradient containing 15 mM TEAP in 5% acetonitrile (pH 3.0) at a flow rate of 0.5 ml/min at room temperature, while monitoring the absorbance at 220 nm. Six peaks were collected and evaporated. Aliquots were used for both amino acid sequence and glycosylation analyses.

**Preparation and Isolation of 2-AA-labeled Glycans**—Glycans released by glycopeptidase A were labeled with the 2-aminobenzoic acid (2-AA) labeling kit (QA-Bio, Palm Desert, CA), and labeled glycans were purified by LudgerClean.
Effect of Collagen on the Stability of DPP—DPP (0.2 mg) was dissolved with 0.2 ml of 0.5 M acetic acid. An aliquot of sample (50 μl) was then mixed with 0.2 ml of 0.5 M acetic acid containing 0.2 mg of acid-soluble human placenta type I collagen (Abcam, Cambridge, MA). Another aliquot (50 μl) was mixed with 0.2 ml of 0.5 M acetic acid only. Two samples with and without the collagen were divided into five tubes (50 μl each). Each sample was incubated and analyzed by SDS-PAGE.

Amino Acid Analysis—The purified peptide samples (0.02–0.03 mg) were hydrolyzed with 6 N HCl at 115 °C for 16 h. The amino acid analyses were performed using a Beckman Model 7300 automatic amino acid analyzer.
calcium, and relatively small cleavage products from the non-DPP portion of DSP (the N-terminal fragments of DSP, DGP, and extended DGP). Very little protein in the AN extract stains with CBB, but this extract is rich in Stains-all-positive DSPP-derived dentin proteins (Fig. 1B). Lower molecular weight protein was removed from the AN extract by size exclusion chromatography, and the high molecular weight DSPP-derived proteins are separated into DPP and DSP components by RP-HPLC. Five RP fractions are evident in the chromatogram (Fig. 1C). R1 contains DPP fragments, R2 contains intact DPP and DPP fragments, R3 contains the central proteoglycan core of DSP, R4 contains DSP, and R5 contains DSP-DGP, or the entire non-DPP region of DSPP (Fig. 1D) (24).

To quantify the relative abundance of DPP and high molecular weight DSP, we isolated and weighed the AN extract from the second molars of 22 individual pigs and isolated the five RP-HPLC fractions containing DSPP-derived proteins. We normalized for variations in protein quantity among the pig extracts by dividing the quantity of protein in each RP fraction by the total amount of protein in the AN extract and calculated the total amount of DPP (R1 + R2) relative to the total amount of high molecular weight DSP in fractions (R3 + R4 + R5). On a weight basis, porcine dentin in the second molars of six-month-old pigs contains between 1.2 and 2.0 times more DPP than high molecular weight DSP (Fig. 1E). There are, however, low molecular weight DSP cleavage products outside of the AN extract (lane A, Fig. 1B). These cannot be efficiently isolated and their quantity was not determined for individual pigs, but including the small DSP cleavage products in the A extract would have brought the ratio of DSP to DPP (on a weight basis) very close to 1:1, which is consistent with the fact that DPP and DSP are synthesized from the same mRNA transcript in 1:1 molar ratios (8, 21).

DPP Post-translational Modifications—An estimate of the gross levels of post-translational modifications in porcine dentin phosphoprotein was made by examining DPP before and after dephosphorylation and deglycosylation (Fig. 2). Porcine DPP in our large scale preparations (that combine dentin powder from 8 pigs) typically appears as two prominent bands, often accompanied by one or two distinct but weaker DPP bands. We originally assumed that the multiple DPP bands, all migrating near 100 kDa on SDS-PAGE, to be due to variations in the amount of DPP post-translational modification. Dephosphorylation of DPP with acid phosphatase, however, causes these DPP bands to shift to a lower apparent molecular weight on SDS-PAGE (Fig. 2A), but does not alter the pattern of multiple bands. N-deglycosylation (Fig. 2B) and O-deglycosylation (Fig. 2C) of the dephosphorylated DPP does not appreciably alter the mobility of DPP on SDS-PAGE and also does not alter the pattern of multiple bands. Each DPP band was excised from the gel and characterized by N-terminal sequencing. All three DPP bands gave the same N terminus: DDPNXXE. These results demonstrate that the various DPP bands extracted from porcine dentin do not differ from each other because of variations in the levels of phosphorylation, glycosylation, or through the use of multiple cleavage sites to release DPP from the parent DSP protein. We then characterized the number of phosphates on DPP and identified DPP glycosylation sites.
Porcine Dentin Sialophosphoprotein

We isolated DPP from 22 individual pigs and determined the numbers of phosphate per molecule using two independent methods. One method measures the amount of phosphoserine released after acid hydrolysis, the other measures inorganic phosphate released by acid phosphatase. Using the apparent molecular mass of DPP (100 kDa) on SDS-PAGE to convert from grams to moles, the number of phosphates per molecule of DPP came to an average of 212 per molecule. This number of phosphates would contribute 17 kDa to the mass of the protein.

To identify specific glycosylation sites, DPP was digested with pronase, a mixture of proteolytic enzymes that digests glycoproteins down to short glycopeptides that are protected from further degradation by their bulky carbohydrate chains. The pronase digestion products were fractionated by size exclusion chromatography and assayed for the presence of glycosylation (Fig. 3A). The glycosylation-positive peak was separated into six fractions (labeled Pr-DPP-1 through Pr-DPP-6) by HILIC (Fig. 3B). The protein in each fraction was deglycosylated and characterized by N-terminal sequencing (Table 2). The glycosylations were fluorescently labeled and analyzed by NP-HPLC (Fig. 4).

**TABLE 2**

| Pr-DPP | Sequence      | Positions | N-glycan |
|--------|---------------|-----------|----------|
| 1      | DDPNXE        | 458–464   | −        |
| 2      | SKEEAE        | 502–508   | +        |
| 3      | AEBDXTSD     | 506–513   | +        |
| 4      | SDSXSXS       | 918–924   | +        |
| 5      | SDES          | 490–494   | −        |
| 6      | DSASD         | 965–969   | −        |

*Numbering for DSPP allele pDSPP<sup>1023</sup>.

A N-terminal sequence of 6 pronase-generated DPP peptides (Fig. 3, C–F). No labeled glycosylations were observed for Pr-DPP-1, Pr-DPP-5, and Pr-DPP-6. Two N-linked glycosylation sites were identified in the other fractions: at Asn<sup>525</sup> (Pr-DPP-2 and Pr-DPP-3) and Asn<sup>937</sup> (numbered for the smallest DSPP allele; Pr-DPP-4). Based upon its retention time on the NP-HPLC column, the glycosylation at Asn<sup>525</sup> can have either no sialic acid (Pr-DPP3) or two sialic acids (Pr-DPP2). The glycosylation at Asn<sup>937</sup> has two sialic acids (Pr-DPP4). These results demonstrate that porcine DPP is glycosylated at two positions, and that variations in glycosylation cannot account for the multiple bands of DPP observed on SDS-PAGE.

**Allelic Polymorphisms as the Source of DPP Length Variations**—DPP protein was separately isolated from 22 pigs and analyzed by SDS-PAGE (Fig. 4A). Each pig showed either one or two DPP bands of equal intensity, migrating near 100 kDa. DPP bands from six pigs representing the full range of DPP size variations were characterized by Edman sequencing and all displayed the same N-terminal sequence: DDPNXE. Dephosphorylation caused the DPP bands to shift to a lower position on the gel, but did not alter the pattern of bands (Fig. 4B). We amplified the DPP coding region (which is not interrupted by any introns) using genomic DNA isolated from the same 22 pigs used to obtain the DPP protein and found length variations in the DPP code that closely correlated with the length variations observed at the protein level (Fig. 4C).

The DPP PCR products from 8 of the pigs were cloned and characterized. Some length variations in the DPP amplification products were identified as PCR artifacts (data not shown), as they could not be reproduced by repeating the experiment. PCR products that represented singular events were consistently ignored in favor of only reproducible outcomes. We characterized 4 groups of DPP cDNAs that were generated from multiple independent PCR cloning experiments, and these clones can account for all of the DPP size variability observed at the protein level. The DPP clones correspond to 4 allelic variations of the DPP code in the Dspp genes of this group of pigs. The 4 clones varied in length because of the presence of multiple in-frame deletions or insertions. All of the length variations localized to the highly redundant region of the DPP code (Fig. 5A). The DPP coding region has 27 sequence variations among the 4 alleles (Fig. 5B). Of these, 11 are point mutations (changing 5 amino acids) and 16 are length variations ranging from 3 to 63 bp that maintain the open reading frame (Table 1 and supplemental data).

The DPP coding regions in the 4 alleles are 1656 (EU419998), 1728 (EU419999), 1770 (EU420000), and

**FIGURE 4. Length polymorphisms of DPP protein and coding region.** A, SDS-PAGE showing variations in the mobility of DPP protein isolated from 22 individual pigs. 9 of 22 pigs showed only a single band of DPP (A, B, C, D, G, H, L, O). Thirteen pigs showed 2 DPP bands (E, F, I, J, K, M, N, P, R, S, T, U, V). Although no individual pig showed more than 2 DPP bands, 6 bands of non-identical sizes could be discerned among the 22 pigs. B, dephosphorylation of DPP from the 22 pigs shifted the bands lower on the gel but did not alter the pattern of bands. C, DPP PCR products using genomic DNA as template generated an identical pattern of 1 or 2 bands.
observed DPP sequence variations are not associated with any loss of function.

**DPP Stability**—The effects of pH and temperature on the stability of DPP were examined (Fig. 6). Dentin phosphoprotein is stable for at least 20 h at 4 °C and at 20 °C within the entire range of pH values tested (pH 4–11). At 37 °C, DPP is sensitive to hydrogen ion concentration, and a significant reduction in the amount of DPP is observed after 20 h at pH 5 and below. The degradation of DPP at low pH 5 occurs more rapidly if the protein is heated at 95 °C for 5 min. The instability of DPP observed at low pH and high temperature is reduced in the presence of type I collagen (Fig. 7, A–C).

**DISCUSSION**

**Relative Abundance of DSP and DPP**—Using the porcine animal model, which provides abundant dentin proteins and having previously surveyed dentin extracts for the presence of DSPP-derived proteins, we were able to quantify levels of DSPP-derived proteins from developing pig molars and determined that there were approximately equal weights of DSP and DPP cleavage products from the DSP-DGP and DPP regions in our dentin extracts. There are differences, however in the way these proteins are degraded. After the release of DSP, DSP-DGP is processed into numerous relatively small cleavage products (~17–35 kDa) that can be found in the A extract. Larger DSP-DGP pieces that contain the two long chondroitin sulfate attachments are in the AN extract. DSP degradation products do not form discrete bands, but rather show a continuum of all sizes that make a Stains-all-positive smear on SDS-PAGE (Fig. 1D, R1 and R2). This pattern might reflect an inherent chemical instability rather than processing by proteases (25).

**DPP Post-translational Modifications**—DPP is the most acidic protein known and has an isoelectric point of 1.1 (26). The low electric point is due to its large number of acidic amino acids and phosphorylated serines. Porcine DPP migrates at about 100 kDa on SDS-PAGE. When DPP is isolated from individual pigs, either one or two DPP bands are observed. We have demonstrated that the presence of multiple bands is due to allelic variation. Porcine DPP is relatively homogeneous with respect to its post-translational modifications. This contrasts with rat DPP where multiple forms have been proposed that vary according to their degree of phosphorylation (3).

We determined that the average number of phosphates per DPP molecule is 212 when using DPPs apparent molecular mass (100 kDa) on SDS-PAGE to calculate moles of DPP in a weighed sample. The results of the cloning and post-transla-
tional modification studies, however, suggest that 73 kDa might be a better estimate for the molecular mass of porcine DPP. The predominant forms of DPP in our population of pigs contain 551 or 575 amino acids. Without post-translational modifica-

tions, the common DPP forms have calculated molecular masses of 54.6 and 57.1 kDa. If DPP were 73 kDa, the average number of phosphates per molecule would be 155 (73 kDa/mol × 212 P/100 kDa/mol = 155 P). This number of phosphates (155/molecule) has a mass of 12.4 kDa (155 P × 80 Da/p = 12.4 kDa). The breakdown for the 73-kDa native protein is: 56-kDa protein + 12.4-kDa p + 4.6-kDa N-glycos = 73 kDa. As there are 310 serines per molecule, half of the serines in DPP are phosphorylated.

The DPP deduced amino acid sequences of the 4 DPP alleles are aligned in Fig. 8. There are eight potential N-linked glyco-
sylation sites (NXS, where x≠P) in the N-terminal region of DPP and five in the C-terminal region. Pronase digests of DPP generated short glycopeptides derived from one site in each of these regions. It seems unlikely that additional sites are glyco-
sylated as the mobility of DPP on SDS-PAGE does not change appreciably following deglycosylation.

**DPP Length Polymorphisms**—The size of DPP varies among species. Rat, bovine, and porcine DPPs have an apparent molec-

ular mass just under 100 kDa (3, 27), while human DPP migrates at about 140 kDa on SDS-PAGE (28). It is becoming increasingly apparent that the size of DPP also varies within species. These observations suggest that DPP function is not affected by length variations in its repeat region. The functional significance (or insignificance) of DPP length variations is important when considering whether a particular sequence

---

**FIGURE 7.** Collagen-DPP interactions increase the stability of DPP and induce collagen fibril assembly. A, SDS-PAGE stained with CBB or CBB plus Stains-all. DPP does not stain with CBB but can be visualized with Stains-all. DPP is not stable when incubated below pH 6 at higher temperature. B, SDS-
PAGE showing type I collagen (Col I). Col I stains with CBB but does not stain with Stains-all (SA). C, SDS-PAGE of Col I and DPP after being mixed together and incubated at 4, 20, 37, or 95 °C. The presence of collagen reduces the instability of DPP at 37 and 95 °C.

**FIGURE 8.** Alignment of the DPP deduced amino acid sequences of 4 porcine DSPP alleles pDSPP1066/DPP594 (black), pDSPP1061/DPP589 (blue), pDSPP1047/DPP575 (green), and pDSPP1023/DPP551 (red). Key: five amino acid substitutions are in bold. Asparagines (N) in the appropriate context for glycosylation are underlined. Asparagines determined to be glycosylated following pronase digestions of DPP are in bold and underlined.
variation is disease-causing or simply a polymorphism. It is also important when considering DPP function.

Our studies of 22 commercial pigs demonstrate that DPP allelic variations at the DNA level translate into size variations at the protein level. The 4 Dspp alleles showed 16 DPP length variations ranging from 1 to 21 amino acids and 5 single amino acid substitutions. There is no apparent loss of function due to this variability. Previously we reported the cloning of two porcine DSPP (pDSPP) cDNA clones encoding a total of 600 and 593 amino acids, respectively: pDSPP600/DPP128 (Acc. AY161863) and pDSPP593/pDPP121 (Acc. AY161862). The DPP part of these clones was short and apparently had most of the DPP region deleted during the cloning procedure, possibly during the reverse transcription reaction (21). Only the DPP region was affected by artifacts, and the early DSPP clones demonstrated that porcine DPP begins at Asp<sup>473</sup> of DSPP. Therefore, the 4 DPP allelic variants identified in the 22 pigs can be designated pDSPP1022/DPP551, pDSPP1047/DPP575, pDSPP1061/DPP589 and pDSPP1066/DPP594. The pDSPP1023/DPP551 and pDSPP1047/DPP575 alleles predominate (38/44 alleles) in our group of pigs and express the two major DPP bands observed in protein preparations from the same animals. By comparison, rat DSPP protein has 970 amino acids, 523 of which are in the DPP portion of the protein (rDSPP970/DPP523 Acc. AJ403971) (29), while human DSPP is larger and may also show allelic length variations (hDSPP1301/DPP839, hDSPP1325/DPP863, hDSPP1347/DPP885, hDSPP1375/DPP913, hDSPP1393/DPP931, hDSPP1415/DPP953, hDSPP1435/DPP975) (30).

**DPP Instability**—The DPP domain is inherently unstable and degrades when heated in vitro (25). Increasing concentrations of SDS helped protect against thermal degradation. We confirmed these findings and add that DPP is unstable at low pH and that type I collagen helps protect DPP from thermal degradation, presumably by forming Col I-DPP complexes.

**DPP Is an Intrinsically Disordered Protein**—Proteins that do not appear to adopt a well-defined conformation under native conditions are increasingly being categorized as intrinsically disordered proteins (IDP) (31). The tendency for proteins to be intrinsically disordered can be predicted from their amino acid sequences. DPP is 29% aspartic acid and 54% serine, with half of the serines phosphorylated. The PONDR computational software for predicting naturally disordered structures gives DPP its highest possible score (1.0) over a continuous region extending for more than 350 amino acids (32). NMR analysis showed that bovine DPP is a molecule of uniformly high mobility, which is consistent with an intrinsically disordered protein (33).

In summary, our results demonstrate that the amounts of protein in porcine dentin extracts that are derived from the DSP-DGP and DPP regions of DSPP are about equal. Porcine DPP is extensively phosphorylated, averaging 155 phosphates per molecule, and is glycosylated at its N-terminal and C-terminal domains. Although porcine DPP shows an apparent molecular mass of 100 kDa on SDS-PAGE, we estimate its actual size to be 73 kDa, based upon the deduced masses of its amino acid chain and post-translational modifications. Allelic length variations in the DPP coding region translate into size variations at the protein level. Such length variations are exceedingly rare in proteins. They occur, however, within the highly redundant DPP region and represent normal variations that do not interfere with protein function.

**Acknowledgments**—We thank Tom Forton, manager of the Michigan State University Meat Laboratory and members of the Michigan State University Department of Animal Science for their kind assistance in obtaining fresh developing molars from pigs. We thank Dr. Myron Crawford, director of W. M. Keck Foundation Biotechnology Resource Laboratory at Yale University and Nancy Williams for the amino acid analysis and protein sequencing.

**REFERENCES**

1. Yamakoshi, Y., Hu, J. C., Fukae, M., Iwata, T., Kim, J. W., Zhang, H., and Simmer, J. P. (2005) *J. Biol. Chem.* 280, 1552–1560
2. Yamakoshi, Y., Hu, J. C., Fukae, M., Zhang, H., and Simmer, J. P. (2005) *J. Biol. Chem.* 280, 17472–17479
3. Butler, W. T., Bhown, M., DiMuzio, M. T., Cothran, W. C., and Linde, A. (1983) *Arch. Biochem. Biophys.* 225, 178–186
4. Huq, N. L., Cross, K. J., Talbo, G. H., Riley, P. F., Loganathan, A., Crossley, M. A., Perich, J. W., and Reynolds, E. C. (2000) *J. Dent. Res.* 79, 1914–1919
5. Butler, W. T., Bhown, M., Dimuzio, M. T., and Linde, A. (1981) *Coll. Relat. Res.* 1, 187–199
6. Butler, W. T. (1998) *Eur. J. Oral. Sci.* 106, 204–210
7. Patel, P. (2001) *Nat. Genet.* 27, 129–130
8. MacDougall, M., Simmons, D., Luan, X., Nydegger, J., Feng, J., and Gu, T. T. (1997) *J. Biol. Chem.* 272, 835–842
9. Qin, C., Brunn, J. C., Baba, O., Wygant, J. N., McIntyre, B. W., and Butler, W. T. (2003) *Eur. J. Oral. Sci.* 111, 235–242
10. Kim, J. W., and Simmer, J. P. (2007) *J. Dent. Res.* 86, 392–399
11. Hart, P. S., and Hart, T. C. (2004) *Cells Tissues Organs* 186, 70–77
12. Sreenath, T., Thyagarajan, T., Hall, B., Longenecker, G., D’Souza, R., Hong, S., Wright, J. T., MacDougall, M., Sauk, J., and Kulkarni, A. B. (2003) *J. Biol. Chem.* 278, 24874–24880
13. Xiao, S., Yu, C., Chou, X., Yuan, W., Wang, Y., Bu, L., Gu, Q., Qian, M., Yang, J., Shi, Y., Hu, L., Han, B., Wang, Z., Huang, W., Liu, J., Chen, Z., Zhao, G., and Kong, X. (2001) *Nat. Genet.* 27, 201–204
14. Malmgren, B., Lindskog, S., Elgadi, A., and Norgren, S. (2004) *Hum. Genet.* 114, 491–498
15. Kim, J. W., Nam, S. H., Jang, K. T., Lee, S. H., Kim, C. C., Hahn, S. H., Hu, J. C., and Simmer, J. P. (2004) *Hum. Genet* 115, 248–254
16. Kim, J. W., Hu, J. C., Lee, J. I., Moon, S. K., Kim, Y. J., Jang, K. T., Lee, S. H., Kim, C. C., Hahn, S. H., and Simmer, J. P. (2005) *Hum. Genet.* 116, 186–191
17. Holappa, H., Nieminen, P., Tolva, L., Lukiinmaa, P. L., and Alaluusua, S. (2006) *Eur. J. Oral. Sci.* 114, 381–384
18. Song, Y., Wang, C., Peng, B., Ye, X., Zhao, G., Fan, M., Fu, Q., and Bian, Z. (2006) *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod.* 102, 360–374
19. Zhang, X., Chen, L., Liu, J., Zhao, Z., Qu, E., Wang, X., Chang, W., Xu, C., Wang, Q. K., and Liu, M. (2007) *BMC Med. Genet.* 8, 52
20. Ritchie, H. H., Hou, H., Veis, A., and Butler, W. T. (1994) *J. Biol. Chem.* 269, 3698–3702
21. Yamakoshi, Y., Hu, J. C., Liu, S., Zhang, C., Oida, S., Fukae, M., and Simmer, J. P. (2003) *Eur. J. Oral. Sci.* 111, 60–67
22. Dong, J., Gu, T., Jeffords, L., and MacDougall, M. (2005) *Am. J. Med. Genet.* 132, 305–309
23. Qin, C., Baba, O., and Butler, W. T. (2004) *Crit. Rev. Oral. Biol. Med.* 15, 126–136
24. Yamakoshi, Y., Hu, J. C., Iwata, T., Kobayashi, K., Fukae, M., and Simmer, J. P. (2006) *J. Biol. Chem.* 281, 38235–38243
25. Ibaraki, K., Shimokawa, H., and Sasaki, S. (1991) *Matrix* 11, 115–124
26. Jonsson, M., and Fredriksson, S. (1978) *J. Chromatogr.* 157, 234–242
27. Termine, J. D., Belcourt, A. B., Miyamoto, M. S., and Conn, K. M. (1980) *J. Biol. Chem.* 255, 9769–9772
28. Chang, S. R., Chiego, D. J., and Clarkson, B. H. (1996) *Calcif. Tissue Int.*
Porcine Dentin Sialophosphoprotein

29. George, A., Bannon, L., Sabsay, R., Dillon, J. W., Malone, J., Veis, A., Jenkins, N. A., Gilbert, D. J., and Copeland, N. G. (1996) J. Biol. Chem. 271, 32869–32873

30. Gu, K., Chang, S., Ritchie, H. H., Clarkson, B. H., and Rutherford, R. B. (2000) Eur. J. Oral Sci. 108, 35–42

31. Hansen, J. C., Lu, X., Ross, E. D., and Woody, R. W. (2006) J. Biol. Chem.

32. Yamakoshi, Y. (2008) J Oral Biosciences, in press

33. Cross, K. J., Huq, N. L., and Reynolds, E. C. (2005) J. Pept. Res. 66, 59–67

34. Thomas, L. J., Panneerselvam, K., Beattie, D. T., Picard, M. D., Xu, B., Rittershaus, C. W., Marsh, H. C., Jr., Hammond, R. A., Qian, J., Stevenson, T., Zopf, D., and Bayer, R. J. (2004) Glycobiology 14, 883–893