Evidence for the Proteolytic Processing of Dentin Matrix Protein 1: Identification and Characterization of Processed Fragments and Cleavage Sites*

Chunlin Qin¹§, Jan C. Brunn¹, Richard G. Cook²,³, Ralph S. Orkiszewski³, James P. Malone⁴, Arthur Veis⁴ and William T. Butler¹

From ¹the Department of Basic Sciences, The University of Texas–Houston Health Science Center, Dental Branch, 6516 M.D. Anderson Blvd., DBB, Rm 4.133, Houston, Texas 77030
²the Department of Immunology and ³the Protein Chemistry Core Laboratory, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030
⁴the Department of Cell and Molecular Biology, Northwestern University, Feinberg Medical School, Ward Bldg. 13-100, 303 E. Chicago Avenue, Chicago, Illinois 60611

Running Title: Proteolytic Processing of DMP1

§Correspondence address:

Chunlin Qin Ph.D.
Department of Basic Sciences,
The university of Texas–Houston Health Science Center, Dental Branch
M.D. Anderson Blvd., DBB, Rm 4.133, Houston, TX 77030, USA
Phone: 713–500–4583; Fax: 713–500–4500, Email Address: Chunlin.Qin@uth.tmc.edu
SUMMARY

Full length cDNA coding for dentin matrix protein 1 (DMP1) has been cloned and sequenced, but the corresponding complete protein has not been isolated. In searching for naturally occurring DMP1, we recently discovered that the extracellular matrix of bone contains fragments originating from DMP1. Shortened forms of DMP1, termed 37K and 57K fragments, were treated with alkaline phosphatase and then digested with trypsin. Resultant peptides were purified by a two-dimensional method: size-exclusion followed by reversed-phase high performance liquid chromatography. Purified peptides were sequenced by Edman degradation and mass spectrometry, and the sequences compared to the DMP1 sequence predicted from cDNA. Extensive sequencing of tryptic peptides revealed that the 37K fragments originated from the NH$_2$-terminal region, and the 57K fragments were from the COOH-terminal part of DMP1. Phosphate analysis indicated that the 37K fragments contained 12 phosphates and the 57K fragments had 41. From 37K fragments, two peptides lacked a COOH-terminal lysine or arginine; instead they ended at Phe$^{173}$ and Ser$^{180}$ and were thus COOH-termini of 37K fragments. Two peptides were from the NH$_2$-termini of 57K fragments, starting at Asp$^{218}$ and Asp$^{222}$. These findings indicated that DMP1 is proteolytically cleaved at four bonds, Phe$^{173}$-Asp$^{174}$, Ser$^{180}$-Asp$^{181}$, Ser$^{217}$-Asp$^{218}$, and Gln$^{221}$-Asp$^{222}$, forming eight fragments. The uniformity of cleavages at the NH$_2$-terminal peptide bonds of aspartyl residues suggests that a single proteinase is involved. Based on its reported specificity, we hypothesize that these scissions are catalyzed by PHEX protein. We envision that the proteolytic processing of DMP1 plays a crucial role during osteogenesis and dentinogenesis.
INTRODUCTION

During osteogenesis and dentinogenesis, osteoblasts and odontoblasts secrete a unique set of noncollagenous proteins (NCPs)\(^1\) into their extracellular matrix (ECM). One category of NCPs is termed the SIBLING family that includes bone sialoprotein (BSP), osteopontin (OPN), dentin sialophosphoprotein (DSPP), dentin matrix protein 1 (DMP1), and matrix extracellular phosphoglycoprotein (1). These SIBLING family members share some common features such as the presence of relatively large amounts of sialic acid, phosphate, and the RGD cell-binding sequence as well as similarities in genomic organization and chromosomal localizations (1, 2). These proteins are very acidic and are secreted into the ECM during the formation and mineralization of these tissues; it is generally accepted that they play key biological roles in the formation of bone and dentin (3, 4); however, details concerning their precise functions are unknown.

Among SIBLING family members, the full-length forms of OPN and BSP have been isolated from ECM and fully characterized (5, 6). Although DSPP has never been detected in tissue, its cleavage products, dentin sialoprotein (DSP) and dentin phosphoprotein (DPP), have been isolated from dentin and thoroughly characterized (7–9). In contrast, there is a lack of information about the characteristics and functions of naturally occurring DMP1 with complete structures (primary sequence with posttranslational modifications such as glycosylation and phosphorylation).

DMP1, first identified by cDNA cloning using a rat odontoblast mRNA library, is a very acidic phosphoprotein (10). It was originally postulated to be dentin-specific, but later, its expression was observed in bone and brain (11, 12). The cDNA of DMP1 has been cloned and sequenced in a number of species including rat (10), mouse (11), bovine (12), human (13),
wallaby (14), opossum (14), platypus (14), caiman (15) and chicken (16). The characteristic feature of DMP1 is that it contains an unusually large number of acidic domains, a property that implicates it as a possible participant in regulating matrix mineralization (10, 17). Using an antibody monospecific to recombinant rat DMP1 to perform Western immunoblotting, George et al. (10) detected a band at an Mr of 61 kDa on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) of dentin extracts and concluded that this 61 kDa band was a full-length DMP1. Later, Srinivasan et al. (18) from the same group reported that the molecular weight of recombinant rat DMP1, determined by SDS-PAGE, was approximately 90-95 kDa, higher than that of the naturally occurring 61 kDa form of DMP1. According to the cDNA-deduced amino acid sequence, more than 55 of the 107 serine residues in the secreted rat DMP1 could be phosphorylated (10). Additionally, 10 threonine residues were potential phosphorylation sites for casein kinase I and II. Thus, the native form of DMP1 with posttranslational modifications would be expected to be larger than the recombinant protein. This discrepancy relative to the experimentally determined molecular weights remains unresolved.

Recombinant DMP1 has been synthesized and used for functional studies (19, 20). Although the use of recombinant DMP1 either produced in prokaryotic or eukaryotic cells may provide significant clues about its biological functions, conclusions drawn from such studies may not fully reflect its actual role in vivo. One of the major shortcomings relative to recombinant protein is a variation in posttranslational modifications such as proteolytic processing, phosphorylation and glycosylation. Although recombinant DMP1 can be posttranslationally modified by eukaryotic cell lines, these cells do not provide a microenvironment that exactly mimics the biosynthetic and transport environments in bone or dentin cells. Hence, there is a necessity for isolating DMP1 from bone or dentin, fully characterizing it and performing
functional studies. In searching for naturally occurring forms of DMP1 in bone and dentin, we discovered two clusters of proteins (37K and 57K fragments) that we speculated were fragments of DMP1 because we cleaved the mixture with trypsin and obtained sequences originating from DMP1 (21). The purpose of the present study was to isolate and characterize these fragments, and in particular, to identify the cleavage site(s) at which DMP1 is proteolytically processed.

EXPERIMENTAL PROCEDURES

**Isolation of 37K and 57K Fragments**

ECM proteins were extracted from rat long bone by standard procedures as described (5, 22). The extracts were first subjected to gel chromatography on Sephacryl S-200 (Amersham Biosciences AB) to separate the high molecular weight protein fraction from smaller-sized proteins; the latter eluted in an included volume. With DEAE-Sephacel (Pharmacia LKB Biotechnology) ion-exchange chromatography, the high molecular weight protein fraction was next separated into seven major fractions that were named D1, D2, D3, D4a, D4b, D5a, and D5b, corresponding to the elution gradient ranging from 0 to 0.53 M NaCl (2, 5). In D4b, fragments of DMP1 (namely 37K and 57K) co-eluted with BSP and OPN (Fig. 1A). Fraction D4b was further divided into five subfractions (2). For purification of 37K fragments, subfractions mostly rich in this protein were passed over a Bio-Gel A15m column (Bio-Rad Laboratories). This step eliminated OPN and partially removed BSP and 57K fragments. Finally, the fractions enriched in 37K fragments from Bio-Gel A15m chromatography were passed over a Mono-Q column (Amersham Pharmacia Biotech AB) connected to a fast protein liquid chromatography (FPLC) system. Thus, a sample for 37K fragments of relatively high purity was obtained (Fig.1B). For the purification of 57K fragments, subfractions rich in this protein were first passed over a Bio-Gel A15m column, and then the fractions enriched in 57K fragments were passed over a Bio-Gel
A50m column (Bio-Rad Laboratories). By this combination, we obtained a preparation of 57K fragments (Fig. 1C), derived exclusively from the COOH-terminus of DMP1 (see later). Please note that the continual presence of either 4M guanidium-HCl (GdmCl) or 6M urea in each step of the preparative procedures prevented artifactual degradation of proteins.

Alkaline Phosphatase Treatment and Trypsin Digestion

Prior to trypsin digestion, 37K and 57K fragments were dephosphorylated with calf intestine alkaline phosphatase (Sigma) attached to 150 µm oxirane-activated macroporous acrylic beads. Four hundred µg of 37K and 57K fragments, respectively, were incubated for 4 h at 37°C in the presence of 4 units of alkaline phosphatase in 100 mM Tris-HCl, pH 8.0 (total volume 0.5 ml). Following dephosphorylation, proteins were digested overnight at 37°C with trypsin (Roche Molecular Biochemicals) at an enzyme:substrate ratio of 1:50 (by weight) in 100 mM Tris-HCl, pH 8.0 (total volume 0.5 ml). The final concentration of trypsin was 16 µg/ml. Prior to the addition of trypsin, the chymotrypsin inhibitor TPCK (Roche Molecular Biochemicals) was added to a final concentration of 100 µg/ml to inhibit possible chymotrypsin activity.

Peptide Purification

Following trypsin digestion, direct separation of peptides by reversed-phase high performance liquid chromatography (HPLC) gave rise to a large number of overlapping peptide peaks; our past experience has shown that this level of separation does not yield optimal data for structure determinations. Therefore, peptides were separated using a two-dimensional approach to produce a series of pure peptides for unambiguous characterization. In the first phase, peptides were separated according to size on Superdex 75 HR 10/30 (Amersham Pharmacia Biotech AB) equilibrated and eluted at 0.5 ml/min in 4 M GdmCl (SigmaUltra, pH 6.0). Using FPLC, the eluant was monitored at 280 nm. The 30 ml elution volume was collected as 60 0.5 ml fractions.
Tryptic peptides from 37K and 57K fragments eluted in a volume of 10–17 ml. Fractions containing peptides ranging in size from 4 amino acids (fraction 32) to 50 amino acids (fraction 22) were selected for a second phase HPLC separation. Each selected fraction was subjected to HPLC using a 2.1 x 250 mm C18 reversed-phase column (Vydac). After applying the sample, the column was washed with water (0% acetonitrile) for 20 min before starting a 0–35% acetonitrile linear gradient in 0.1% trifluoroacetic acid (100 min at a flow rate of 300 µl/min). For all HPLC separations, the eluant was monitored at 218 nm and peaks were collected manually.

**Sequence Analysis and Determination of Mass**

Most purified peptides from HPLC were first sequenced by Edman degradation on an Applied Biosystems ABI Procise 492cLC or 477A Sequencer using standard techniques. Next, those that could not be fully sequenced by Edman degradation were further analyzed for molecular mass and sequence by mass spectrometry (MS). For MS, the samples were dried under vacuum and redissolved in aqueous solution containing approximately 50% methanol and 1% formic acid. Aliquots of the solutions were deposited either in metal coated glass nanoelectrospray capillary tubes for analysis on a PE Sciex API 3000 triple quadrupole mass spectrometer (Concord, Ontario, Canada) equipped with a Protana nanoelectrospray source (Odense, Denmark) or directly to the target of the ABI Voyager-DE STR (Foster City, CA) and 1 µl α-cyano-4-hydroxycinnamic acid (10 µg/µl) laser desorption time of flight analysis. The samples were analyzed by one or more of the following procedures. For determination of peptide molecular weight, full scan mass spectra were recorded utilizing either Q1 as the resolving analyzer in positive ion mode on the triple quadrupole instrument or by time of flight in the positive ion reflectron mode of analysis. Product ion spectra were acquired utilizing Q1 to
transmit the precursor ion of interest to the radio frequency-only collision cell Q2 under collisionally activated decomposition conditions. Nitrogen was the collision gas and collision energies in the range of 20–50 eV product ions were analyzed using Q3 to determine peptide sequence.

**Prediction of Secondary Structure for Amino Acid Sequences of DMP 1**

Using the Lasergene sequence analysis software package (DNASTAR, Madison WI), followed by molecular modeling and energy minimization (Consistant Valence Force Field) with Accelrys software (San Diego CA) on an SGI workstation (Silicon Graphics Inc.), secondary structure predictions were made on four segments of DMP1 sequences. Each sequence contained 20 amino acids from selected regions (ESSHGDGSEFDDEGMQSDDP, SEFDDEGMQSDDPGSTRSDR sequences around the two identified COOH-termini of 37K fragments and DHEPTSTQDSDDSQDVEFSS, TSTQDSDDSQDVEFSSRKSF around the two identified NH2-termini of 57K, see later). Chou-Fasman, Garnier-Robson, Eisenberg and Karplus-Shulz algorithm predictive results were used to make models that were minimized using conjugate gradient iterations until the root mean square deviation (rmsd) was below 0.0001. Results were displayed with Insight II (2000, Accelyrs) and the secondary structure render Kabsch-Sandler algorithm was utilized.

**Amino Acid and Phosphate Analysis**

For amino acid analysis, aliquots from samples of 37K and 57K fragments in water were dried, hydrolyzed by HCl gas (6M HCl) and analyzed with an ABI 420A Amino Acid Analyzer (Applied Biosystems). For each sample, three amino acid analyses were performed and average values were calculated.
For phosphate analysis, aliquots of the 37K and 57K fragment solutions used for amino acid analysis were hydrolyzed in 0.2M NaOH at 37°C overnight. Concentrations of inorganic phosphate were measured (23). For each sample, three independent values were obtained and averaged for calculating the phosphate contents of 37K and 57K fragments. Based on amino acid analysis, the molar concentrations of phosphate for each sample were determined and these values were then used to calculate the number of phosphates/mol.

RESULTS

The following section is organized in a systematic manner, designed to help the reader to more easily comprehend our findings and the scientific implications. Thus, the presentation of the results begins with isolation and characterization of fragments, gives analytical results for all the tryptic peptides and leads to the summation of the major conclusions about the sites of proteolytic processing.

Isolation of 37K and 57K Fragments

The purity of the 37K and 57K fragments was assessed by SDS-PAGE and Stains-All staining. As shown in Fig. 1B, the 37K fragments gave rise to a cluster of bands migrating between ~29 and ~37 kDa on 5-15% SDS-PAGE. On 7.5% SDS-PAGE, the major band of this set migrated at 37 kDa (hence named 37K fragments). On 5-15% SDS-PAGE, the 57K fragments gave rise to a cluster of bands migrating between ~45 and ~60 kDa (Fig. 1C). On 7.5% SDS-PAGE, the major band of this sample migrated at 57 kDa (hence designated 57K fragments).

Despite the apparent heterogeneity, we concluded that both sets of protein fragments were derived from DMP1 as evidenced by the following. Amino acid sequencing (see later) of the tryptic peptides from each showed that every peptide we obtained was solely derived from
DMP1. Some minor peptides from trypsin itself were also detected, but we found no evidence for the presence of tryptic peptides from any contaminating proteins. Note that only 2% of the total protein was trypsin.

As shown in Figs. 1B and 1C, both 37K and 57K fragments displayed relatively broad clusters of protein bands on SDS-PAGE. To examine if the apparent heterogeneity in migration rate is due to variations in phosphorylation, we performed experiments to compare the migration rates of dephosphorylated 37K and 57K fragments with those of the non-dephosphorylated. Although the dephosphorylated 37K and 57K fragments displayed migration rates slightly slower than the non-dephosphorylated (data not shown), the observed heterogeneity in migration rate remained the same before and after dephosphorylation. These results indicated that this heterogeneity in migration rate was not caused by difference in phosphorylation level. We speculate that this apparent heterogeneity in migration rate within the cluster of 37K or 57K fragments may be due to difference in fragment length and/or glycosylation level.

**Separation of Tryptic Peptides Derived from 37K and 57K Fragments**

When we first used trypsin to digest 37K and 57K fragments without dephosphorylation, only a few peptides were obtained in very low amounts. Realizing that a majority of trypsin cleavage sites might be blocked by protein bound phosphate, we then used alkaline phosphatase to remove phosphates prior to trypsin digestion. This significantly increased the level of cleavage, allowing a thorough analysis of a number of pure peptides from the 37K and 57K fragments.

Examples of HPLC separation of tryptic peptides derived from 37K and 57K fragments are shown in Figs. 2 and 3, respectively. For both 37K and 57K fragments, we performed HPLC
separations on Superdex 75 fractions 22 through 32. These 11 Superdex fractions included peptides ranging in size from 4 to 50 amino acids.

Fig. 2A shows an example of HPLC separation of 37K fragment tryptic peptides from Superdex fraction 32; peak 2 represents the DMP1 NH$_2$-terminus (residues 1-5). Fig. 2B is an HPLC chromatogram of Superdex fraction 26 from 37K fragments. Curiously, peaks 4 and 5 showed exactly the same sequence, V$_{157}^{180}$GGGSEGESSHGDGSEFDDEGMQS (a COOH-terminal sequence from the 37K fragments). Fig. 3A shows an HPLC separation of tryptic peptides in Superdex fraction 29, in which peak 3 represents an NH$_2$-terminal peptide of 57K fragments (residues 218-228). The peptides in both peaks 2 and 4 of Fig. 3A had sequences from residues 252-262 of DMP1, except that they differed in one residue. From peak 4, we obtained the sequence E$_{252}^{262}$TQSVSTEDFR, in agreement with the cDNA-deduced sequence. The sequence for peak 2 was E$_{252}^{262}$TQSDSTEDFR. This single difference of Asp$_{256}$ in place of Val$_{256}$ probably represents an allelic variant. We also found two 57K fragment peptides (peaks 5 and 6 in Fig. 3B) that gave identical sequences: K$_{450}^{473}$LIVDAYHNKPIGDQDDNDNDCQDGY derived from the COOH-terminus of DMP1. Please note that Lys$_{450}$ was modified by a 43 Da carbamoyl group, a substance resulting from the reaction of the epsilon-amino group with cyanate, present in urea (24).

**Structures of Tryptic Peptides Derived from 37K and 57K Fragments**

Fig. 4A is the cDNA-deduced amino acid sequence of rat DMP1 NH$_2$-terminal region (10). Amino acids are numbered from the NH$_2$-terminus of DMP1, excluding the 16 amino acid signal peptide. Fig. 4B is a summary of tryptic peptides derived from 37K fragments, sequenced in this study. Totally, we sequenced 12 peptides covering 154 amino acid residues; all originated from the NH$_2$-terminal region of DMP1. Sequences were determined by either Edman
degradation or MS or by both. Please note that in peptide L\textsuperscript{1}PVAR\textsuperscript{5}, the NH\textsubscript{2}-terminal residue Leu\textsuperscript{1} was modified by a 43 Da carbamoyl group (24); this peptide could only be sequenced by MS since its NH\textsubscript{2}-terminus was blocked. Among the 12 peptides sequenced, two (peptides 9 and 10) represent the COOH-terminal region of 37K fragments (see below), ending at Phe\textsuperscript{173} and Ser\textsuperscript{180}, respectively.

Fig. 5A shows the cDNA-deduced amino acid sequence of rat DMP1 COOH-terminal region (10). Amino acids are numbered as described above. Fig. 5B is a summary of tryptic peptides derived from 57K fragments, analyzed in this study. Totally, we sequenced 15 tryptic peptides covering 146 residues; all originated from the COOH-terminal region of DMP1. Two of the 15 peptides (peptides 4 and 5) represent NH\textsubscript{2}-terminal regions of 57K fragments (see below), beginning at Asp\textsuperscript{218} and Asp\textsuperscript{222}; they could not be sequenced by Edman degradation since the two NH\textsubscript{2}-terminal residues were modified by the 43 Da carbamoyl moiety (24).

**Identification of COOH-termini of 37K Fragments**

The identification of COOH-terminal peptides from 37K fragments, resulting from trypsin digestion, was made by searching for peptides with a COOH-terminal amino acid other than lysine or arginine. For a complete search, we sequenced all the major HPLC peaks containing peptides ranging in size from 4 to 50 amino acids (Superdex fractions 22–32). We identified two tryptic peptides from the COOH-termini of 37K fragments (Fig. 4B, peptides 9 and 10): V\textsuperscript{157}GGGSEGESSHGDFSEF\textsuperscript{173} and V\textsuperscript{157}GGGSEGESSHGDFSEFDDEGMQS\textsuperscript{180}. These two peptides were first partially sequenced by Edman degradation and then fully analyzed by MS. The presence of these two residues, Phe\textsuperscript{173} and Ser\textsuperscript{180} at the end of 37K fragments, along with comparison to the cDNA-deduced sequence (Fig. 4A), indicated that peptide bonds Phe\textsuperscript{173}-Asp\textsuperscript{174} and Ser\textsuperscript{180}-Asp\textsuperscript{181}, were cleaved (discussed later). We also found one peptide that
contained uncleaved Phe\(^{173}\) -Asp\(^{174}\) and Ser\(^{180}\) -Asp\(^{181}\) bonds; peptide 11 (Fig. 4B) was V\(^{157}\) GGGSEGESSHGDGSEFDDEGMQSDDPGSTR\(^{187}\). Another tryptic peptide (peptide 12, Fig. 4B) gave the sequence M\(^{195}\) SSAGIR\(^{201}\) that was also beyond the COOH-termini of 37K fragments described above. After sequencing all the major HPLC peaks containing peptides ranging from 4 to 50 amino acids, we did not detect any additional COOH-terminal peptides.

**Identification of NH\(_2\)-termini of 57K Fragments**

NH\(_2\)-terminal tryptic peptides from 57K fragments were identified by searching for peptides with an NH\(_2\)-terminal amino acid not preceded by lysine or arginine (as indicated from the cDNA-deduced amino acid sequence). We identified two peptides from the NH\(_2\)-termini of 57K fragments, D\(^{218}\) DSQDVEFSSR\(^{228}\) and D\(^{222}\) VEFSSR\(^{228}\) (Fig. 5B, peptides 4 and 5). In fact, the observation that Asp\(^{218}\) and Asp\(^{222}\) were modified by the 43 Da carbamoyl groups resulting from the reaction of the amino groups with cyanates in urea (24), provided additional evidence that these two residues were the NH\(_2\)-termini of 57K fragments. The existence of the two NH\(_2\)-termini, Asp\(^{218}\) and Asp\(^{222}\) indicated that peptide bonds Ser\(^{217}\) -Asp\(^{218}\) and Gln\(^{221}\) -Asp\(^{222}\) had been cleaved (discussed later). As shown in Fig. 5B, the structures of three peptides, M\(^{195}\) SSAGIR\(^{201}\) (peptide 1), S\(^{202}\) EESKGDHEPTSTQDSDDSQDVEFSSR\(^{228}\) (peptide 2) and G\(^{207}\) DHEPTSTQDSDDSQDVEFSSR\(^{228}\) (peptide 3), indicated that bonds Ser\(^{217}\) -Asp\(^{218}\) and Gln\(^{221}\) -Asp\(^{222}\) were not totally cleaved. After sequencing all the major HPLC peaks containing 57K fragment tryptic peptides ranging in size from 4 to 50 amino acids (fractions 22–32), we were unable to detect any NH\(_2\)-terminal peptides other than the two described above.

**Secondary Structure Prediction for Sequences around Cleavage Sites**

The identification of the COOH-termini of 37K fragments revealed that peptide bonds Phe\(^{173}\) -Asp\(^{174}\) and Ser\(^{180}\) -Asp\(^{181}\) were cleaved. The results of secondary structure predictions
revealed that these two cleavages occurred at the NH$_2$-terminal and COOH-terminal edges, respectively, of a predicted $\alpha$-helix formed by residues D$_{174}^{174}$DEGMQ$_{179}^{179}$ (Fig. 6A). Secondary structure predictions for the sequences around the cleaved peptide bonds Ser$_{217}^{217}$-Asp$_{218}^{218}$ and Gln$_{221}^{221}$-Asp$_{222}^{222}$ (as revealed by the NH$_2$-termini of 57K fragments) indicated that Ser$_{217}^{217}$-Asp$_{218}^{218}$ was located 4 residues away from the NH$_2$-terminal edge of a predicted $\alpha$-helix formed by residues D$_{222}^{222}$VEFSSR$_{228}^{228}$ whereas Gln$_{221}^{221}$-Asp$_{222}^{222}$ was at the edge (Fig. 6B). Note that the surrounding chain segments did not favor formation of any regular conformation.

**Amino Acid and Phosphate Analysis**

Amino acid analysis showed that the compositions of 37K and 57K fragments agreed closely with those calculated from the cDNA-deduced amino acid sequence, using Ser$_{180}^{180}$ as the end of 37K fragments, and Asp$_{218}^{218}$ as the beginning of 57K fragments (data not shown). Using the molar concentrations of these 37K and 57K fragments, and the phosphate contents of the same solutions, we calculated that 37K fragments had 12 phosphates/mol while 57K had 41. Using Ser$_{180}^{180}$ as the end of 37K fragments, we calculated that the average ratio of phosphate to amino acid residue was 1:15. With Asp$_{218}^{218}$ as the beginning of 57K fragments, this ratio for 57K was 1:6. These data indicate that the 57K fragments are more highly phosphorylated than 37K fragments.

**DISCUSSION**

In the present study, we isolated and characterized fragments derived from naturally occurring form of DMP1 in rat bone. Extensive sequencing of tryptic peptides derived from these fragments (Figs. 4B, 5B) shows unequivocally that they are derived from DMP1 and strongly suggest that this protein in bone is proteolytically processed into fragments, including 37K fragments from the NH$_2$-terminal region and 57K fragments from the COOH-terminal
region. In order to locate the cleavage sites at which DMP1 is processed, we sequenced all of the major HPLC peaks containing peptides ranging in size from 4 to 50 amino acids. From 37K fragments, we identified two COOH-termini, Phe\textsuperscript{173} and Ser\textsuperscript{180}, and from 57K fragments, we detected two NH\textsubscript{2}-termini, Asp\textsuperscript{218} and Asp\textsuperscript{222}. Based on these findings along with comparison to the cDNA-deduced amino acid sequence, we conclude that four peptide bonds, Phe\textsuperscript{173}-Asp\textsuperscript{174}, Ser\textsuperscript{180}-Asp\textsuperscript{181}, Ser\textsuperscript{217}-Asp\textsuperscript{218} and Gln\textsuperscript{221}-Asp\textsuperscript{222} are cleaved during DMP1 processing. In other words, DMP1 is processed into eight potential fragments, 4 pieces of 37K and 4 pieces of 57K (Fig. 7A). As shown in Fig. 7A, the longest 37K fragment has 221 amino acids (37K-221AA), the shortest one with 173 (37K-173AA). For 57K fragments, the longest fragment consists of 300 amino acids (57K-300AA) while the shortest one contains 252 amino acids (57K-252AA). A comparison of sequences around the four cleavage sites showed that all of the cleavages occurred at the NH\textsubscript{2}-terminus of an aspartic acid residue, three of them being immediately followed by an Asp-Asp motif (Fig. 7B). These findings suggest that DMP1 is processed by a single proteinase that cleaves DMP1 at the NH\textsubscript{2}-terminus of an Asp residue. Although the proteinase catalyzing these scissions has not been directly identified at this time, we speculate that a tissue-specific enzyme, designed to activate DMP1 by cleaving it into 37K and 57K fragments, is involved. One possible candidate for the catalytic cleavage of DMP1 is a protein encoded by PHEX (phosphate-regulating gene with homologies to endopeptidases on X chromosome) which is expressed predominantly in bone and tooth (25–30). This conclusion is based on reports that PHEX protein has a strong preference for cleavage of peptide bonds at the NH\textsubscript{2}-terminus of aspartic acid residues (28, 31, 32). As discussed above, our data indicate that DMP1 is proteolytically processed in a similar fashion (Figs. 7A, 7B). Mutations in the PHEX gene are associated with X-linked hypophosphatemic rickets in humans (33) and the murine homolog, Hyp (26). PHEX
protein is a cell membrane-associated glycoprotein with structural homologies to members of the M13 family of metalloendopeptidases (28, 31, 33). The M13 endopeptidases are zinc-containing type II integral membrane glycoproteins with a relatively short cytoplasmic NH₂-terminal region, a single transmembrane domain, and a long extracytoplasmic domain, which contains the active site of the enzyme (34). Immunohistochemical studies have demonstrated that signals for PHEX protein are stronger in osteocytes than in osteoblasts and localized predominantly on the plasma membrane (35). Likewise, DMP1 in bone is more highly expressed in osteocytes than in osteoblasts (36, 37). This coincidental localization of DMP1 and PHEX protein in osteocytes and osteoblasts strengthens the possibility that PHEX protein may be the proteinase involved in the processing of DMP1. Based on these data, it is tempting to speculate that, during or after secretion, DMP1 is processed by PHEX protein on cell surfaces, converting this SIBLING family member from an inactive precursor into active fragments.

Since there are many aspartic acid residues and quite a few Asp-Asp motifs in the DMP1 amino acid sequence, the fact that cleavages only occur at the four sites discussed above suggests that the scissions by the enzyme may relate to spatial structure of DMP1 in addition to conserved aspartic acid residues. The structure predictions suggest that three of the cleaved bonds, Phe₁⁷³-Asp₁⁷⁴, Ser₁⁸⁰-Asp₁⁸¹, and Gln²²¹-Asp²²², are located at the edges of a possible short (6-7 residues) α-helix. The Ser²¹⁷-Asp²¹⁸ cleavage site was 4 residues NH₂-terminal to the α-helix. Although we cannot draw a firm conclusion based on the secondary structure predictions for the sequences around the cleavage sites, such data may provide clues about the selectivity of DMP1 bonds that are proteolytically processed. Further studies, designed to focus on the tertiary structure of DMP1, and on the substrate specificity of PHEX, are beyond the scope of this study.
Fig. 7C shows an alignment of amino acid sequences of DMP1 from four species: rat (10), mouse (11), bovine (12) and human (13). We selected a portion corresponding to rat DMP1 residues 155–234 that covers the four cleavage sites, for this alignment. An alignment of complete DMP1 amino acid sequences among the four species can be found in reference 13. It is worth noting that the flanking region (corresponding to rat DMP1 residues 165–183) of two cleaved bonds, Phe$^{173}$-Asp$^{174}$ and Ser$^{180}$-Asp$^{181}$ shows a high level of conservation among these species. In fact, this region is one of the three most highly conserved domains (the other two highly conserved domains are located in the NH-terminal and COOH-terminal regions of DMP1 sequence, respectively) (11, 13, 15). Curiously, the sequences around the other two cleaved bonds, Ser$^{217}$-Asp$^{218}$ and Gln$^{221}$-Asp$^{222}$ display no apparent sequence conservation among these species.

DMP1 amino acid sequence deduced from cDNA indicates that the secreted rat DMP1 has 65 potential phosphorylation sites (10). If fully phosphorylated, rat DMP1 would have an average of 0.137 phosphates for each residue. Phosphate analysis in the present study revealed that a full-length DMP1 contained 53 phosphates (12 for 37K fragments plus 41 for 57K fragments). In our previous studies with OPN and DSP peptides (24, 38), we have noted that phosphoserines juxtaposed to lysine or arginine can prevent these residues from being cleaved by trypsin. In agreement with this observation, dephosphorylation of 37K and 57K fragments in the present study resulted in a more efficient digestion of these proteins by trypsin. In addition, dephosphorylation of DMP1 fragments reduces their acidity, thus increasing the hydrophobicity of many of their tryptic peptides, since phosphates are negatively charged. We believe that without dephosphorylation, certain numbers of the peptides obtained in the present study would
not have been detected due to their inability to bind onto a reversed-phase C-18 column. One disadvantage of dephosphorylation is that we were unable to identify sites of phosphorylation.

We suggest that the present data provide insights into the processing of DSPP. It is well known that DSPP is processed into two fragments, DSP and DPP found in the ECM of dentin (39). A previous study in our laboratory identified and characterized two COOH-terminal peptides for rat DSP (24), the major one ending at Tyr$^{421}$ and the minor one at His$^{406}$, which suggests that the proteolytic processing of rat DSPP to DSP and DPP involves proteolytic cleavage of Tyr$^{421}$-Asp$^{422}$ (giving rise to the COOH-terminus of DSP) and Gly$^{430}$-Asp$^{431}$ (the NH$_2$-terminus of DPP) and to lesser extent His$^{406}$-Ser$^{407}$ (a second, minor COOH-terminus). Interestingly, the two major cleavages in DSPP also occur at the NH$_2$-terminus of Asp, one being immediately followed by an Asp-Asp motif (Ref. 24, Fig. 4). Based on these similarities, we speculate that same enzyme (possibly, PHEX protein) and perhaps substrate conformation properties may be involved in the proteolytic processing of DMP1 and DSPP.

Phosphate analysis revealed that 57K fragments (COOH-terminal portion of DMP1) are more highly phosphorylated than 37K (NH$_2$-terminal portion). We know that DPP (COOH-terminal portion of DSPP) is much more highly phosphorylated than DSP (NH$_2$-terminal portion) (9, 39, 40). Experiments on the in vitro formation of hydroxyapatite and growth of the mineral crystals indicate that DSP has little or no effect on mineralization (41) whereas DPP actively regulates the formation and growth of hydroxyapatite in vitro (42-44). We have previously hypothesized that DSPP is an inactive precursor and its proteolytic processing is an activation step to release the functional fragment, DPP (45). Based on the similarities in chemical features and proteolytic processing between DMP1 and DSPP, we have formed a similar hypothesis about DMP1: that full-length DMP1 is an inactive precursor that requires processing to release
functional molecules, 37K and/or 57K fragments. We believe that this proteolytic processing plays a significant, crucial role for the biological functions of DMP1 during osteogenesis and dentinogenesis. To address these issues, further investigations are needed to test the effects of 37K and 57K fragments as well as full-length DMP1 on mineralization.

Additionally, we have observed that dentin ECM contains 37K and 57K fragments. We have isolated 37K fragments from rat dentin. The HPLC chromatographic profile of tryptic peptides derived from dentin 37K fragments was identical to that obtained from bone 37K fragments. Selective sequencing of tryptic peptides derived from 37K fragments indicated that the NH₂-termini and COOH-termini of dentin 37K fragments were identical to those of bone. Attempts to obtain pure dentin 57K fragments for sequence analysis have been unsuccessful.

REFERENCES


**FOOTNOTES**

*This work was supported by NIH Grant DE 05092 (to W.T.B.) and Grants AR 13921 and DE 01374 (to A.V.).

§ Correspondence address:

Chunlin Qin, Department of Basic Sciences, The university of Texas–Houston Health Science Center, Dental Branch, M.D. Anderson Blvd., DBB, Rm 4.133, Houston, TX 77030, USA. Phone: 713–500–4583; Fax: 713–500–4500, Email Address: Chunlin.Qin@uth.tmc.edu

1 The abbreviations used are: NCPs, noncollagenous proteins; ECM, extracellular matrix; BSP, bone sialoprotein; OPN, osteopontin; DSPP, dentin sialophosphoprotein; DMP1, dentin matrix protein 1; DSP, dentin sialoprotein; DPP, dentin phosphoprotein; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; GdmCl, guanidium-HCl; FPLC, fast protein liquid chromatography; HPLC, high performance liquid chromatography; MS, mass spectrometry; PHEX, phosphate-regulating gene with homologies to endopeptidases on X chromosome; Hyp, X-linked hypophosphatemic rickets in mice.
FIGURE LEGENDS

Fig. 1A. SDS-PAGE and Stains-All staining of a subfraction from D4b of bone DEAE-Sephacel chromatography (Ref. 2).

Lane 1, molecular weight standard; lane 2, sample form bone D4b. Fifty µl of urea containing sample from bone D4b was loaded on a 5-15% gel and proteins were visualized with Stains-All. Note the presence of BSP and OPN along with DMP1 fragments.

Fig. 1B. SDS-PAGE and Stains-All staining of purified 37K fragments.

Lane 1, molecular weight standard; lane 2, 37K fragment sample. On 5-15% SDS-PAGE, 37K fragments appear as a cluster of bands migrating between ~29 and ~37 kDa.

Fig. 1C. SDS-PAGE and Stains-All staining of purified 57K fragments

Lane 1, molecular weight standard; lane 2, 57K fragment sample. On 5-15% SDS-PAGE, 57K fragments occur as a cluster of bands migrating between ~45 and ~60 kDa.

Fig. 2. Examples of HPLC separation of tryptic peptides from 37K fragments.

A: Separation of Superdex 75 fraction 32, a dot under Leu\(^1\) in the NH\(_2\)-terminal peptide L\(^1\)PVAR\(^5\) (peak 2) indicates that this residue was modified by a 43 Da carbamoyl group.

B: Separation of fraction 26, one of the COOH-terminal peptides obtained in this study, V\(^{157}\)GGGSEGESSHGDSFDDEGMS\(^{180}\) was detected in this fraction (peaks 4 and 5). Arrow indicates that peptide in peak 1 was not fully sequenced.

Fig. 3. Examples of HPLC separation of tryptic peptides from 57K fragments.

A: Separation of fraction 29, peak 3 was one of the NH\(_2\)-terminal peptides obtained in this study, D\(^{218}\)DSQDVEFSSR\(^{228}\). A dot under D\(^{218}\) indicates the modification of Asp by a 43 Da carbamoyl group. Please compare peaks 2 and 4 and note the elution time delay caused by a change from Val to Asp.
B: Separation of fraction 26, peaks 5 and 6 represent the same sequence, K$^{450}$LIVDAYHNKPGIDQDDNDCQDGY$^{473}$, from the COOH-terminus of DMP1.

Fi. 4. cDNA-deduced amino acid sequence of rat DMP1 (10) and tryptic peptides obtained from 37K fragments.

A: cDNA-deduced amino acid sequence of the DMP1 NH$_2$-terminal region. Amino acids are numbered from the NH$_2$-terminus of secreted DMP1. Residues sequenced in the present study are shaded. Arrow heads indicate the cleavage sites.

B: A summary of data for tryptic peptides derived from 37K fragments, analyzed in this study. 

a, peptides we have sequenced, numbered (P#) starting from the NH$_2$-terminal region. 
b, peptide sequence deduced from DMP1 cDNA (cDNASeq). 
c, theoretical molecular mass (TMM; Da) of peptide sequence deduced from cDNA. 
d, peptide sequence determined by Edman degradation (EDSeq). 
e, molecular mass determined by MS (MSMM). 
f, peptide sequence determined by MS (MSSeq). 
g, A dot under a residue indicates the modification of that residue by a 43 Da carbamoyl group. 
h, not determined. 
i, Interrupted line indicates that Edman degradation could not be run to the end of a peptide. 
j, - indicates an absence of a PTH-derivative during Edman degradation. 
k, COOH-terminal peptide not ending at Arg or Lys.

Fi. 5. cDNA-deduced amino acid sequence of rat DMP1 (10) and tryptic peptides obtained from 57K fragments.

A: cDNA-deduced amino acid sequence of the DMP1 COOH-terminal region. Amino acids are numbered as in Fig. 4A. Residues sequenced in the present study are shaded. Arrow heads indicate the cleavage sites.
B: A summary of tryptic peptides obtained from 57K fragments, analyzed in this study. The marks are the same letters designated (a – i) as those in Fig. 4B, except for the two NH$_2$-terminal peptides indicated by stars (*).

**Fig. 6. Prediction of secondary structure for sequences around the cleavage sites.**

**A:** Cleaved bonds Phe$^{173}$-Asp$^{174}$ and Ser$^{180}$-Asp$^{181}$ locate at the NH$_2$-terminal and COOH-terminal edges of a predicted $\alpha$-helix formed by residues D$^{174}$DEGMQ$^{179}$. Arrows indicate cleavages sites.

**B:** Cleaved bond Gln$^{221}$-Asp$^{222}$ locates at the NH$_2$-terminal edge of a predicted $\alpha$-helix formed by residues D$^{222}$VEFSSR$^{228}$ while another cleaved bond Ser$^{217}$-Asp$^{218}$ is 4 residues NH$_2$-terminal to the $\alpha$-helix. Arrows indicate cleavages sites.

**Fig. 7A. Schematic illustration of DMP1 processing.**

DMP1 is processed at 4 cleavage sites, forming 8 fragments. Each fragment is named in accordance with the amino acid number it contains. For example, the biggest 37K fragment with 221 amino acids is named 37K-221AA; the largest 57K fragment with 300 amino acids is designated 57K-300AA. Vertical arrows indicate cleavage sites.

**Fig. 7B. A comparison of sequences around the four cleavage sites.**

All cleavages occur at the NH$_2$-terminus of an aspartic acid residue, three of them being immediately followed by an Asp-Asp motif. Vertical arrows indicate cleavage sites.

**Fig. 7C. An alignment of amino acid sequences of DMP1 from rat (10), mouse (11), bovine (12) and human (13).** We selected a portion corresponding to rat DMP1 residues 155–234 (around the four cleavage sites) for this alignment. Amino acids are numbered in the left column starting from the NH$_2$-terminus of the secreted rat DMP1. Names of species are written in the right column. Identical residues among the species are indicated by stars (*). Letters indicate
differing amino acids. Dashes indicate absent residues. The cleavage sites are marked with vertical arrows. Note that there is a high level of conservation in the region (rat DMP1 residues 165–183) flanking the two cleaved bonds, Phe\textsuperscript{173} - Asp\textsuperscript{174} and Ser\textsuperscript{180} - Asp\textsuperscript{181}, but no obvious conservation with regard to the other two cleaved bonds, Ser\textsuperscript{217} - Asp\textsuperscript{218} and Gln\textsuperscript{221} - Asp\textsuperscript{222}.
$1 = Q^{92} WGGPSR^{98}$

$2 = L^{1} PVAR^{5}$

Fig. 2A
Fig. 2B

1 = T\textsuperscript{18}G\text{NLAQ}
2 = V\textsuperscript{157}G\text{GGSEGESSHGDGSEFDDEGMQSDDPGSTR}\textsuperscript{187}
3 = D\textsuperscript{131}H\text{HSDEADSREAGDSTQDSSESEYR}\textsuperscript{156}
4 = 5 = V\textsuperscript{157}G\text{GGSEGESSHGDGSEFDDEGMQS}\textsuperscript{180}
Fig. 3A

1 = $^3_{319}$GDPDNTSQTGDQ$^2_{332}$R
2 = $^2_{252}$TSQDSTEDFR
3 = $^2_{218}$DSQDVEFSSR
4 = $^2_{252}$TQSVSTEDFR

AU

Minutes
Fig. 3B

1 = S^207\text{KEESNSTGSTSSSEEDNHPKNIEADNR}^{449}
2 = G^207\text{DHEPTSTQDSDDSQDVEFSSR}^{228}
3 = S^202\text{EESKGDHEPTSTQDSDDSQDVEFSSR}^{228}
4 = V^236\text{SEEDDRGELADSNSR}^{251}
5 = 6 = K^450\text{LIVDAYHNKPIGDQDDNDJCQDGY}^{473}
1  LPVARYQNTESSESSEERTGNAQSPPPPMDANSIDTSDSES
41  GEELGSDR SQYRPAGGLSKSAQMDADKEEEDDSSGDTSFG
81  DEDNGPGPEERQWGGPSRLDSDEDSADTTQSSEDSTSQEN
121  SAQDTPSDSDKHSSDEADSRPEAGDSTQDSESEEEYRVGGG
161  SEGESSHGDGSEFDDEGMQSDDPGSTRSDRGHTRMSSAGI
201  RSEESKGDEPSTQDSDDSQ(221)

Fig. 4A
<table>
<thead>
<tr>
<th>P#</th>
<th>cDNASeq ²</th>
<th>TM²</th>
<th>EDSQ ²</th>
<th>MSEQ ²</th>
<th>MSQ ²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>L¹PVAR⁶</td>
<td>555.4</td>
<td>N-terminus blocked</td>
<td>598.4</td>
<td>LPVAR⁷</td>
</tr>
<tr>
<td>2</td>
<td>Y⁶QNTESSESEE¹⁷</td>
<td>1458.6</td>
<td>ND⁸</td>
<td>1458.4</td>
<td>YQNTESSESEE</td>
</tr>
<tr>
<td>3</td>
<td>T¹⁸GNL AQSPPFMMRPSGDMTSSESSEQEMGS²⁰</td>
<td>TUNLAQSPPPMP</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>S⁶⁸QYRPAGGLSK⁵⁹</td>
<td>1163.6</td>
<td>SQYRPAGGLSK</td>
<td>1163.5</td>
<td>SQYRPAGGLSK</td>
</tr>
<tr>
<td>5</td>
<td>S⁶⁰AGMADEDEDEDGDDTFQGEDNPGPEER⁹¹</td>
<td>SAGMDADKEE...</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Q²WGGPSR²⁸</td>
<td>QWGPGSR</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>L⁹⁹DSDEDASDTQSSSEDSTQENSAQDTPSDSK¹³⁰</td>
<td>LDSDEDSADTTQSSEDS</td>
<td>LDSDEDSADTTQSSEDS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>D¹³¹HHSDEASRPEAGDSTQDSESEEYR¹⁵⁶</td>
<td>DHHSDEASRPEAG</td>
<td>DHHSDEASRPEAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>V¹⁵⁷GGSEGEGSHGDGSEP¹⁷³</td>
<td>1594.2</td>
<td>VGGS...</td>
<td>1594.7</td>
<td>VGGSSEGEGSHGDGSEP²⁴</td>
</tr>
<tr>
<td>10</td>
<td>V¹⁵⁷GGSEGEGSHGDGSEPDEGMQS¹⁸⁰</td>
<td>2357.3</td>
<td>VGGGS...</td>
<td>2358.0</td>
<td>VGGSSEGEGSHGDGSEPDEGMQSK²⁴</td>
</tr>
<tr>
<td>11</td>
<td>V¹⁵⁷GGSEGEGSHGDGSEPDEGMQSDDPGSTR¹⁹⁷</td>
<td>3084.2</td>
<td>VGGGS...</td>
<td>3085.9</td>
<td>VGGSSEGEGSHGDGSEPDEGMQSDDPGSTR²⁴</td>
</tr>
<tr>
<td>12</td>
<td>A¹⁹⁸SSAGIR²⁰¹</td>
<td>720.9</td>
<td>ND</td>
<td>721.4</td>
<td>MSSAGIR</td>
</tr>
</tbody>
</table>

**Fig. 4B**
Fig. 5A
<table>
<thead>
<tr>
<th>P#</th>
<th>cDNASeq</th>
<th>TMM</th>
<th>EDSeq</th>
<th>MSMM</th>
<th>MSSeq</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M&lt;sup&gt;195&lt;/sup&gt; SSAGIR&lt;sup&gt;201&lt;/sup&gt;</td>
<td>720.9</td>
<td>MSSAGIR</td>
<td>720.4</td>
<td>MSSAGIR</td>
</tr>
<tr>
<td>2</td>
<td>S&lt;sup&gt;202&lt;/sup&gt; EESKGDHEPTSTQ DSDSQDVDFESSR&lt;sup&gt;228&lt;/sup&gt;</td>
<td>3000.2</td>
<td>SEESKGDHEPTSTQ&lt;sup&gt;g&lt;/sup&gt; DSDSQDVDFESSR</td>
<td>3043.2</td>
<td>SEESKGDHEPTSTQ&lt;sup&gt;g&lt;/sup&gt; DSDSQDVDFESSR</td>
</tr>
<tr>
<td>3</td>
<td>G&lt;sup&gt;207&lt;/sup&gt; DHEPTSTQDS DDSQDVDFESSR&lt;sup&gt;228&lt;/sup&gt;</td>
<td>2438.9</td>
<td>ND&lt;sup&gt;h&lt;/sup&gt;</td>
<td>2439.6</td>
<td>GDHEPTSTQDS DDSQDVDFESSR</td>
</tr>
<tr>
<td>4</td>
<td>D&lt;sup&gt;218&lt;/sup&gt; DDSQDVDFESSR&lt;sup&gt;228&lt;/sup&gt;</td>
<td>1284.5</td>
<td>N-terminus blocked</td>
<td>1327.6</td>
<td>DDSQDVDFESSR&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>5</td>
<td>D&lt;sup&gt;222&lt;/sup&gt; VEFSSR&lt;sup&gt;228&lt;/sup&gt;</td>
<td>838.4</td>
<td>N-terminus blocked</td>
<td>882.4</td>
<td>DVEFSSR&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>6</td>
<td>K&lt;sup&gt;223&lt;/sup&gt; SFR&lt;sup&gt;232&lt;/sup&gt;</td>
<td>KSFR</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>7</td>
<td>V&lt;sup&gt;236&lt;/sup&gt; SEEDDR&lt;sup&gt;242&lt;/sup&gt;</td>
<td>VSEEDDR</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>8</td>
<td>V&lt;sup&gt;236&lt;/sup&gt; SEEDDRGELADSNR&lt;sup&gt;251&lt;/sup&gt;</td>
<td>VSEEDDRGELADSNR</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>9</td>
<td>E&lt;sup&gt;252&lt;/sup&gt; TQSVSTEDFR&lt;sup&gt;262&lt;/sup&gt;</td>
<td>1298.6</td>
<td>ETQSVSTEDFR</td>
<td>1298.7</td>
<td>ETQSVSTEDFR</td>
</tr>
<tr>
<td>10</td>
<td>E&lt;sup&gt;252&lt;/sup&gt; TQSDSTEDFR&lt;sup&gt;262&lt;/sup&gt;</td>
<td>1314.5</td>
<td>ETQSDSTEDFR</td>
<td>1314.6</td>
<td>ETQSDSTEDFR</td>
</tr>
<tr>
<td>11</td>
<td>S&lt;sup&gt;269&lt;/sup&gt; ETQEDTAETQSQEDSQEGQDPSESSEEAGEPSQESSESQEGVASESR&lt;sup&gt;218&lt;/sup&gt;</td>
<td>SETQEDTAETQSQED</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>12</td>
<td>G&lt;sup&gt;319&lt;/sup&gt; DNDNNTSQTGQR&lt;sup&gt;332&lt;/sup&gt;</td>
<td>1504.6</td>
<td>GDNDNNTSQTGQR</td>
<td>1504.6</td>
<td>GDNDNNTSQTGQR</td>
</tr>
<tr>
<td>13</td>
<td>S&lt;sup&gt;422&lt;/sup&gt; KEESESTGSTSSS EEDNHPKNIEADNR&lt;sup&gt;449&lt;/sup&gt;</td>
<td>3049.3</td>
<td>SKEESESTGSTSSS EEDN&lt;sup&gt;·&lt;/sup&gt;</td>
<td>3094.5</td>
<td>SKEESESTGSTSSS EEDNHPKNIEADNR</td>
</tr>
<tr>
<td>14</td>
<td>N&lt;sup&gt;443&lt;/sup&gt; IEADNR&lt;sup&gt;449&lt;/sup&gt;</td>
<td>NIEADNR</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>15</td>
<td>K&lt;sup&gt;450&lt;/sup&gt; LIVDAYHKPI GDQDDNDCQDGY&lt;sup&gt;473&lt;/sup&gt;</td>
<td>2735.2</td>
<td>KLIIVDAYHKPI GDQDDN&lt;sup&gt;·&lt;/sup&gt;</td>
<td>2792.1</td>
<td>KLIIVDAYHKPI GDQDDNDCQDGY</td>
</tr>
</tbody>
</table>

Fig. 5B
cDNA deduced “full-length” DMP1

\[ L^1_{\text{37K-173AA}} F^{173} \downarrow D^{174} 57K-300AA Y^{473} \]
\[ L^1_{\text{37K-180AA}} S^{180} \downarrow D^{181} 57K-293AA Y^{473} \]
\[ L^1_{\text{37K-217AA}} S^{217} \downarrow D^{218} 57K-256AA Y^{473} \]
\[ L^1_{\text{37K-221AA}} Q^{221} \downarrow D^{222} 57K-252AA Y^{473} \]

Fig. 7A

37K \leftrightarrow 57K

37K-173AA \leftarrow ESHGDGSEF \downarrow DDEGMQSDDP \rightarrow 57K-300AA
37K-180AA \leftarrow SEFDDEGMQS \downarrow DDPGSTRSDR \rightarrow 57K-293AA
37K-217AA \leftarrow DHEPTSTQDS \downarrow DDSQDVEFSS \rightarrow 57K-256AA
37K-221AA \leftarrow TSTQDSDDSQ \downarrow DVEFSSRKSF \rightarrow 57K-252AA

Fig. 7B
155  YRVGGGSEGESSHGDGSFDDDEGMQSDDPGSTRSDRGHTR  Rat
    Q********Q**************************E********A*  Mouse
   HW********D************************AY**E**NS*  Bovine
   HW*****D********L**********E*I**E**NS*  Human

195  MSSAGIRSEESKGDH-EPTSTQDSDDSDQDFSSRKSFRR  Rat
  ****************R-**********S********** Mouse
     I**DK*TQ*****DE*QA*****HE*PAAYPR**F**K  Bovine
*N***MK*K**-ENS*QAN*****GG**LL*HP***I**K  Human

Fig. 7C
Evidence for the proteolytic processing of dentin matrix protein 1: Identification and characterization of processed fragments and cleavage sites
Chunlin Qin, Jan C. Brunn, Richard G. Cook, Ralph S. Orkiszewski, James P. Malone, Arthur Veis and William T. Butler

J. Biol. Chem. published online June 17, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M305315200

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

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/early/2003/06/17/jbc.M305315200.citation.full.html#ref-list-1