FAM20C-Mediated Phosphorylation of MEPE and Its Acidic Serine- and Aspartate-Rich Motif

Brian Christensen,1 Gitte N Schytte,1 Carsten Scavenius,1,2 Jan J Enghild,1,2 Marc D McKee,3 and Esben S Sørensen1,2

1Department of Molecular Biology and Genetics, Aarhus University, Aarhus, Denmark
2Interdisciplinary Nanoscience Center, Aarhus University, Aarhus, Denmark
3Faculty of Dentistry and Department of Anatomy and Cell Biology, McGill University, Montreal, Quebec, Canada

ABSTRACT

Matrix extracellular phosphoglycoprotein (MEPE) is expressed in bone and teeth where it has multiple functions. The C-terminus of MEPE contains a mineral-binding, acidic serine- and aspartate-rich motif (ASARM) that is also present in other noncollagenous proteins of mineralized tissues. MEPE-derived ASARM peptides function in phosphate homeostasis and direct inhibition of bone mineralization in a phosphorylation-dependent manner. MEPE is phosphorylated by family with sequence similarity 20, member C (FAM20C), which is the main kinase phosphorylating secreted phosphoprotein. Although the functional importance of protein phosphorylation status in mineralization processes has now been well-established for secreted bone and tooth proteins (particularly for osteopontin), the phosphorylation pattern of MEPE has not been previously determined. Here we provide evidence for a very high phosphorylation level of this protein, reporting on the localization of 31 phosphoresidues in human MEPE after coexpression with FAM20C in HEK293T cells. This includes the finding that all serine residues located in the canonical target sequence of FAM20C (Ser-x-Glu) were phosphorylated, thus establishing the major target sites for this kinase. We also show that MEPE has numerous other phosphorylation sites, these not being positioned in the canonical phosphorylation sequence. Of note, and underscoring a possible important function in mineralization biology, all nine serine residues in the ASARM were phosphorylated, even though only two of these were positioned in the Ser-x-Glu sequence. The presence of many phosphorylated amino acids in MEPE, and particularly their high density in the ASARM motif, provides an important basis for the understanding of structural and functional interdependencies in mineralization and phosphate homeostasis. © 2020 The Authors. JBMR Plus published by Wiley Periodicals, Inc. on behalf of American Society for Bone and Mineral Research.

KEY WORDS: ASARM; FAM20C; MEPE; PHOSPHORYLATION

Introduction

Matrix extracellular phosphoglycoprotein (MEPE) is highly expressed by osteoblasts, osteocytes, and odontoblasts, and its secretion by these cells plays multiple roles in mineralized tissue biology, including in the regulation of calcification.1,2 Human MEPE consists of 508 amino acids and is a member of the SIBLING protein family (small integrin-binding ligand, N-linked glycoproteins) that are also phosphorylated. The five SIBLING proteins are MEPE, dentin matrix protein-1, bone sialoprotein (BSP), dentin sialophosphoprotein, and osteopontin (OPN).3,4 The SIBLING proteins are intrinsically disordered, yet they share common structural features such as having an integrin-binding RGD sequence, many potential phosphorylation sites, an abundance of acidic amino acids (Asp and Glu), and a conserved acidic serine and aspartate-rich motif (ASARM).3,5 The SIBLING proteins play key roles in biomineralization, and their functionalities are highly influenced by post-translational modifications such as phosphorylation and proteolytic cleavage.3,6–8

Several activities have been demonstrated for MEPE in a variety of in vivo and in vitro systems. MEPE-deficient mice exhibit increased bone formation compared with WT mice.9,10 In other studies, regulation of mineralization by MEPE has been linked to the ASARM motif positioned near the C-terminus of the protein10,11; the ASARM peptide can be released from MEPE by cathepsin B-catalyzed proteolysis.12 In in vitro studies, the MEPE-ASARM peptide inhibits de novo hydroxyapatite formation in a cell-free system,17 and it inhibits mineralization of MC3T3-E1 osteoblast cultures and bone marrow stromal cell cultures.16,11 Furthermore, in tooth cell culture studies, MEPE-ASARM impairs odontogenic differentiation and matrix
mineralization of dental pulp stem cells. MEPE, and particularly the ASARM peptide, have also been implicated in X-linked hypophosphatemic rickets (XLH). XLH is caused by inactivating mutations of the cell-surface metalloprotease PHEX (phosphate-regulating gene with homologies to endopeptidases on the X-chromosome). ASARM peptides derived from the SIBLING proteins, and OPN in particular, are the only known substrates for PHEX. Accumulation of ASARM peptides has been observed in XLH patients and in hypophosphatemic mice leading to mineralization defects in bone and teeth. In addition, some evidence exists that ASARM peptides upregulate the expression of fibroblast growth factor 23, a circulating factor that inhibits reabsorption of phosphate in the kidneys and intestinal phosphate uptake leading to hypophosphatemia and decreased bone mineralization.

Posttranslational modifications of MEPE—such as phosphorylation and proteolytic processing—have significant effects on the biological functions of the protein. The full-length phosphorylated MEPE protein can act as a promoter of hydroxyapatite formation and growth in a cell-free in vitro gel-diffusion system. The phosphorylated MEPE-ASARM peptide is an effective inhibitor of osteoblast cell culture extracellular matrix mineralization, directly binding to and inhibiting hydroxyapatite mineral crystal growth, and having greatly reduced or absent inhibition activity after dephosphorylation. Similar to this, in a cell culture model designed to study the underlying mechanisms of dental defects observed in XLH, only phosphorylated MEPE-derived ASARM peptides inhibited odontogenic cell differentiation and matrix mineralization. In terms of the enzymatic degradation of ASARM peptides, PHEX has been shown to rescue the inhibition of mineralization by both phosphorylated MEPE-ASARM and OPN-ASARM peptides, and the cleavage sites in ASARM have been determined. PHEX has also been shown to bind directly to the C-terminal ASARM motif of MEPE, which protects MEPE and its ASARM from proteolytic degradation. In this same work, the interaction between MEPE and PHEX was shown to be inhibited by phosphorylated MEPE-ASARM peptide, but much less effectively by nonphosphorylated MEPE-ASARM peptide. Similarly, the phosphorylated, but not the nonphosphorylated MEPE-ASARM inhibited the enzyme activity of PHEX. Collectively, these studies demonstrate the importance of phosphorylation in the different functions of MEPE, as well as for the MEPE-ASARM peptide.

Family with sequence similarity 20C (FAM20C) is the main kinase phosphorylating secreted phosphoproteins; it is highly expressed in mineralized tissues. FAM20C is a serine/threonine kinase with the canonical recognition motif being Ser-x-Glu/pSer (phosphoserine), but phosphorylation of secreted proteins has also been reported at a number of other motifs. The SIBLING proteins, including MEPE, are well-established substrates for FAM20C. MEPE contains many serine residues having FAM20C recognition motifs, and the C-terminal ASARM motif contains a polyserine stretch that potentially could be targeted by the kinase. Heretofore, no studies have investigated the phosphorylation of MEPE and its ASARM peptide, either in vivo or in vitro. Previous work has used synthetic phosphorylated ASARM peptides to provide important functional information, but no information is yet available on actual phosphorylation sites in MEPE. Given this context of the importance of SIBLING protein phosphorylation, we have investigated the phosphorylation pattern for MEPE in an attempt to understand the mechanisms of its (and its peptides) mineralization-regulating function and interactions with PHEX.

Here, we characterize the phosphorylation of MEPE coexpressed with FAM20C in human embryonic kidney 293 cells (HEK293T), and identify a remarkably high total of 31 residues that can be phosphorylated by FAM20C, of which nine are clustered in the ASARM motif.

Materials and Methods

Cloning

Human FAM20C cDNA encoding residues 1–584 was amplified by PCR using the forward primer 5’ ACCAAAGCTGCTAGATAT-GAAGATGATGCTGGTGCGCCG 3’ and the reverse primer 5’ GTTCGCGGCCAAGCTTTCTGGCCAGGGCCGCTCG 3’ (LGc Biosearch Technologies, Risskov, Denmark). Amplified cDNA was cloned into the pcDNA3.1/myc His(−)A (Invitrogen, Carlsbad, CA, USA) via NheI and HindIII restriction sites using the In-Fusion HD cloning technology (TaKaRa Bio, Otsu, France). MEPE-pcDNA3.1/myc His(−)A (residue 1–525 including signal peptide) was purchased from GenScript (Piscataway, NJ, USA). These constructs encoded human FAM20C and MEPE, both followed immediately by residues KLG, the myc epitope (EQKLISEEDL), residues NSAVD and HK6. The MEPE- (K509A/K515A-) plasmid was prepared by GenScript by mutation of K509L to ALGP and EQK115LISEEDL to EQALISEEDL in the MEPE plasmid.

The plasmids were propagated in Escherichia coli STELLAR competent cells (TaKaRa Bio), and all constructs were verified by sequence analysis. Plasmid DNA for transfection was prepared using PureLink HiPure Plasmid Filter Maxiprep (Invitrogen).

Protein expression and purification

HEK293T cells were maintained in Dulbecco’s modified Eagle’s medium with GlutaMAX (Invitrogen), supplemented with 10% FBS and 1% antibiotics (penicillin/streptomycin). HEK293T cells (8 × 10^6) were seeded in T25-cell-culture flasks and grown to confluency of 30% to 70%; at which point, the medium was changed to Opti-MEM-reduced serum medium supplemented with 5% FBS and 1% penicillin/streptomycin (all from Gibco, Grand Island, NY, USA). Then 4-μg plasmid DNA in Opti-MEM was mixed with lipofectamine LTX and PLUS transfection reagent (Invitrogen), and added to the cells. For cotransfection, 1-μg FAM20C plasmid DNA and 4-μg MEPE plasmid DNA were applied. After 48 hours, the supernatants were collected and centrifuged at 8,000g for 20 min.

His-tagged recombinant MEPE was purified using a metal-chelate affinity column (1 mL; QIAGEN, Valencia, CA, USA) charged with nickel ions. Bound protein was eluted with 250mM imidazole in PBS. The fractions containing MEPE were identified by Western blotting.

Western blotting

Recombinant MEPEs were loaded onto NuPAGE 10% Bis-Tris gels (Invitrogen), fractionated by SDS/PAGE, and electrophoretically transferred onto polyvinylidene difluoride membranes for immunodetection. The membranes were blocked in 2% Tween in Tris-buffered saline before addition of a c-myc monoclonal antibody (1 μg/mL) (9E10; Thermo Fisher Scientific, Waltham, MA, USA). MEPE was detected with alkaline phosphatase (ALP)- conjugated secondary immunoglobulins.
Phospho-imaging (pIMAGO) detection of protein phosphorylation

Protein phosphorylation was detected with the pIMAGO HRP Phosphoprotein Detection Kit (Tymora Analytical, West Lafayette, IN, USA). Proteins were separated by SDS/PAGE and electrophoretically transferred onto a polyvinylidene fluoride membrane. The membrane was blocked with 0.5% BSA and 0.1% polyamidoamine dendrimers. Next, the membrane was incubated with the pIMAGO Ti/3/biotin-dendrimer reagent and washed with 50mM 2,5-dihydroxybenzoic acid in 0.1% trifluoroacetic acid. Avidin-horseradish peroxidase was applied, and phosphoproteins were visualized with 3,3′-diaminobenzidine. For dephosphorylation, 1-μg MEPE was incubated with 30 μM of bovine ALP (Merck & Co., Kenilworth, NJ, USA) in 50mM ammonium bicarbonate overnight at 37°C.

Purification of ASARM peptides

There were 20-μg MEPE and MEPE-(K526A/K532A) mutant expressed with and without FAM20C, which were digested with trypsin (Merck) using an enzyme-to-substrate ratio of 1:50 (w/w), in 50mM ammonium bicarbonate, at 37°C for 6 hours. The tryptic digests were separated by reverse-phase high-performance liquid chromatography (RP-HPLC) on an Aeris C4 widepore column (Phenomenex, Torrance, CA, USA) connected to a Shimadzu HPLC System (Shimadzu, Kyoto, Japan). The peptides were separated at 40°C in 0.1% trifluoroacetic acid (TFA; buffer A) and eluted with a gradient of 60% acetonitrile in 0.1% TFA (buffer B). The gradient was developed over 59 min (0 to 5 min: 0% buffer B; 5 to 49 min: 0-98% buffer B; 49 to 59 min: 98% buffer B) at a flow rate of 0.85 mL/min. The peptides were detected by measuring the absorbance at 226 nm and the resulting fractions were collected for analyses.

To identify fractions containing myc-tagged peptides, 20 μL of each fraction were mixed with 80-μL PBS and coated on 96-well MaxiSorp immunoassay plates (Thermo Fisher Scientific) overnight at 4°C. The plate was extensively washed with PBS and blocked with 2% ovalbumin (Merck) in PBS for 1 hour at 37°C. All washes and dilutions after this were performed in PBS-Tween (10 mM disodium phosphate anhydrate (pH 7.2), 150 mM NaCl, 0.05% (v/v) Tween 20). Next, 0.25-μg c-myc monoclonal antibody (9E10; Thermo Fisher Scientific) were added to each well, followed by incubation with horseradish peroxidase-conjugated polyclonal goat anti-mouse IgG (DAKO, Carpentaria, CA, USA) (diluted 1:1000). Both antibodies were incubated for 1 hour at 37°C. Color development was obtained with TMB-one substrate (Kem-En-Tec Diagnostics, Taastrup, Denmark), and the reaction was stopped by the addition of 0.2M H2SO4. Color intensity was measured at 450 nm using an ELISA reader (BioTek, Winooski, VT, USA). Fractions containing myc-tagged peptides were further analyzed by MALDI-MS. For dephosphorylation, 20 μL of HPLC fractions containing ASARM peptides were dried in a vacuum centrifuge and incubated with ALP as described above.

Mass spectrometry

MALDI-MS was performed using a Bruker Autoflex III instrument operated in linear mode and calibrated using Protein Calibration Standard I (Bruker Daltonics, Billerica, MA, USA). Samples for MS analyses were prepared by mixing the fraction to be analyzed with 2,5-dihydroxybenzoic acid (10 mg/mL) in a 1:1 ratio directly on the MS-target probe. The theoretical peptide masses were calculated using the GPMAW program (Lighthouse Data, Odense, Denmark).

LC–MS/MS analysis of trypsinized MEPE samples was performed on a QExactive+ mass spectrometer (Thermo Fisher Scientific) equipped with an EASY-nLC 1000 (Thermo Fisher Scientific). All samples were micropurified using 3-M Empore C18 discs (3M Co., Saint Paul, MN, USA) packed in pipette tips. The lyophilized samples were dissolved in 0.1% formic acid, injected on a trap column (2 cm × 100 μm inner diameter), and separated on an analytical column packed inhouse in a pulled emitter (15 cm × 75 μm inner diameter). All columns were packed with ReproSil-Pur C18-AQ 3-μm resin (courtesy of Dr Marisich, Ammerbuch-Entringen, Germany). Peptides were eluted directly into the mass spectrometer with a linear gradient from 5% to 35% solvent B (0.1% formic acid in acetonitrile) for 20 min with a flowrate of 250 nL/min, followed 10 min with 100% solvent B. The acquired raw data was converted to Mascot Generic Format using the Raw Converter (version 1.1.0.18). The generated peak lists were searched against the Swiss-Prot database using an inhouse Mascot search engine (Matrix Science, London, UK). The following search parameters were used: Homo sapiens, trypsin, maximal two missed cleavages, carbamidomethyl as fixed modification, and oxidation and phosphorylation as variable modifications. Peptide tolerance was set to 10 ppm and MS/MS tolerance was set to 0.1 Da. MS/MS spectra of the targeted phosphorylated peptides were validated manually.

**Fig 1** Schematic representation of the two matrix extracellular phosphoglycoprotein (MEPE) variants used in this study. The recombinant human MEPE proteins were expressed in HEK293T cells and contained a C-terminal myc- and -His tag. The numbering of amino acids corresponds to the position in the MEPE sequence consisting of 508 amino acids. The integrin-binding RGD sequence, serines located in the FAM20C recognition motif (Ser-x-Glu), and the acidic serine- and aspartate-rich motif (ASARM) are indicated in the figure. The two lysine residues mutated to alanine in MEPE (K509A/K515A) are shown in bold.

---

**JBMR® Plus**

PHOSPHORYLATION OF MEPE 3 of 8
Multiple sequence alignment

Sequence analyses were performed on MEPE from the following 14 mammalian species found in the UniProt database https://www.uniprot.org/ (UniProt release 2020_01): Human, Homo sapiens (Q9NQ76); chimpanzee, Pan troglodytes (A0A2I3ST89); Bornean orangutan, Pongo pygmaeus (D6C6N8); bovine, Bos taurus (E1BG36); pig, Sus scrofa (I3LGA7); horse, Equus caballus (F6XRW9); sheep, Ovis aries (W5NXZ2); rabbit, Oryctolagus cuniculus (G1SG94); cat, Felis catus (A0A337S4X3); dog, Canis lupus familiaris (D6C6L9); guinea pig, Cavia porcellus (D6C6M4); mouse, Mus musculus (Q8K4L6); rat, Rattus norvegicus (Q9ES02); and dolphin, Lipotes vexillifer (A0A340VT16). Sequence alignment was performed with the UniProt website alignment tool.

Results

To investigate the FAM20C-mediated phosphorylation of MEPE, two recombinant forms of the MEPE protein (Fig. 1) were coexpressed with FAM20C in HEK293T cells. First, HEK293T cells were cotransfected with FAM20C and MEPE, and the resulting phosphorylation of MEPE was analyzed using pIMAGO, which specifically detects phosphorylated proteins. The pIMAGO reagent recognized MEPE coexpressed with FAM20C, but not MEPE expressed without the kinase (Fig. 2). Following treatment with ALP, the MEPE coexpressed with FAM20C was not detected in the pIMAGO assay, indicating FAM20C phosphorylation of MEPE. As control, an anti-c-myc antibody detected all MEPE forms (Fig. 2).

Table 1 LC–MS/MS Characterization of Phosphorylated Matrix Extracellular Phosphoglycoprotein (MEPE) Peptides

| Peptide                      | P residue | ppm | Counts | Expect | Conservation (%) |
|------------------------------|-----------|-----|--------|--------|------------------|
| 37^INQELSSKENIVQER^51        | 42        | 0.14| 2      | 1,60E-07| 36               |
| 36^INQELSSKENIVQER^51        | 43        | 0.9 | 39     | 1,10E-08| 21               |
| 52^KDLSLEASEN^64            | 56        | 0.055| 7      | 2,90E-06| 64               |
| 51^KDSLLEASEN^64            | 58        | 0.56 | 2      | 2,90E-04| 21               |
| 54^DSLLEASENKGSSK^68        | 61        | 1   | 7      | 1,20E-06| 43               |
| 54^DSLLEASENKGSSK^68        | 66        | 0.55 | 1      | 4,20E-04| 86               |
| 103^STGNKFGEDGDDAISK^118     | 103 or 104| 0.11| 1      | 6,00E-07| 79,79            |
| 158^NVNLIAPSMNYAK^71        | 166       | 2.74| 3      | 4,30E-05| 36               |
| 214^KIPSDFEGSTYTLQER^230     | 222       | 0.87| 20     | 6,10E-09| 100              |
| 231^GDNDSIPFSGDGQPK^246     | 236       | 0.06| 1      | 5,60E-05| 86               |
| 231^GDNDSIPFSGDGQPK^246     | 239       | 0.36| 6      | 4,90E-08| 100              |
| 263^DIQTGFAGPSEAESTHDLTK^282 | 272       | 0.92| 1      | 5,40E-07| 79               |
| 263^DIQTGFAGPSEAESTHDLTK^282| 276       | 1.02| 7      | 7,00E-12| 71               |
| 263^DIQTGFAGPSEAESTHDLTK^282| 277       | 2.75| 3      | 1,50E-06| 50               |
| 309^EADAVDSLSEGNSDIMGSTNK^331| 316       | 5.3 | 1      | 1,30E-09| 93               |
| 337^EGNVRDAGSQAHOQK^352      | 345       | 0.69| 1      | 2,70E-06| 100              |
| 355^VEFHYPAP^363            | 362       | 0.67| 30     | 4,30E-05| 93               |
| 366^RKEGSSDAESTYNEIPK^288    | 370 or 371| 1.43| 4      | 3,80E-07| 79,64            |
| 414^GKSQGLIPPSR^424         | 416       | 0.39| 4      | 8,20E-05| 64               |
| 432^NEMDSFNGPSHENITHGR^450  | 436       | 0.39| 5      | 1,60E-05| 93               |
| 432^NEMDSFNGPSHENITHGR^450  | 441       | 1.56| 31     | 4,6E-09 | 29               |
| 432^NEMDSFNGPSHENITHGR^450  | 436,441   | 1.1 | 1      | 5,60E-04| 93,29            |
| 467^GMPQGGKGSW^477           | 474       | 0.05| 2      | 4,80E-04| 86               |

Note. The phosphorylated residues are underlined and shown in bold. If the phosphorylation had an equal possibility to be located on two different residues, the residue is shown in italics. The sequence position of phosphorylated residues located in the FAM20C recognition motif Ser-x-Glu is shown in bold. The ppm differences between the measured and theoretical calculated masses are listed. The spectral counts include all identified peptides containing the specific phosphorylated residue. The shown expect value is for the highest-scoring peptide containing the residue. Conservation indicates how many of the analyzed mammalian MEPE sequences (Supplementary Fig. S1) contain the specific phosphorylated residue.
MEPE expressed alone, or coexpressed with FAM20C, in HEK293T cells was digested with trypsin, and the resulting peptides were characterized by LC–MS/MS. By this approach, 22 sites of phosphorylation (20 to 21 serine and 1 to 2 threonine residues) were identified in MEPE coexpressed with FAM20C (Table 1). Note, that this analysis did not include the ASARM motif, as the C-terminal part of MEPE was not detected in sequence, as the C-terminal part of MEPE was not detected in (Table 1). Note, that this analysis did not include the ASARM motif, as the C-terminal part of MEPE was not detected in (Table 1). Note, that this analysis did not include the ASARM motif, as the C-terminal part of MEPE was not detected in (Table 1). Note, that this analysis did not include the ASARM motif, as the C-terminal part of MEPE was not detected in (Table 1). Note, that this analysis did not include the ASARM motif, as the C-terminal part of MEPE was not detected in (Table 1). Note, that this analysis did not include the ASARM motif, as the C-terminal part of MEPE was not detected in (Table 1). Note, that this analysis did not include the ASARM motif, as the C-terminal part of MEPE was not detected in (Table 1). Note, that this analysis did not include the ASARM motif, as the C-terminal part of MEPE was not detected in (Table 1). Note, that this analysis did not include the ASARM motif, as the C-terminal part of MEPE was not detected in (Table 1). Note, that this analysis did not include the ASARM motif, as the C-terminal part of MEPE was not detected in (Table 1). Note, that this analysis did not include the ASARM motif, as the C-terminal part of MEPE was not detected in (Table 1). Note, that this analysis did not include the ASARM motif, as the C-terminal part of MEPE was not detected in (Table 1). Note, that this analysis did not include the ASARM motif, as the C-terminal part of MEPE was not detected in (Table 1). Note, that this analysis did not include the ASARM motif, as the C-terminal part of MEPE was not detected in (Table 1). Note, that this analysis did not include the ASARM motif, as the C-terminal part of MEPE was not detected in (Table 1). Note, that this analysis did not include the ASARM motif, as the C-terminal part of MEPE was not detected in (Table 1). Note, that this analysis did not include the ASARM moti...
The present study describes the phosphorylation of human MEPE coexpressed with FAM20C in kidney epithelial HEK293 cells, showing a remarkably high level of phosphorylation, similar to what we have shown previously for OPN. HEK293 cells have a very low endogenous expression of FAM20C, and the SIBLING proteins OPN, BSP, and MEPE expressed without FAM20C in HEK293 cells have previously been found not to be phosphorylated, thus demonstrating the specificity of this phosphorylation event for SIBLING proteins. Here, it was first shown that MEPE expressed in HEK293 cells without FAM20C was not recognized by the pIMAGO phosphorylation-detecting reagent (Fig. 2), thus indicating that there was no (or very little) endogenous phosphorylation of MEPE by this cell line. After coexpression with FAM20C, 22 phosphorylation sites were identified in the MEPE sequence by LC–MS/MS analysis (not including the ARSARM sequence). Spectral counts from the analyses indicated that the major FAM20C phosphorylation sites are Ser, Ser, Ser and Ser, in Ser-x-Glu motifs, the canonical phosphorylation sequence of FAM20C: These three serine residues constitute all residues located in this motif outside of the ARSARM sequence. Ser is not located in a motif that is considered a FAM20C phosphorylation site. MEPE also contains four serine residues located in Ser-x-Asp motifs, sites that have been shown to be phosphorylated in another SIBLING protein, OPN. Of these sites, only Ser, and potentially Ser, were found to be phosphorylated by FAM20C, and the spectral counts indicate that they are minor sites of phosphorylation. Collectively, this emphasizes that FAM20C preferably phosphorylates serine residues in Ser-x-Glu motifs. Nineteen phosphorylated residues were located in motifs not complying with the canonical consensus sequence of FAM20C, Ser-x-Glu/pSer (Table 1). This finding is in line with other studies that have shown that FAM20C is capable of phosphorylating residues in a great variety of motifs. Of note, FAM20C can phosphorylate serine residues in an SxQxxDEE motif where the acidic residues in position n + 5 to n + 7 are critical determinants. The two phosphorylated residues of MEPE, Ser and Ser, are located in sequences with acidic residues in the n + 5 to n + 7 positions.

LC–MS/MS analysis showed that Ser was the only phosphorylated residue in MEPE without coexpressed FAM20C. Ser, is located in a Ser-x-Glu motif, and the phosphorylation can most likely nevertheless be ascribed to the low endogenous expression of FAM20C by HEK293 cells. The phosphorylation was not detected by the pIMAGO reagent indicating that the level of phosphorylation is low (Fig. 2). In FAM20C KO mice, OPN has been shown to be still phosphorylated at FAM20C recognition motifs. This indicates that other kinases apart from FAM20C can phosphorylate the SIBLING proteins. In the present study, 22 phosphorylated residues were identified by LC–MS/MS in MEPE (not including the ARSARM sequence) after coexpression with FAM20C, whereas only a single phosphoserine residue was identified when MEPE was expressed alone. The overexpression of FAM20C and MEPE in the HEK293 cells can potentially result in phosphorylation of residues that are not phosphorylated in vivo. However, the results strongly indicate that FAM20C is the primary kinase responsible for MEPE phosphorylation, as has been shown for other SIBLING proteins, particularly for OPN.

The C-terminal part of MEPE containing the ARSARM peptide was not identified in the LC–MS/MS analysis most likely because of the acidic nature of the peptide. To identify and characterize peptides containing the ARSARM motif, we identified the MEPE(K509A/K515A) mutant with and without coexpression of FAM20C in HEK293 cells. The mutation of the two lysine residues made it possible to generate tryptic peptides of a size and charge that could be analyzed by MS. Moreover, the substitution of the lysine residues leads to tryptic peptides that contain the myc- and His-tag. The MS/MS analysis showed that Ser was the only phosphorylated residue in MEPE without coexpressed FAM20C. Ser, is located in a Ser-x-Glu motif, and the phosphorylation can most likely nevertheless be ascribed to the low endogenous expression of FAM20C by HEK293 cells. The phosphorylation was not detected by the pIMAGO reagent indicating that the level of phosphorylation is low (Fig. 2). In FAM20C KO mice, OPN has been shown to be still phosphorylated at FAM20C recognition motifs. This indicates that other kinases apart from FAM20C can phosphorylate the SIBLING proteins. In the present study, 22 phosphorylated residues were identified by LC–MS/MS in MEPE (not including the ARSARM sequence) after coexpression with FAM20C, whereas only a single phosphoserine residue was identified when MEPE was expressed alone. The overexpression of FAM20C and MEPE in the HEK293 cells can potentially result in phosphorylation of residues that are not phosphorylated in vivo. However, the results strongly indicate that FAM20C is the primary kinase responsible for MEPE phosphorylation, as has been shown for other SIBLING proteins, particularly for OPN.

The C-terminal part of MEPE containing the ARSARM peptide was not identified in the LC–MS/MS analysis most likely because of the acidic nature of the peptide. To identify and characterize peptides containing the ARSARM motif, we identified the MEPE(K509A/K515A) mutant with and without coexpression of FAM20C in HEK293 cells. The mutation of the two lysine residues made it possible to generate tryptic peptides of a size and charge that could be analyzed by MS. Moreover, the substitution of the lysine residues leads to tryptic peptides that contain the myc-tag located just after the ARSARM sequence (see Fig. 1). The ARSARM-containing peptide was efficiently phosphorylated by FAM20C, as MALDI-MS analysis showed that it contained up to
Sequence alignment showed that Ser222 and Ser362 were highly phosphorylation based on spectral counts (Table 1). Other residues indicated as minor FAM20C Ser-x-Glu motifs were only present in approximately one-third of conserved (Table 1, Fig. S1), whereas Ser43 and Ser441 located in different positions in MEPE expressed without FAM20C. This provides additional evidence for some endogenous FAM20C activity in HEK293 cells.

Ser43, Ser222, Ser362, and Ser441 were indicated to be major FAM20C phosphorylation based on spectral counts (Table 1). Sequence alignment showed that Ser222 and Ser362 were highly conserved (Table 1, Fig. S1), whereas Ser43 and Ser441 located in Ser-x-Glu motifs were only present in approximately one-third of the aligned sequences. Other residues indicated as minor FAM20C target sites, Ser46, Ser236, Ser239, Ser316, Ser345, Ser336, and Ser374, are conserved in more than 80% of the aligned sequences. Only Ser222, Ser239, and Ser345 are conserved in all species. In comparison, other important motifs like N-glycosylation sites and the integrin binding RGD-sequence are conserved in approximately 40% to 60% and approximately 93% of the sequences, respectively (Supplementary Fig. S1). The conservation of these serine residues suggests that they may be important for MEPE function even if they are not located in the FAM20C recognition motif. In particular, the alignment highlights the importance of the ASARM, as all aligned sequences contained nine serine residues except guinea pig, where only Ser694 is missing (Fig. 6, Supplementary Fig. S1).

Phosphorylation of the MEPE-ASARM peptide is a requirement for its ability to act as an inhibitor of mineralization, and for its involvement in phosphate homeostasis by binding and inhibiting the enzyme activity of PHEX. In these previous studies, a synthetic phosphorylated ASARM peptide, RDDSSESSDSGpSSpSDFSDGD, was utilized. This synthetic phosphorylated ASARM peptide contains three phosphoserines, which we show are also phosphorylated by FAM20C, as all nine serine residues in the ASARM are. In our study, we have used a genetically altered construct to facilitate purification and MS identification of the tryptic ASARM peptide. Additional experiments are needed to confirm the phosphorylation pattern in vivo.

The conservation of all serine residues in the ASARM motif (Fig. 6) indicates that phosphorylation could influence the functionality of the potent MEPE-ASARM peptide. For instance, the phosphorylation degree of the OPN-ASARM peptide influences its activity, where it was shown that PHEX addition to bone cell cultures rescued inhibition of mineralization by the peptide when containing three phosphoserines, but not when containing five phosphoserines. Taken together, these findings imply that the extent of phosphorylation of SIBLING proteins is likely used as a fine-tuning mechanism for a variety of functions in mineralized tissue biology. Paramount among these functions appears to be the strong influence of the phosphorylation status of SIBLING proteins such as MEPE and OPN in regulating extracellular matrix mineralization. Particularly noteworthy in these two cases of high phosphorylation for secreted proteins is that there is quite a spread (albeit with clusters) of phosphorylation sites along the full-length of the proteins. This provides a means by which to distribute additional negative charge (beyond many Asp and Glu residues) maximally across proteins that are intrinsically disordered and seemingly designed by nature to be flexible and versatile polyelectrolytes, capable of binding to the many different mineral phases found in different types of organisms across many phyla in the animal kingdom. Phosphorylation of secreted proteins is a relatively new area of research as compared with our knowledge on intracellular kinase phosphorylating pathways, yet recent and forthcoming data align well with these important roles for this posttranslational modification in the assembly and mineralization of extracellular matrices. Indeed, the importance of secreted-protein phosphorylation in mineralized tissues can be recently exemplified by studies showing that mutation of even the single phosphorylation site (Ser16) in the major secreted enamel protein (amelogenin) results in faulty mineralization involving precursor-phase transitions and disorganized and hypoplastic tooth enamel.

In summary, our findings describe the first full characterization of MEPE phosphorylation. We show that coexpression in a eukaryotic cell line of the SIBLING protein MEPE and the kinase FAM20C results in up to 31 different residues in MEPE being phosphorylated, the major sites being Ser43, Ser222, Ser362, and Ser441. Moreover, the C-terminal ASARM peptide of MEPE implicated in various mineralization functions contains nine serine residues that are maximally phosphorylated by FAM20C, further supporting the importance of this conserved, posttranslationally modified ASARM peptide in mineralized tissue biology.

Disclosures

MDM currently has ongoing funding discussions with Ultragenyx regarding its research and clinical programs for the treatment of X-linked hypophosphatemia. All other authors declare no conflicts of interest.

Acknowledgments

This work was supported by The Danish Council for Independent Research-Natural Sciences (Grant no. 4181-00131).

Fig 6 Sequence alignment of the acidic serine- and aspartate-rich motif (ASARM) sequence in matrix extracellular phosphoglycoprotein (MEPE). Sequence alignment of 14 mammalian MEPE ASARM sequences from the UniProt database (UniProt release 2020_01). Phosphorylated residues identified in this study are highlighted in black.

| Human   | ---RRDK-pSSpSDFSDGD 508 |
| Chimpanzee | ---RRDK-pSSpSDFSDGD 508 |
| Bornean orangutan | ---RRDK-pSSpSDFSDGD 508 |
| Bovine  | ---KKHSESSDSGpSSpSDFSDGD |
| Pig     | ---RKHSESSDSGpSSpSDFSDGD |
| Horse   | ---RKHSESSDSGpSSpSDFSDGD |
| Sheep   | ---RRDK-pSSpSDFSDGD 508 |
| Rabbit  | ---RKHSESSDSGpSSpSDFSDGD |
| Cat     | ---RKHSESSDSGpSSpSDFSDGD |
| Dog     | ---RKHSESSDSGpSSpSDFSDGD |
| Guinea pig | ---RRDK-pSSpSDFSDGD 508 |
| Mouse   | ---RRDK-pSSpSDFSDGD 508 |
| Rat     | ---RRDK-pSSpSDFSDGD 508 |
| Dolphin | ---KKHSESSDSGpSSpSDFSDGD |
Authors’ roles: BC designed and performed experimental work and wrote the manuscript. GNS and CS performed experimental work and data analysis. JJE assisted with MS and data analysis. MDM and ESS contributed to the conception and design of the study and wrote the manuscript. All authors have participated in the critical revision of the manuscript and approved the final version. BC, GNS, and ESS take responsibility for the integrity of the data analysis.

References
1. Rowe PSN, Kumagai Y, Gutierrez G, et al. MEPE has the properties of an osteoblastic phosphoglycoprotein and minihinin. Bone. 2004;34(2):303–19.
2. Rowe PSN. The chicken or the egg: PHEX, FGF23 and SIBLINGs unscrambled. Cell Biochem Funct. 2012;30(5):355–75.
3. Qin C, Bao O, Butler WT. Post-translational modifications of sibling proteins and their roles in osteogenesis and dentinogenesis. Crit Rev Oral Biol Med. 2004;15(3):126–36.
4. Fisher LW, Fedarko NS. Six genes expressed in bones and teeth encode the current members of the SIBLING family of proteins. Connect Tissue Res. 2003;44(Suppl 1):33–42.
5. Minamizaki T, Yoshiko Y. The bioactive acidic serine- and aspartate-rich motif peptide. Curr Protein Pept Sci. 2015;16(3):196–202.
6. Martin A, David V, Laurence JS, et al. Degradation of MEPE, DMP1, and osteoblastic phosphatonin and minihinin. Bone. 2004;34(2):1998–202.
7. Boskey AL, Chiang P, Fermanis A, et al. MEPE degradation of MEPE cleavage by Phex and accumulation of osteopontin fragments in Hyp mouse bone, the murine model of X-linked rickets (HYP). J Mol Med. 2005;3(15):10574–92.
8. Rappisberger J, Mann M, Ishihama Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat Protoc. 2007;2(8):1056–66.
9. Zhang SF, Li Y, Zhao W, et al. Secreted kinase phosphorylates extracellular proteins that regulate birefringeral biomechanization. Science. 2012;336(6085):1150–3.
10. Ishikawa HO, Xu A, Oghara E, Manning G, Irvine KD. The Raine syndrome protein FAM20C is a Golgi kinase that phosphorylates mineralization proteins. PLoS One. 2012;7(8):e42988.
11. Rappisberger J, Wiley SE, Guo X, et al. A single kinase generates the majority of the secreted phosphoproteome. Cell. 2015;161(7):1619–32.
12. Cozza G, Moro E, Black M, et al. The Golgi ‘casein kinase’ Fam20C is a genuine ‘phosvitin kinase’ and phosphorylates polyserine stretches devoid of the canonical consensus. FEBS J. 2018;285(24):4674–83.
13. Kato A, Kato N, et al. Secreted kinase phosphorylates extracellular proteins that regulate birefringeral biomechanization. Science. 2012;336(6085):1150–3.
14. Ishikawa HO, Xu A, Oghara E, Manning G, Irvine KD. The Raine syndrome protein FAM20C is a Golgi kinase that phosphorylates bio-

Acknowledgments: We acknowledge the support of the NIH intramural program and the NIH Center for Inherited Bone and Mineral Diseases (BMD1). We also wish to thank the reviewers for their comments and suggestions, which helped improve the manuscript. All authors have approved the final version of the manuscript.