Title: Dynamic Processing of Recombinant Dentin Sialoprotein-Phosphophoryn Protein

Authors:
V. Godovikova
H. H. Ritchie

Department of Cariology, Restorative Sciences and Endodontics
University of Michigan School of Dentistry

Running title:
DSP-PP precursor protein processing

Keywords:
DSP-PP<sub>240</sub> precursor protein, self-processing, Dentin Sialoprotein, Phosphophoryn, proteolytic activity

Corresponding author:
Helena H. Ritchie
Room 2393 Dental Building
University of Michigan School of Dentistry
1011 N. University Ave.
Ann Arbor, MI 48109-1078
U.S.A.
E-mail: helenar@umich.edu
Tel: 1-734-763-3746
Fax: 1-734-936-1597
Summary

Dentin sialoprotein (DSP) and phosphophoryn (PP) are the two noncollagenous proteins classically linked to dentin, but more recently found in bone, kidney and salivary glands. These two proteins are derived from a single copy DSP-PP gene. While this suggests that the DSP-PP gene is first transcribed into DSP-PP mRNAs, which later undergo processing to yield the DSP and PP proteins, this mechanism has not yet been demonstrated because of the inability to identify a DSP-PP precursor protein from any cell or tissue sample. To study this problem we utilized a baculovirus expression system to produce recombinant DSP-PP precursor proteins from a DSP-PP<sub>240</sub> cDNA, which represents one of several endogenous DSP-PP transcripts that influence various tooth mineralization phases. Our in vitro results demonstrate that DSP-PP<sub>240</sub> precursor proteins are produced by this system, and are capable of self-processing to yield both DSP and PP proteins. We further demonstrated that purified recombinant DSP-PP<sub>240</sub>, purified recombinant PP<sub>240</sub>, as well as the native highly phosphorylated protein (HP; equivalent to the PP<sub>523</sub> isoform) have proteolytic activity. These newly identified tissue proteases may play key roles in tissue modeling during organogenesis.
Introduction

Dentin sialoprotein (DSP) and phosphophoryn (PP) are the two most abundant noncollagenous proteins in dentin. In rats DSP and PP coding sequences are derived from DSP-PP transcripts (1,2), a finding confirmed in the mouse (3), and in human DSP-PP transcripts (4). Immunohistochemistry and in situ studies showed that DSP/PP proteins and DSP-PP mRNA expression are tightly associated with dentin mineralization (5-8). Mutations in the DSP-PP gene are linked to dentinogenesis imperfecta II and hearing loss (9,10) while DSP-PP null mouse exhibit hypomineralization and dentin dysplasia (11). These studies, taken together with the early finding that PP binds Ca++ and can initiate hydroxyapatite formation in vitro (12-14), strongly support the assertion that DSP and PP proteins play significant roles in dentin mineralization. However, more recent findings by Godovikova et al. (15) demonstrate DSP-PP mRNA expression occurring not only in teeth, but also in bone, kidney and salivary glands suggesting that the DSP-PP gene may participate in a variety of processes during organogenesis.

Since DSP and PP proteins are derived from a single copy DSP-PP gene, it is commonly assumed that the DSP-PP gene is first transcribed into DSP-PP mRNAs, translated to become DSP-PP precursor proteins, then enzymatically cleaved to yield DSP and PP proteins found in dentin. Yet, dentin DSP protein has been estimated at 5-8% of the dentin NCP content (16) and PP has been estimated to be >50% of dentin NCP content (17). To date, this discrepancy between the observed 1:6 DSP:PP dentin protein ratio and the expected 1:1 ratio has not been explained. A key-missing element in this story has been the inability to identify a DSP-PP precursor protein from any cell or tissue sample. And, without this putative DSP-PP precursor protein, it has not been possible to study DSP-PP post-translational processing and cleavage, leaving unanswered such questions as where DSP-PP cleavage occurs (i.e., intracellularly, or extracellularly) and what cleavage enzyme(s) may be involved.

To answer these DSP-PP protein processing questions, we utilized a baculovirus expression system to produce recombinant DSP-PP precursor proteins from a DSP-PP240 cDNA, which represents one of several endogenous DSP-PP transcripts (18,19) believed to play different roles during dentin mineralization. Our in vitro results demonstrate that DSP-PP240 precursor proteins are produced by this system, and are capable of self-processing to yield both DSP and PP proteins.

Materials and Methods

Construction and expression of DSP370 cDNA and DSP-PP240 cDNA using the pVL941 and pVL1392 baculovirus expression systems. A diagram of the DSP-PP gene, showing the relative positions of DSP-PP240 cDNA and DSP370 is shown in Figure 1A. DSP-PP240 cDNA contains the 17 amino acid leader sequence, DSP-PP240 coding sequence, a stop codon and a 200 bp 3’ noncoding sequence (see Figure 1A). DSP-PP240 cDNA was subcloned into the baculovirus expression vector pVL1392 at Xba I and Bam HI sites. The DSP370 cDNA, containing a signal sequence encoding the 17 amino acid leader sequence and the partial DSP coding sequence for the first 370 amino acids, was subcloned into the baculovirus expression vector pVL941 at the BamHI site. This DSP370...
cDNA construct yielded a fusion protein, which contains a 370 amino acid DSP peptide sequence as well as an additional 18 amino acid peptide derived from the viral sequence (see Figure 1C). An antisense DSP\textsubscript{370} cDNA construct was produced as a control.

The pVL941-DSP\textsubscript{370} cDNA, pVL941-antisense-DSP\textsubscript{370} cDNA and pVL1392-DSP-PP\textsubscript{240} cDNA constructs were individually cotransfected with a linearized BaculoGold baculovirus DNA (PharMingen’s BaculoGold transfection kit, PharMingen, San Diego, CA) (20) into insect sf9 cells to obtain virus stock. To produce recombinant proteins, insect sf9 cells infected with recombinant virus stock at a multiplicity of 10, were grown in Grace’s insect cell medium (Invitrogen, Carlsbad, CA) supplemented with 10% FCS to a density of 2x10\textsuperscript{6} cells/T25 flask. Supernatants were harvested on days 1, 2, 3, and 4 after infection and partially purified using a polyanion extraction protocol (see below).

**Partial Purification of recombinant DSP-PP\textsubscript{240} and DSP\textsubscript{370} proteins using polyanion extraction**

This protocol takes advantage of the finding that acidic proteins such as DSP and PP are soluble in 5% TCA (21). For DSP\textsubscript{370} purification, the supernatant from pVL941-DSP\textsubscript{370} cDNA infected sf9 cells was diluted 1:20 with 100% TCA. The majority of culture medium proteins were precipitated and removed by centrifugation. The TCA soluble portion was further neutralized with 1/5 original volume of 3 M Tris.HCl, pH 8.8 and precipitated with 1/10 volume of 1 M CaCl\textsubscript{2}. This new precipitate was dissolved again in 5% TCA and precipitated with 3 M Tris.HCl, pH 8.8 and 1 M CaCl\textsubscript{2}. This second CaCl\textsubscript{2} precipitate, containing recombinant DSP\textsubscript{370}, was dissolved in 1/10 original volume of 0.1 M EDTA. Purified DSP-PP\textsubscript{240} was similarly obtained. The recombinant proteins were stored in 0.1 M EDTA and were stable for 2 years.

**SDS-Polyacrylamide Gel Electrophoresis**

SDS-PAGE was performed using 7.5%, 10%, and 4-15% polyacrylamide gels. Samples were dissolved in Laemmli sample buffer (Bio-Rad, Hercules, CA, USA). Electrophoresis was carried out at 60mA for 45 min. The gels were stained with Bio-Safe Coomassie blue R250 (Bio-Rad, Hercules, CA, USA) or Stains–All (Sigma, St. Louis, MO, USA). The apparent molecular weights of the protein bands were estimated by comparison with Bench Mark Pre-stained Protein Ladder standards (Invitrogen, Carlsbad, CA). The gels were air-dried in a cellophane membrane overnight. Stains–All stains acidic proteins (i.e., DSP or PP) blue, while neutral proteins (such as BSA) appear orange red.

**Gel Purification of recombinant DSP-PP\textsubscript{240} precursor proteins**

The partially purified recombinant DSP-PP\textsubscript{240} precursor proteins were electrophoresed on 7.5% or 10% SDS-PAGE gels. The DSP-PP\textsubscript{240} band was excised, electro-eluted, and concentrated. All steps were performed in the presence of a protease inhibitor IP cocktail (Sigma, St. Louis, MO.).

Dentin extract preparation. Dentin extract (DE) was prepared from rat incisors (22).
**Highly phosphorylated protein preparation.** Highly phosphorylated protein (HP) was prepared from rat incisors following the method of Marsh (23). The purified HP contained 2.9 nmole Pi/µg HP and the N-terminal sequence was determined as DDPN.

**Rabbit anti-rat DSP.** Rat DSP peptide (N’-CPSGQSQNQGLETEGSSTGN-C’) was synthesized and purified by reverse phase HPLC (Genemed Synthesis Inc, San Francisco, CA). This peptide was then conjugated to keyhole limpet hemocyanin (KLH) and used for generating rabbit anti-rat DSP antibodies. These anti-DSP antibodies (1:200 dilution) were used to perform Western blot analyses to identify the expressed recombinant DSP proteins.

**Western blot analyses**
Proteins were electrophoresed and transferred to nitrocellulose filters using a semi-dry apparatus. The nitrocellulose filters were then hybridized with 1% blocking agent to block nonspecific antibody binding then incubated with a 1:200 dilution of primary antibodies (i.e., rabbit anti-rat DSP antibodies) overnight at 4°C. The nitrocellulose filters were washed with TBS three times and incubated with secondary antibodies (goat anti-rabbit antibodies conjugated with alkaline phosphatase at 1:2000 dilution) for 3 hours. The filters were washed and NBT/BCIP (nitroblue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate (BioRad Laboratories, Hercules, CA) was added for color development.

**N-Terminal Amino Acid Sequence Analysis**
After gel electrophoresis separation, the proteins were transferred onto Immobilon–P (PVDF) membrane (Millipore) by semi-dry technique in CAPS-methanol buffer (0.8mA/cm, 1-2 h). After transfer, the membranes were washed and lightly stained with Stains All and the protein bands were excised. The N-Terminal amino acid sequence of the proteins was determined by Procise Protein Sequencer 494 HT (Applied Biosystems, Foster City, CA) at the Protein Core Facility, University of Michigan, using reagents and methods recommended by the manufacturer.

**Mass spectral analyses**
Gradient 4-15% SDS-polyacrylamide gel samples were stained with Stains-All, then excised, transferred to a 96 well plate and destained. The gel samples were then subjected to reduction and alkylation, then washed, dehydrated, and digested with trypsin using a MassPrep robot. The peptides were extracted from the gel plugs with 2% acetonitrile and 1% formic acid. The extracted peptides (30 µl) were transferred to another 96 well plate, where 5 µl of matrix (α-Cyano) was added to the sample well. The samples were then vaporized to dryness and redissolved in 5 µl of 60% acetonitrile and 0.1% TFA. Peptide samples were then spotted on a MALDI-TOF/TOF target plate for MS and MSMS analyses. MSMS or tandem mass spectrometry is a mass spectrometric method in which a peptide is fragmented and the masses of the resultant fragment ions are recorded in a spectrum. The analyses were performed using the ABI 4800 Maldi TOF/TOF (Applied Biosystems, Foster City, CA) at the Michigan Protein Consortium.
Searches for homologies between the amino acid sequences obtained and those of other known proteins in GenBank, GenPept, SwissProt were performed using BLAST software.

The Michigan Proteome Consortium provided proteomics data at the University of Michigan.

**Gelatin zymography**

Zymogram gels, prepared with 7.5% or 10% SDS polyacrylamide containing 0.1% gelatin, were used to detect and characterize protease activity in gel purified and eluted DSP-PP\textsubscript{240} precursor protein and PP\textsubscript{240} protein samples. Protease activity was also examined in the purified native rat HP. Following electrophoresis, the gel was washed in renaturing buffer (50mM Tris-HCl, 5mM CaCl\textsubscript{2}, 2.5% Triton X100, 0.02% NaN\textsubscript{3}, pH 7.5) with gentle agitation for 30 minutes at room temperature, equilibrated with developing buffer (50mM Tris-HCl, 5mM CaCl\textsubscript{2}, 1% Triton X100, 0.02% NaN\textsubscript{3}, pH 7.5) for 30 min, replaced with fresh developing buffer, and incubated at 37°C overnight. The gel was then stained with Coomassie Blue, de-stained and digitally scanned.

**Results**

*Identification of baculovirus-derived recombinant DSP-PP\textsubscript{240} and DSP\textsubscript{370} protein profiles*

Both DSP-PP\textsubscript{240} and DSP\textsubscript{370} baculovirus constructs contained leader sequences, thus the baculovirus-derived recombinant proteins would be expected to be secreted into the insect culture medium. Insect sf9 cells were infected with baculovirus containing either DSP-PP\textsubscript{240} cDNA or DSP\textsubscript{370} cDNA. After infection, the cells were incubated for 4 days and the recombinant proteins in the harvested cell media were partially purified and concentrated using polyanion extraction, separated by SDS-PAGE, and stained with Stains-All (see Methods). The resulting protein profiles are shown in Figure 2. Five major blue-staining bands (i.e., bands 1-5; BSA stains orange) were present in media obtained from DSP-PP\textsubscript{240} bacteriophage infected cells, while 3 major blue-staining bands (i.e., bands 1’-3’; BSA stains orange red) were present in the DSP\textsubscript{370} cell infected medium.

To characterize the identities of the 120 kDa (band 1) and 95 kDa (band 2) protein bands, the bands were transferred to a PVDF membrane, excised and subjected to Edman degradation for N-terminal amino acid sequence analyses. The N-terminal sequence for the 120 kDa band was determined to be IPVPQ which corresponded correctly to that of the presumed DSP-PP\textsubscript{240} precursor protein. The N-terminal sequence for the 95 kDa band was determined to be IPVPQ, suggesting that this band represented DSP.

We further determined the identity of band 1 with mass spectra analyses. DSP-PP\textsubscript{240} cDNA encodes a 687 amino acid peptide, which contains a 17 amino acid signal peptide, as well as DSP and PP peptide sequences (see Figure 3B). Mass spectra analyses of band 1, following trypsin digestion, identified a number of tryptic peptides across the presumed DSP-PP\textsubscript{240} precursor protein (see Figure 3A and 3B). For example, MS
analyses detected peptide \textsubscript{18-28} (i.e., IPVPQLVPLER, corresponding to amino acid sequence positions 18-28; Figure 3B), the actual N-terminal 11 amino acid sequence of DSP-PP\textsubscript{240} precursor protein. Additional MS peptide sequences corresponding to DSP protein that we identified included peptide\textsubscript{70-79} (i.e., QVHSNGGYER, corresponding to amino acid sequence positions 70-79); peptide\textsubscript{93-109} (i.e., SSPTQPLANAQGNSAK, corresponding to amino acid sequence positions 93-109); peptide\textsubscript{136-148} (i.e., GQVGIAENAEEAK, corresponding to amino acid sequence positions 136-148); peptide\textsubscript{266-286} (i.e., ESHDGTEGHEQSSGGNNDNR, corresponding to amino acid sequence positions 266-286); peptide\textsubscript{287-308} (i.e., GQGSVSTEDDDSKEQEGSPNGR, corresponding to amino acid sequence positions 287-308); and peptide\textsubscript{385-397} (i.e., DSNGHHGMLCKKR, corresponding to amino acid sequence positions 385-397). The N-terminal sequence of mature PP is DDPN, located at amino acid positions 448-451. And the detection of a PP matching amino acid sequence at positions 524-542 (i.e., DKDESDNSHDNDSDESK). Thus band 1 encompasses the DSP coding sequence starting with the N-terminal sequence IPVPQ and an additional 6 peptides (located between the DSP N-terminal and PP N-terminal sequences) as well as the PP sequence.

To identify the 33 kDa band 4 in Figure 2, the recombinant protein profile generated from a DSP\textsubscript{370} cDNA construct was compared to that generated from a DSP-PP\textsubscript{240} cDNA construct (Fig. 2, lanes 1 and 2). The DSP\textsubscript{370} cDNA construct encodes a recombinant protein with a size of 388 amino acids (i.e., 370 amino acid DSP protein plus an additional 18 amino acid viral derived sequence). Band 2' represents the 388 amino acid DSP protein, which was recognized by anti-DSP antibody (not shown). This DSP\textsubscript{370} recombinant protein does not contain a PP sequence and is shorter than band 2. From Figure 2, band 4 (33 kDa) and band 5 (29 kDa) are not present in the DSP\textsubscript{370} profile. Thus, these bands most likely represent two PP related proteins.

To confirm that band 4 represented PP\textsubscript{240} we again used MS analyses. Because the N-terminal sequence of the deduced PP\textsubscript{240} protein is DDPN as indicated in Figure 3C, we expect that among the trypsinized fragments we should see peptide\textsubscript{524-542} (located from amino acid 524-542; 19 amino acids), peptide\textsubscript{660-687} (28 amino acids), peptide\textsubscript{448-523} (76 amino acids), and peptide\textsubscript{543-647} (105 amino acids). Mass spectra analyses of band 4 detected both peptide\textsubscript{524-542} (i.e., DKDESDNSHDNDSDESK) and peptide\textsubscript{660-687} (i.e., SGNGNSDSDSDSDSDEGSASNHSSTSD) (see Figure 3C). However, because the mass spectra detection ranges from 10-30 amino acids, it is therefore understandable that peptide\textsubscript{448-523} and peptide\textsubscript{543-647} were beyond the MS and MSMS detection range. However, we did observed a mass with a molecular weight of ~7700 Da, which agrees well with that of peptide\textsubscript{448-523}. From these mass spectra data, we conclude that band 4 represents PP\textsubscript{240}.

In contrast to the MS obtained from band 4, MS spectra from band 5 only detected peptide\textsubscript{524-542} (i.e., DKDESDNSHDNDSDESK) and did not detect the last 28 amino acid peptide\textsubscript{660-687} (i.e., SGNGNSDSDSDSDSDEGSASNHSSTSD). Thus band #5 most likely represents PP\textsubscript{211}, which is missing the C-terminal 28 amino acids (see Figure 3D). The calculated molecular weight ratio of PP\textsubscript{240} (blue band #4) and PP\textsubscript{211} (blue band #5) is 1.14. The apparent molecular weight ratio of blue band #4 (33 kDa) and blue band
#5 (29 kDa) is 1.14. Taken together, band #4 likely represents the full length PP₂₄₀ and band #5 represents PP₂₁₁.

As shown in Figure 2, blue band 2 present in DSP-PP₂₄₀ lane has a higher molecular weight than blue band 2' (containing 388 amino acids) present in the DSP₃₇₀ lane. Thus this blue band 2 likely represents the DSP₄₃₀ protein, which is comprised of DSP₅₅₀ and dentin glycoprotein (DGP₈₀). In Figure 2, blue band 1' (180 kDa), present in the DSP₃₇₀ lane, most likely represents a dimer of DSP (band 2'; 90 kDa). Blue band 3 (70 kDa) and blue band 3', present in both DSP-PP₂₄₀ and DSP₃₇₀ lanes respectively in Figure 2, were not always present in recombinant protein samples. Mass spectra analyses suggest that blue band 3 is a telokin-like-20 protein (i.e., baculovirus related, data not shown). Thus, blue band 3 could be produced by the baculovirus system. Smaller weak blue bands between 22 kDa and 6 kDa were present in both DSP-PP₂₄₀ derived products and DSP₃₇₀ derived products. These bands are likely derived from DSP.

Dynamic processing of DSP-PP₂₄₀ protein

Sf9 cells were infected for 4 days with baculovirus containing DSP-PP₂₄₀ cDNA. At various times following infection, recombinant DSP-PP₂₄₀ and related proteins were purified from the culture medium and run on a 10% SDS-PAGE, which was then stained with Stains-All. After 2 days infection, bands located at 120 kDa (i.e., DSP-PP₂₄₀) and 33 kDa (i.e., PP₂₄₀) were observed (Fig. 4A, lane 2). These two bands increased in intensity on days 3 and 4 (Fig. 4A, lanes 3 and 4), with two additional minor bands appearing at 95 kDa (i.e., DSP) and 70 kDa (equivalent to blue band 3 in Figure 2) on day 4. A Western blot, using anti-DSP antibodies of day 4 infected media, recognized both the 120 kDa and 95 kDa bands, confirming that they both contain DSP.

The purified DSP-PP₂₄₀ precursor protein can undergo processing in the absence of insect cell condition medium

Our results thus far demonstrate that recombinant DSP-PP₂₄₀ protein can be processed to produce both DSP₄₃₀ and PP₂₄₀ peptides over time (see Figure 4). To determine whether this processing was due to proteases that were present in the insect culture medium, we obtained a SDS-PAGE purified DSP-PP₂₄₀ protein via electro-elution (see Methods) and incubated this purified precursor in 25 mM Tris·HCl, pH 7.5 for 30 minutes at 37°C. The resulting protein profile is shown in Figure 5. The initial electro-eluted DSP-PP₂₄₀ is present as a single band (Figure 5, lane 1), which then undergoes significant protein processing within 30 minutes to yield a DSP₄₃₀ band and a PP₂₄₀ band (Figure 5, lane 2). Thus DSP-PP₂₄₀ precursor protein processing is likely not due to proteases in the insect medium suggesting that it may instead undergo self-processing.

These findings are summarized in Figure 6. Here we show that the DSP-PP₂₄₀ precursor protein (120 kDa), following secretion into the culture medium, was further processed into a 95 kDa DSP protein (containing 430 amino acids), and a major 33 kDa PP protein (containing 240 amino acids). The DSP₃₇₀ protein profile contains a secreted 90 kDa protein (containing 388 amino acids with 370 DSP amino acids and 18 viral amino acids;
band 2') and no PP bands. Figure 6 also displays the cleavage site responsible for removing the leader sequence from the DSP-PP$_{240}$ precursor protein and the cleavage site responsible for generating DSP$_{430}$ and PP$_{240}$.

**DSP-PP$_{240}$ has gelatinolytic activity**

Our results demonstrate that DSP-PP$_{240}$ precursor protein is capable of self-processing to yield both DSP$_{430}$ and PP$_{240}$. To test whether this process could be due to proteolysis, gel purified DSP-PP$_{240}$ and PP$_{240}$ were electrophoresed on 10% SDS-PAGE gels containing 0.1% gelatin as a proteolytic substrate. Following electrophoresis, the gels were incubated overnight in renaturing buffer (see Methods), then stained with Coomassie blue. The destained gels showed two white bands at 120 kDa (i.e., DSP-PP$_{240}$) and 33 kDa (i.e., PP$_{240}$) (see Figure 7A, Lane 1 and Figure 7B, lane 1) indicative of gelatinolytic activity. In the control groups, no white bands were observed in polyanion extracts of sf9 insect cell condition medium and in polyanion extracts of media derived from sf9 insect cells infected with the baculovirus containing anti-sense DSP$_{370}$ cDNA (see Figure 7A, lanes 2 and 3, and Figure 7B, lane 2). This is the first evidence that DSP-PP and PP possess proteolytic activity.

**Native Highly Phosphorylated protein (HP) also exhibits protease activity**

Since DSP-PP$_{240}$ and PP$_{240}$ possess proteolytic activity as shown in Figure 7, purified rat incisor native HP protein (23) was next tested to determine whether naturally occurring HP was capable of digesting gelatin. HP is equivalent to isoform PP$_{523}$, which is derived from the DSP-PP$_{523}$ transcript. Rat HP, isolated using the polyanion extraction method, was analyzed by Edman degradation. Its N-terminal sequence was identified as DDPNSSDESNGSD (24), indicating that it was free of contamination by other proteins. As shown in Figure 8, lanes 2 and 3, both native HP and heat denatured HP were stained blue by Stains-All staining. Coomassie blue did not stain native HP or heat denatured HP (Figure 8, lanes 5 and 6). When native rat HP was electrophoresed along with heat denatured HP on a 7.5% SDS-PAGE gel containing 0.1% gelatin, after 3 hour incubation in renaturing buffer, followed by Coomassie staining single proteolytic bands were observed around 90 kDa for native rat HP but not for heat denatured HP (Figure 8, lanes 7 and 8). Furthermore, under the same conditions, but in the presence of 0.1 M EDTA, no clear band was detected in either the native HP or denatured HP lanes (Figure 8, lanes 9 and 10). Taken together, these data demonstrate that both PP$_{240}$ and PP$_{523}$ exhibit proteolytic activity.

**Discussion**

To date, three different DSP-PP transcripts (i.e., DSP-PP$_{523}$, DSP-PP$_{240}$, and DSP-PP$_{171}$), giving rise to 3 PP isoforms, have been identified in rat tooth extracts (1,18,19). These PP isoforms, which include PP$_{523}$, PP$_{240}$, and PP$_{171}$, are speculated to play different roles during tooth development and mineralization by helping to fine-tune mineral nucleation and hydroxyapatite growth at different stages of the mineralization program (19). We used DSP-PP$_{240}$ infected sf9 insect cells to produce and secrete a recombinant...
120 kDa protein product into the conditioned cell medium (band 1, Fig. 2). MS and MSMS analysis identified a number of recognizable tryptic peptides across the DSP portion of the presumed DSP-PP\textsubscript{240} precursor protein, and the detection of a PP\textsubscript{240} matching amino acid sequence at positions 524-542 (i.e., DKDESDNSHNDSDE SESK), demonstrated that the 120 kDa protein band contains both DSP and PP\textsubscript{240} sequences. MS identification of the N-terminal DSP-PP\textsubscript{240} amino acid sequence IPVPQLVPLER confirmed that (i) it is likely that the signal peptide sequence (i.e., MKTKIIYICIWATAWA) was cleaved from the nascent DSP-PP\textsubscript{240} peptide in the endoplasmic reticulum during the secretory process, and (ii) that the DSP-PP\textsubscript{240} precursor protein was secreted into the extracellular medium. These data demonstrate for the first time that DSP-PP\textsubscript{240} transcripts are capable of producing and secreting full-length DSP-PP\textsubscript{240} proteins into the extracellular space. Moreover, we were able to produce recombinant DSP-PP\textsubscript{240} precursor protein in sufficient quantities to allow us to follow its processing over time using standard SDS-PAGE followed by Stains-All staining.

**In vitro DSP-PP\textsubscript{240} protein processing**

When we infected Sf9 cells with baculovirus containing DSP-PP\textsubscript{240} cDNA, we were able to identify not only the DSP-PP\textsubscript{240} precursor protein in the conditioned medium, but also found DSP\textsubscript{430}, PP\textsubscript{240} and minor amounts of PP\textsubscript{211}. During the 4 day incubation period both DSP-PP\textsubscript{240} and PP\textsubscript{240} bands increased in intensity. DSP\textsubscript{430} appeared on days 3 and 4, but was significantly weaker than the PP\textsubscript{240} band (Fig. 4A). Interestingly, dentin DSP protein has been estimated at 5-8% of the dentin NCP content and PP has been estimated to be >50% of dentin NCP content (16,17). Therefore the actual DSP/PP ratio in the dentin matrix is estimated to be 1:6, rather than the expected 1:1 ratio. Using the NIH image J program, we determined the relative densities of recombinant DSP\textsubscript{430} and PP\textsubscript{240} from day 4 culture medium to be ~1:6 (see Figure 4B). Thus, our *in vitro* DSP-PP processing results agree very well with the measured DSP/PP ratio in dentin tissue.

The cleavage of DSP-PP in baculovirus conditioned medium initially prompted us to consider that dentin matrix metalloproteases (MMPs) might be responsible for this proteolytic activity since Yamakoshi, et al, (25) recently reported on the ability of dentin resident MMPs to cleave DSP-DGP, where DSP-DGP is a proteoglycan having 457 amino acids. According to these authors, porcine DSP-DGP-PP is first cleaved on the N-terminal side of ASP\textsuperscript{458} to split DSP-DGP (equivalent to our DSP\textsubscript{430}) from PP. They also claimed that this cleavage is rapid since they were unable to detect intact DSP-DGP-PP protein in the dentin matrix. Without DSP-DGP-PP to use as a substrate, they were unable to identify the protease responsible for the proposed cleavage at ASP\textsuperscript{458}. However, they were able to identify 12 different cleavage products from developing porcine molars by N-terminal sequencing. They then compared these fragments with fragments generated from DSP-DGP digested under *in vitro* conditions with either MMP-2 or MMP-20. They found that both MMP-2 and MMP-20 were capable of cleaving DSP-DGP at similar specific sites *in vitro* to those identified...
from *in vivo* isolations of low molecular weight DSP and DGP. Thus, they concluded that MMP-20 cleaved DSP-DGP from both ends and MMP-2 cleaved DSP-DGP within the DSP C-terminal region as well as within the DGP region. However, as shown by Jo, et al., the baculovirus-sf9 insect-cell system expresses neither gelatinolytic MMP2 or TIMP-2 (26). Furthermore, during our polyanion extraction of sf9-conditioned culture medium, we found no 72 kDa or 60-65 kDa protein bands stained with Coomassie blue that would suggest the presence of MMP-2. And we were able to show that SDS-PAGE isolated DSP-PP240 could undergo cleavage in a Tris-buffered salt solution. Thus it is unlikely that MMP2 participates in DSP-PP240 precursor protein cleavage in the baculovirus system.

**DSP-PP240, PP240, and HP (PP523) proteolytic activity**

Our in vitro studies, using purified recombinant DSP-PP240, demonstrate that the major cleavage products are DSP430 and PP240 (see Figure 5). Edman degradation, Western blot analysis and comparison of the SDS-PAGE protein profiles derived from DSP-PP240 cDNA and DSP370 cDNA transfected sf9 cells support our findings that DSP430 and PP240 are products of the proteolytic cleavage of DSP-PP240. As mentioned above, our DSP430 is equivalent to porcine DSP-DGP. We found that no DSP350 and no DGP cleavage products were produced during the 30-minute incubation time used to cleave our purified DSP-PP240 precursor protein. We also found that no DSP350 and no DGP cleavage products were produced in 4d culture medium. Using 0.1% gelatin gels, we also found that DSP-PP240, PP240, as well as native HP (i.e., PP523), were capable of degrading gelatin.

Because PP does not stain with Coomassie Blue, there is a possibility that the concentrated PP on the gelatin gel might yield a clear band. To test whether after Coomassie Blue staining, the clear HP band present on the gelatin gel was caused by the inability of HP to be stained by Coomassie Blue, we ran both native HP and heat denatured HP on a 0.1% gelatin gel. As shown in Figure 8, lane 7, after 3 hours incubation in renaturing buffer, 0.5 µg of native HP showed a clear band in gelatin zymography while heat denatured HP displayed no clear band (Figure 8, lane 8). This data suggests that native HP possesses proteolytic activity which can be inactivated by heating at 95°C for 5 minutes. Furthermore, when both native HP and heat denatured HP were incubated with renaturing buffer in the presence of 0.1 M EDTA, no clear bands were detected in gelatin zymography (Fig. 8, lanes 9 and 10). This data also demonstrates that EDTA can inhibit HP proteolytic activity. These studies demonstrate that the HP clear band detected in gelatin zymography is due to HP proteolytic activity and not due to the inability of Coomassie blue to stain HP.

Taken together, these studies suggest that newly synthesized DSP-PP precursor proteins (derived from three DSP-PP multiple transcripts) undergo a rapid self-processing step, which yields DSP430 plus the associated PP isoforms (i.e., PP171, PP240 and PP523). Perhaps, over a longer time period, MMP-2
present in dentin, then acts on DSP\textsubscript{430} to yield DSP\textsubscript{350} and DGP\textsubscript{80}.

**Developmental implications**

DSP-PP promoter-driven LacZ expression appears in kidney, in postnatal day 3 transgenic mice (15), in alveolar bone in newborn mice prior to its appearance in the incisor (15), and in salivary glands obtained from newborn mice (Ritchie, unpublished). Our current data, demonstrating that DSP-PP\textsubscript{240}, PP\textsubscript{240} and PP\textsubscript{523} all exhibit proteolytic activity suggests that these proteins may play important roles in tissue modeling during organ development.

**Acknowledgement**

We want to express special thanks to Dr. David G. Ritchie for helpful discussion and critiques on this manuscript. We thank Mr. Ryan Adams and Mr. Ke Wan for technical support. We thank Dr. Mary E. Marsh for kindly providing rat HP. Proteomics data were provided by the Michigan Proteome Consortium (www.proteomeconsortium.org) which is supported in part by funds from the Michigan Life Sciences Corridor (State of Michigan, MEDC Grant #GR239). We want to thank Ms. Mary Hurley, and Dr. Maureen Kachman at the University of Michigan Protein Consortium for MS and MSMS analyses. This work is supported by DE11442 (H.H.R.).
References
1. Ritchie, H. H., and Wang, L.-H. (1996) *J Biol Chem* **271**, 21695-21698
2. Ritchie, H. H., and Wang, L.-H. (1997) *Biochem Biophys Res Co* **231**, 425-428
3. MacDougall, M., Simmons, D., Luan, X., Nydegger, J., Feng, J., and Gu, T. T. (1997) *J Biol Chem* **272**, 835-842
4. Gu, K., Chang, S., Ritchie, H., Clarkson, B., and Rutherford, R. (2000) *Eur J Oral Sci* **108**(1), 35-42
5. D’Souza, R. N., Bronckers, A. L. J. J., Happonen, R.-P., Doga, D. A., Farach-Carson, M. C., and Butler, W. T. (1992) *J. Histochem. Cytochem.* **40**, 359-366
6. Bronckers, A., Farach-Carson, M., Wavern, E., and Butler, W. (1994) *J Bone Miner Res* **9**, 833-841
7. Ritchie, H. H., Shigeyama, Y., Strayhorn, C., Hanks, C. T., Somerman, M. J., and Butler, W. T. (1996) *J Dent Res* **75**, 153
8. Ritchie, H. H., Berry, J. E., Somerman, M. J., Hanks, C. T., Bronckers, A. L. J. J., Hotton, D., Papagerakis, P., Berdal, A., and Butler, W. T. (1997) *Eur J Oral Sci* **105**, 405-413
9. Xiao, S., Yu, C., Chou, X., Yuan, W., Wang, Y., Bu, L., Fu, G., Qian, M., Yang, J., Shi, Y., Hu, L., Han, B., Wang, Z., Huang, W., Liu, J., Che, Z., Zhao, G., and Kong, X. (2001) *Nat. Genet.* **27**, 201-204
10. Zhang, X., Zhao, J., Li, C., Gao, S., Qiu, C., Liu, P., Wu, G., Qiang, B., Lo, W., and Shen, Y. (2001) *Nat. Genet.* **27**, 151-152
11. Sreenath, T., Thyagarajan, T., Hall, B., Longenecker, G., D’Souza, R., Hong, S., Wright, J., MacDougall, M., Sauk, J., and Kulkarni, A. (2003) *J Biol Chem* **278**, 24874-24880
12. Veis, A. (1985) *In Chemistry and Biology of Mineralized Tissues, (Butler, W.T., ed), pp. 170-184, EBSCO Media, Birmingham, AL, USA.*
13. Linde, A. (1985) *In Chemistry and Biology of Mineralized Tissues, (Butler, W.T., ed), pp. 344-355, EBSCO Media, Birmingham, AL, USA.*
14. Linde, A. (1989) *Anatomical Record* **224**, 154-156
15. Godovikova, V., Li, X.-R., Saunders, T., and Ritchie, H. (2006) *Development Biology* **289**, 507-516
16. Butler, W. T., Bhowm, M., D’Souza, R. N., Farach-Carson, M. C., Happonen, R. P., Schrohenloher, R. E., Seyer, J. M., Somerman, M. J., Foste, R. A., Tomana, M., and van Dijk, S. (1992) *Matrix* **12**, 343-351
17. Linde, A., Bhowm, M., and Butler, W. (1980) *J. Biol. Chem.* **255**, 5931-5942
18. Ritchie, H. H., and Wang, L.-H. (2000) *Biochim Biophys Acta* **1493**, 27-32
19. Ritchie, H. H., Wang, L. H., and Knudtson, K. (2001) *Biochim Biophys Acta* **1520**, 212-222
20. Crossen, R., and Gruenwald, S. (1996) *3rd. Edition*
21. Marsh, M. E. (1989) *Biochemistry* **28**, 339-345
22. Butler, W. (1987) *Meth. Enzymol.* **145**, 290-303
23. Marsh, M. E. (1989) *Biochemistry* **28**, 346-352
24. Chang, S., Chiego, D., and Clarkson, B. (1996) *Calcif Tissue Int* **59**, 149-153
25. Yamakoshi, Y., Hu, J., Iwata, T., Kobayashi, K., Fukae, M., and Simmer, J. (2006) *J Biol Chem* **281**, 38235-38243
26. Jo, Y., Yoon, D., Kim, M., Lee, Y., Kim, H., and Lee, S. (1999) *J Biochem Mol Biol* 32, 60-66
Figure legends

**Figure 1.** Baculovirus constructs. A: Rat DSP-PP gene structure and the relative position of DSP-PP240 and DSP370 cDNAs. The rat DSP-PP gene is comprised of five exons (i.e., E1, E2, E3, E4, E5) and four introns (I1, I2, I3, I4). The translation start site is designated as ATG. Arrows indicate the locations of the DSP N-terminal and PP N-terminal sequences. Light gray bars, DSP exons; black bar, PP exon; white bar, 3’ noncoding sequence. Stop codon is represented by . Two polyadenylation sites (aataaa) are represented by ●. B: baculovirus pVL1392 containing DSP-PP240 cDNA, which was inserted at EcoRI and BamHI sites. C: baculovirus pVL941 containing DSP370 cDNA was inserted at BamHI site. This DSP370 construct generated a recombinant protein containing a 370 amino acid DSP sequence and an additional 18 amino acid viral-derived sequence.

**Figure 2.** Recombinant Proteins derived from DSP-PP240 and DSP370 baculovirus infected sf9-condition medium. Recombinant proteins were partially purified using polyanion extraction (see Methods), electrophoresed on a 4-15% SDS-PAGE gradient gel and stained with Stains-All. Lane M: Protein size markers. Lane DSP-PP240: DSP-PP240 baculovirus infected sf9-condition medium. Lane DSP370: DSP370 baculovirus infected sf9-conditioned medium. Blue stained bands represent secreted proteins while the orange-red bands represent serum-derived proteins.

**Figure 3.** Mass spectrometry analyses of DSP-PP240 recombinant protein blue bands #1, #4 and #5. A: Mass spectra data for blue bands #1, #4 and #5. B: DSP-PP240 cDNA deduced amino acid peptide sequence. This deduced peptide sequence contains a 17 amino acid signal peptide sequence (i.e., indicated by italicized letters), and DSP and PP amino acid sequences. The locations of MS/MS identified peptide sequences for band #1 are labeled in bold. C: Peptide sequence of mature PP240. The locations of MS/MS identified peptide sequences for band #4 are labeled in bold. D: Peptide sequence of PP221. The locations of MS/MS identified peptide sequences for band #5 are labeled in bold.Italicized and underlined bold letters DDPN represent the N-terminal PP sequence. The trypsin cleavage sites are located C-terminal to the underlined amino acids R and K.

**Figure 4.** Dynamic processing of recombinant proteins. A: Sf9 cells were infected with baculovirus containing DSP-PP240 cDNA. Sf9-conditioned medium was removed on days 1, 2, 3, and 4 and protein samples were prepared and subjected to SDS-PAGE using 10% gels (see Methods). The gels were then stained with Stains-All (lanes 1-4) or directly blotted onto a nitrocellulose membrane. Western blot (lanes 5-6) was performed with anti-DSP antibody (see Methods). Lane 5: day 4 Sf9 DSP-PP240 virus infected sample. Lane 6 control: day 4 Sf9 wild type baculovirus infected sample. B: The expression levels of recombinant DSP430 and PP240 proteins from day 4 Sf9 DSP-PP240 virus infected sample as determined by NIH image program.
Figure 5. Recombinant DSP-PP<sub>240</sub> precursor protein undergoes self-processing to generate DSP<sub>430</sub> and PP<sub>240</sub>. Purified DSP-PP<sub>240</sub> precursor protein was obtained by gel-purification and electro-elution. The purified DSP-PP<sub>240</sub> precursor protein was incubated in 25 mM Tris.HCl pH 7.5 at 37°C for 30 min. The samples were then electrophoresed on a 10% SDS PAGE and stained with Stains-All. Lane M: protein size marker. Lane 1: purified DSP-PP<sub>240</sub> precursor protein. Lane 2: DSP-PP<sub>240</sub> precursor protein incubated at 37°C for 30 min.

Figure 6. Production and processing of recombinant DSP-PP<sub>240</sub> and DSP<sub>370</sub> proteins in a baculovirus expression system. Panel A: Recombinant DSP-PP<sub>240</sub> precursor protein, containing a leader sequence, was synthesized from DSP-PP<sub>240</sub> cDNA in a baculovirus expression system. The leader sequence was removed and DSP-PP<sub>240</sub> precursor protein was secreted into the culture medium (see Figure 2). During a 30 minute incubation at 37°C in 25 mM Tris-HCl pH 7.5, the DSP-PP<sub>240</sub> precursor protein was further processed into DSP<sub>430</sub> and PP<sub>240</sub> (see Figure 5). The DSP-PP<sub>240</sub> precursor protein leader sequence is depicted in gray and the N-terminal sequence is depicted in black bold letters as represented by MKIKIIYICWIWATAWAIPVPQLVPLERDI. The gray triangle indicates the cleavage site responsible for leader sequence removal. SYDEDDESMQDDPNSDDSN represents the junction amino acid sequence between DSP and PP sequence. Within this sequence, bold letters represent amino acids from the DSP sequence and italic letters represent amino acids from the PP sequence. The black triangle indicates the cleavage site responsible for generating DSP<sub>430</sub> and PP<sub>240</sub>. IPVPQL.

DDPNSS. ............... SYDEDDESMQ represents DSP<sub>430</sub> N-terminal and C-terminal sequences. SYDEDDESMQDDPNSDDSN represents DSP<sub>430</sub> N-terminal and C-terminal sequences. Panel B: A recombinant DSP<sub>370</sub> protein-viral sequence (with a leader sequence) was synthesized from DSP<sub>370</sub> cDNA in a baculovirus expression system. The leader sequence was removed and DSP<sub>370</sub> protein-viral sequence was secreted into the culture medium. The size of DSP<sub>370</sub> protein-viral sequence (i.e., 388 amino acids) is smaller than DSP<sub>430</sub> and no PP related bands were present in the recombinant DSP<sub>370</sub> protein gel profile (see Figure 2).

Figure 7. Gelatin zymography of recombinant DSP-PP<sub>240</sub> precursor protein and recombinant PP<sub>240</sub> protein. A: Recombinant DSP-PP<sub>240</sub> precursor protein obtained by gel purification and electro-elution was loaded onto a 10% SDS PAGE containing 1 mg/ml gelatin. Lane 1: recombinant DSP-PP<sub>240</sub> precursor protein. Lane 2: polyanion protein extracts derived from insect cell condition medium as a control. Lane 3: polyanion protein extracts derived from baculovirus containing anti-sense DSP<sub>370</sub> cDNA as a control. B: Recombinant PP<sub>240</sub> protein obtained by gel purification and electro-elution was loaded onto a 10% SDS PAGE gel containing 1 mg/ml gelatin. Lane 1: recombinant PP<sub>240</sub> protein. Lane 2: polyanion protein extracts derived from baculovirus containing anti-sense DSP<sub>370</sub> cDNA as a control. Following electrophoresis protease activity was detected as described in Methods.
Figure 8. Stains-All staining, Coomassie blue staining and gelatin zymography of rat native highly phosphorylated protein (HP). Column purified native rat HP (a PP523 equivalent isoform protein) and heat denatured HP were electrophoresed on 7.5% SDS-PAGE or 7.5% SDS-PAGE-gelatin gels. Lanes 1-3: Stains-All staining. Protein size markers (Lane 1), native rat HP (0.5 µg, Lane 2), heat denatured (95°C, 5 min) rat HP (0.5 µg, Lane 3) were electrophoresed on a 7.5% SDS-PAGE gel and stained with Stains-All. Lanes 4-6: Coomassie blue staining. Protein size markers (Lane 4), native rat HP (0.5 µg, Lane 5), heat denatured (95°C, 5 min) rat HP (0.5 µg, Lane 6) were electrophoresed on a 7.5% SDS-PAGE gel and stained with Coomassie blue. Lanes 7-8: Coomassie blue staining. Native rat HP (0.5 µg ; Lane 7) and heat denatured (95°C, 5 min) rat HP (0.5 µg ; Lane 8) were electrophoresed on a 7.5% SDS-PAGE-gelatin gel, incubated for 3 hours with renaturing buffer (see Methods), then stained with Coomassie blue. Lanes 9-10: Coomassie blue staining. Native rat HP (0.5 µg , Lane 9) and heat denatured (95°C, 5 min) rat HP (0.5 µg , Lane 10) were electrophoresed on a 7.5% SDS-PAGE-gelatin gel, incubated for 3 hours with 0.1 M EDTA and renaturing buffer, then stained with Coomassie blue.
Fig. 1A
Fig. 1B
Fig. 1C
| Band | Ion Score | C.I. % | Ion Score | Observed Mass | PPM | Match Error | Best Peptide Sequence | Calculated Mass | Match Error Da |
|------|-----------|--------|-----------|--------------|-----|-------------|-----------------------|----------------|----------------|
| 1    | 99.97     | 47     | 1260.78   | 1260.78      | 9   | 0.01        | IPVPQLVPLER           | 1260.77        |                |
| 1    | 99.97     | 47     | 1146.55   | 1146.55      | 23  | 0.03        | QVHSNGGYER            | 1146.53        |                |
| 1    | 93.32     | 25     | 1129.52   | 1129.52      | 17  | 0.02        | QVHSNGGYER            | 1129.50        |                |
| 1    | 100.00    | 73     | 1683.89   | 1683.89      | 13  | 0.02        | SSPTQPIALANAQNSAK     | 1683.87        |                |
| 1    | 99.99     | 55     | 1315.67   | 1315.67      | 18  | 0.02        | GQVIAEAAEAK           | 1315.65        |                |
| 1    | 100.00    | 83     | 2183.98   | 2183.98      | 4   | 0.02        | ESHDGTEGHEGQSSGNNNDNR | 2183.86        |                |
| 1    | 100.00    | 114    | 2137.80   | 2137.80      | 14  | 0.03        | DKDESNSNHNDSDSSEK     | 2137.80        |                |
| 1    | 79.68     | 20     | 1511.74   | 1511.74      | 49  | 0.07        | DSNGHGMELDKR          | 1511.67        |                |
| 1    | 100.00    | 61     | 2137.80   | 2137.80      | 14  | 0.03        | DKDESNSNHNDSDSSEK     | 2137.80        |                |
| 4    | 100.00    | 61     | 2137.80   | 2137.80      | 14  | 0.03        | DKDESNSNHNDSDSSEK     | 2137.80        |                |
| 4    | 97.92     | 28     | 2779.00   | 2779.00      | 36  | 0.10        | SGNGNSDSDDSDSDDSDSEGSDSNSHNTSSDD | 2779.90 |                |
| 5    | 99.99     | 54     | 2137.80   | 2137.80      | 0   | 0.00        | DKDESNSNHNDSDSSEK     | 2137.80        |                |

Figure 3
Figure 5
Figure 6
Fig. 7
Dynamic processing of recombinant dentin sialoprotein-phosphophoryn protein
Valentina Godovikova and Helena H. Ritchie

J. Biol. Chem. published online August 13, 2007

Access the most updated version of this article at doi: 10.1074/jbc.M702605200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts