Protein Science (1995), *4*:2040-2049. Cambridge University Press. Printed in the USA.
Copyright © 1995 The Protein Society

Posttranslational modifications of bovine osteopontin: Identification of twenty-eight phosphorylation and three *O*-glycosylation sites

ESBEN S. SØRENSEN,[1] PETER HØJRUP,[2] AND TORBEN E. PETERSEN[1]
[1] Protein Chemistry Laboratory, University of Aarhus, Science Park, DK-8000 Aarhus C, Denmark
[2] Department of Molecular Biology, University of Odense, DK-5230 Odense M, Denmark

(RECEIVED May 4, 1995; ACCEPTED July 25, 1995)

Abstract

Osteopontin (OPN) is a multiphosphorylated glycoprotein found in bone and other normal and malignant tissues, as well as in the physiological fluids urine and milk. The present study demonstrates that bovine milk osteopontin is phosphorylated at 27 serine residues and 1 threonine residue. Phosphoamino acids were identified by a combination of amino acid analysis, sequence analysis of *S*-ethylcysteine-derivatized phosphopeptides, and mass spectrometric analysis. Twenty-five phosphoserines and one phosphothreonine were located in Ser/Thr-X-Glu/Ser(P)/Asp motifs, and two phosphoserines were found in the sequence Ser-X-X-Glu/Ser(P). These sequence motifs are identical with the recognition sequences of mammary gland casein kinase and casein kinase II, respectively. Examination of the phosphorylation pattern revealed that the phosphorylations were clustered in groups of approximately three spanned by unphosphorylated regions of 11-32 amino acids. This pattern is probably of importance in the multiple functions of OPN involving interaction with Ca^{2+} and inorganic calcium salts. Furthermore, three *O*-glycosylated threonines (Thr 115, Thr 124, and Thr 129) have been identified in a threonine- and proline-rich region of the protein. Three putative *N*-glycosylation sites (Asn 63, Asn 85, and Asn 193) are present in bovine osteopontin, but sequence and mass spectrometric analysis showed that none of these asparagines were glycosylated in bovine mammary gland osteopontin. Alignment analysis showed that the majority of the phosphorylation sites in bovine osteopontin as well as all three *O*-glycosylation sites were conserved in other mammalian sequences. This conservation of serines, even in otherwise less well-conserved regions of the protein, indicates that the phosphorylation of osteopontin at specific sites is essential for the function of the protein.

Keywords: mineralization; *O*-glycosylation; osteopontin; phosphorylation; phosphoserine; phosphothreonine; *S*-ethylcysteine

The phosphorylation of serine, threonine, and tyrosine residues in intracellular proteins is an essential and well-documented mechanism in the regulation of cell physiology. Less attention has been directed to the localization and effects of phosphorylation in extracellular proteins. As the number of examples of phosphorylated proteins has intensified, it has become apparent that a majority of these contain multiple phosphorylations (Roach, 1991). In extracellular proteins, these multiple phosphorylations are often of structural importance in the formation of mineralized tissues. In this report we have identified the posttranslational modifications of bovine osteopontin, a multiphosphorylated extracellular protein that plays an important role in several normal and malignant calcification processes.

OPN is an acidic glycoprotein rich in aspartic acid, glutamic acid, and serine, where many of the serines are phosphorylated (for reviews, see Denhardt & Guo, 1993; Saavedra, 1994). OPN was first isolated from the mineralized matrix of bovine bone (Franzén & Heinegård, 1985) and has been characterized in calcified tissues from several species, including rat (Oldberg et al., 1986), human (Fisher et al., 1987), mouse (Smith & Denhardt, 1987), pig (Wrana et al., 1989), and chicken (Castagnola et al., 1991). In bone, OPN is synthesized by preosteoblasts, osteoblasts, and osteocytes, secreted into the osteoid, and subsequently incorporated in the bone (Mark et al., 1987a, 1987b). In addition to its presence in bone, OPN is also found in physiologic fluids such as urine (Kohri et al., 1992, 1993; Shiraga

Reprint requests to: Esben S. Sørensen, Protein Chemistry Laboratory, University of Aarhus, Science Park, Gustav Wieds Vej 10, DK-8000 Aarhus C, Denmark; e-mail: ess@biobase.dk.

Abbreviations: OPN, osteopontin; PDMS, plasma desorption mass spectrometry; MALDI-MS, matrix-assisted laser desorption ionization mass spectrometry; MGCK, mammary gland casein kinase; CKII, casein kinase II; RGDS, Arg-Gly-Asp-Ser; OPA, ortho-phthaldialdehyde; TFA, trifluoroacetic acid; PTH, phenylthiohydantoin.

et al., 1992) and milk (Senger et al., 1989b; Sørensen & Petersen, 1993a). In urine, OPN stereospecifically regulates the solubility of calcium salts (Shiraga et al., 1992), and it has been suggested that OPN is involved in the pathological calcification of urinary stones (Kohri et al., 1992, 1993) and atherosclerotic plaques (Giachelli et al., 1993; Ikeda et al., 1993; Shanahan et al., 1994). Furthermore, OPN is expressed in activated T-cells and metastasizing tumor cells where roles in the early resistance to bacterial infections (Fet et al., 1989) and binding of tumor cells at secondary sites (Craig et al., 1990), respectively, have been hypothesized.

The functions of OPN are dependent on the unique structure of the protein and especially the posttranslational modifications have been assigned significant importance. OPN isolated from rat bone has been estimated to contain 12 phosphoserines and 1 phosphothreonine (Prince et al., 1987), but the phosphorylations have not been localized in the amino acid sequence. Bovine OPN contains 41 serines, 17 threonines, and 2 tyrosines, constituting a total of 60 potential phosphorylation sites in the protein (Kerr et al., 1991). In earlier reports, we have described the isolation of OPN from bovine milk (Sørensen & Petersen, 1993a) and identified two phosphorylation motifs in the protein, namely that of the mammary gland casein kinase and that of casein kinase II (Sørensen & Petersen, 1993b, 1994). Here we report the complete in vivo phosphorylation pattern of bovine OPN, containing a total of 28 phosphorylation sites. Moreover, three *O*-glycosylation sites were identified in a proline-rich region of the protein. The nature and localization of the posttranslational modifications are discussed in relation to the biology of the protein.

Results

Strategy

The phosphorylation sites in bovine OPN were identified by a combination of *S*-ethylcysteine derivatization, sequence and mass spectrometric analysis (Sørensen et al., 1995). Peptides were obtained by endoproteinase Lys-C digestion and separated by reverse-phase HPLC (Fig. 1). Fractions containing phosphoserine and/or phosphothreonine residues (E1–E7) were identified by amino acid analysis and purified. Fraction E6 (Fig. 1) was subdigested with thermolysin prior to repurification (Fig. 2). Peptides containing phosphoserine were subjected to β-elimination followed by addition of ethanethiol, resulting in the conversion of phosphoserine to *S*-ethylcysteine (Meyer et al., 1986), thereby enabling localization of phosphoserines by Edman sequencing. In regions where sequencing was impaired, plasma desorption mass spectrometry or matrix-assisted laser desorption ionization mass spectrometry was used to assign phosphorylations and all phosphorylations were verified by mass spectrometric analysis. The combined results are shown in Table 1.

Localization of phosphoserines

Bovine OPN contains 41 serine residues, of which 23 are located in Ser-X-Glu/Ser(P) sequences. Amino acid sequencing of ethanethiol-treated phosphopeptides combined with mass spectrometric analysis of the underivatized peptides showed that 22 of these serines (Ser 8, Ser 10, Ser 11, Ser 44, Ser 46, Ser 47, Ser 60, Ser 62, Ser 65, Ser 99, Ser 105, Ser 108, Ser 172, Ser 189, Ser 194, Ser 217, Ser 224, Ser 229, Ser 240, Ser 251, Ser 256, Ser 258) were phosphorylated in milk OPN. PTH-*S*-ethylcysteine was identified in the corresponding sequence cycles in all cases except two (Ser 65 and Ser 229). The absence of *S*-ethylcysteine at Ser 65 could be explained by the neighboring Pro 66, which interferes with the ethanethiol addition. At Ser 229, no *S*-ethylcysteine is formed as the residue is located as the C-terminal amino acid in the peptide (peptide 17, Fig. 3) (Meyer et al., 1991). Nevertheless, mass spectrometric analysis of the respective peptides (Table 1) enabled us to assign phosphorylations at Ser 65 and Ser 229. Ethanethiol treatment of peptide 13 (Fig. 3) and subsequent sequencing gave no PTH-*S*-ethylcysteine at the position corresponding to Ser 198, showing that this serine

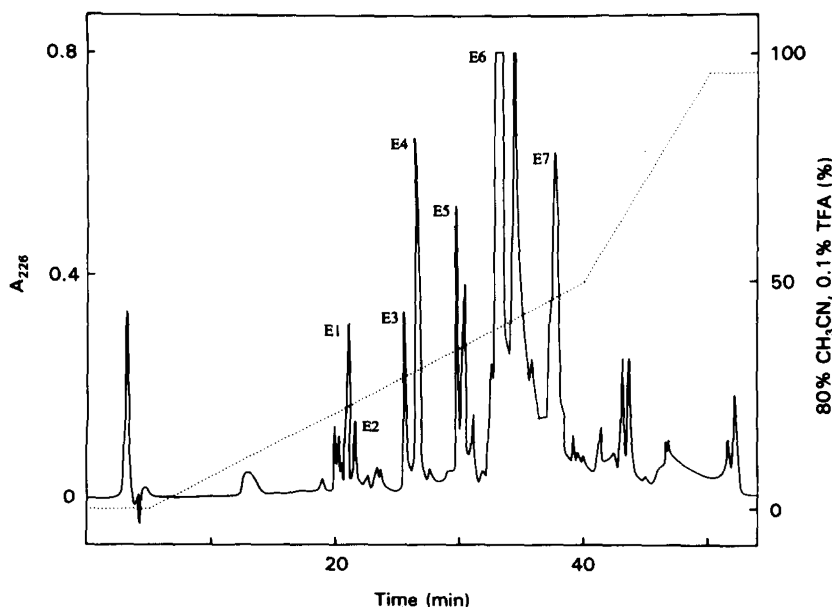

Fig. 1. Reverse-phase HPLC separation of an endoproteinase Lys-C digest of OPN. OPN was digested with endoproteinase Lys-C as described in the Materials and methods. Peptides were eluted with a gradient of 80% acetonitrile in 0.1% TFA (dotted line) on a Vydac C_{18} (10 μm) column (4 × 250 mm). The column was operated at 40 °C and the flow rate was 0.85 mL/min. Peptides were detected in the effluent by recording the absorbance at 226 nm (solid line) and collected manually. Phosphate-containing fractions were identified by amino acid analysis as described in the Materials and methods.

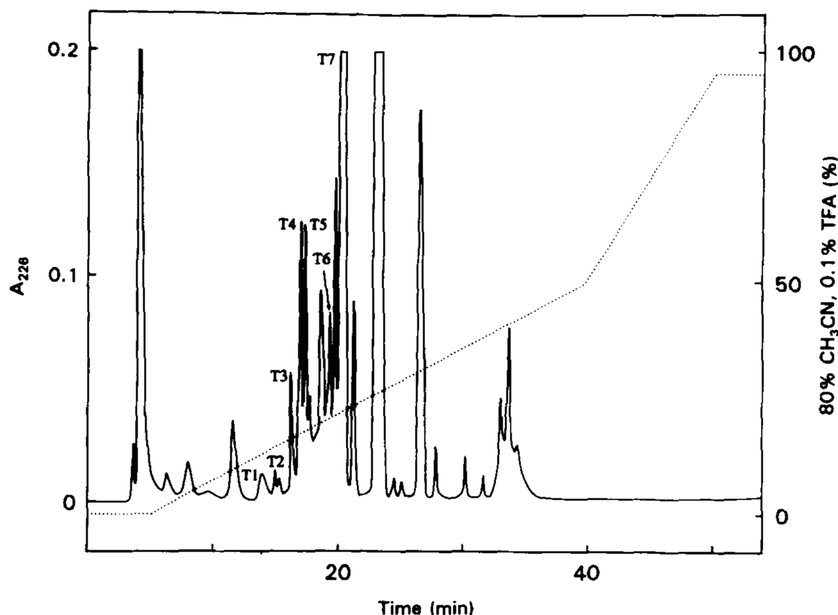

Fig. 2. Reverse-phase HPLC separation of a thermolysin digest of fraction E6 from the endoproteinase Lys-C digest (Fig. 1). Fraction E6 (Fig. 1) was digested with thermolysin as described in the Materials and methods. The digest was separated by reverse-phase HPLC as described in Figure 1.

Table 1. *Peptides from osteopontin*

Peptide no.[a]	Peak designation[b]	Sequence[c]	Observed MW[d]	Calculated MW[e]	Mass difference[f]
1	E1	L^{1}-K^{14}	1,766.1 (P)	1,445.6	320.5 (~4 H_2PO_3−)
2	–[g]	Y^{20}-K^{35}	–[h]		
3	E3	Q^{36}-K^{54}	2,350.4 (P)	2,110.2	240.2 (~3 H_2PO_3−)
4	E6T6	T^{57}-Q^{68}	1,558.4 (L)	1,316.4	242.0 (~3 H_2PO_3−)
5	E6T7	T^{57}-D^{71}	–[i]		
6	E6T4	L^{72}-D^{81}	1,232.7 (L)	1,151.0	81.7 (~1 H_2PO_3−)
7	E6T3	V^{82}-T^{92}	1,328.0 (L)	1,166.1	161.9 (~2 H_2PO_3−)
8	E6T5	T^{93}-E^{110}	2,294.8 (L)	2,053.9	240.9 (~3 H_2PO_3−)
9	E7	V^{111}-T^{119}	–[h]		
10	E7	I^{120}-A^{132}	–[h]		
11	–[g]	G^{137}-K^{145}	–[h]		
12a[j]	E5	F^{151}-N^{155}	679.1 (P)	678.8	
12b[j]	E5	V^{156}-D^{165}	1,170.8 (P)	1,090.9	79.9 (~1 H_2PO_3−)
12c[j]	E5	F^{166}-K^{181}	2,046.7 (P)	1,886.1	160.6 (~2 H_2PO_3−)
13	E2	T^{182}-K^{199}	2,152.7 (P)	1,992.1	160.6 (~2 H_2PO_3−)
14	–[k]	A^{205}-N^{262}	7,268.8 (L)	6,712.2	556.6 (~7 H_2PO_3−)
15	E6T7	L^{214}-N^{220}	–[i]		
16	E6T2	L^{223}-E^{226}	–[i]		
17	E6T1	F^{227}-S^{229}	469.5 (P)	389.4	80.1 (~1 H_2PO_3−)
18	E6T5	L^{236}-K^{244}	1,180.2 (P)	1,100.2	80.0 (~1 H_2PO_3−)
19	E4	I^{248}-N^{262}	1,897.5 (P)	1,656.8	240.7 (~3 H_2PO_3−)

[a] Peptide numbers correspond to those of Figure 3.
[b] Peak designations correspond to those of Figures 1 and 2. E numbers specify elution position in the endoproteinase Lys-C digest (Fig. 1) and T numbers specify elution position in the thermolysin digest (Fig. 2).
[c] Amino acid sequence identified by sequence and/or mass spectrometric analysis.
[d] Molecular mass determined by mass spectrometric analysis. (P) denotes mass determined by PDMS, (L) denotes mass determined by MALDI-MS.
[e] Calculated average masses.
[f] Difference between observed and calculated molecular weight. The number of phosphate groups corresponding to the mass difference is given in parentheses.
[g] Data on peptides 2 and 11 originate from an earlier study on osteopontin structure (Sørensen & Petersen, 1994).
[h] Because no phosphoamino acids were detected, mass spectrometric analyses of these peptides were not performed.
[i] No ions were detected by mass spectrometric analysis of the peptide.
[j] Peptides 12a–c are generated by thermolysin digestion of fraction E5 (Fig. 1).
[k] Peptide 14 represents the C-terminal part of osteopontin obtained by thrombin cleavage at Lys 204–Ala 205.

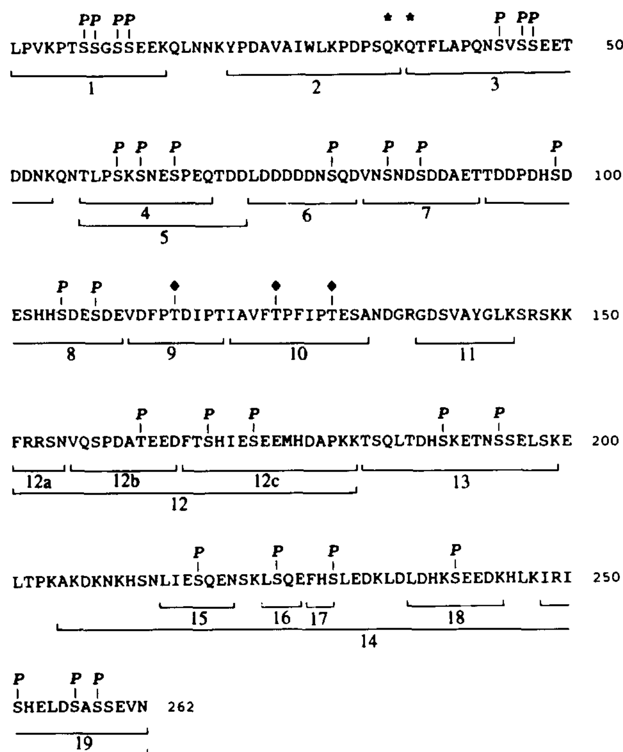

Fig. 3. Localization of posttranslational modifications in bovine OPN. The amino acid sequence was deduced from the cDNA sequence of bovine OPN (Kerr et al., 1991). A single sequence discrepancy was observed at position 40, which we identified as an alanine. Solid lines indicate isolated and characterized peptides. P denotes identified phosphorylation. Filled diamonds symbolize identified glycosylation. Asterisks symbolize transglutaminase-reactive glutamines (Sørensen et al., 1994). Peptides are numbered consecutively and references to elution profiles are given in Table 1.

residue, though located in a Ser-X-Glu sequence, is not phosphorylated. This finding was verified by mass spectrometric analysis of the peptide (Table 1), which showed a mass of 2,152.7 Da, corresponding to the peptide with two phosphorylations (Ser 189, Ser 194).

Four serines are located in Ser-X-Asp sequences (Ser 79, Ser 84, Ser 87, and Ser 158). Ethanethiol treatment and subsequent sequencing of peptide 7 (Fig. 3) showed PTH-*S*-ethylcysteine in cycles corresponding to Ser 84 and Ser 87. In the case of Ser 79 (peptide 6, Fig. 3), *S*-ethylcysteine was not clearly detected at the corresponding position due to difficulties in sequencing through the aspartic acid-rich region preceding this residue. However, mass spectrometric analysis of the peptide showed a mass discrepancy of 81.7 Da compared to the calculated peptide mass (Table 1). This difference can be accounted for by a phosphorylation at Ser 79, the only hydroxyamino acid in the peptide. Sequencing of peptide 12b (Fig. 3) gave normal yield of PTH-Ser in the cycle corresponding to Ser 158. Furthermore, mass spectrometric analysis of the peptide showed a mass consistent with the calculated peptide mass, thereby confirming that Ser 158 is not modified.

Four serines (Ser 7, Ser 102, Ser 168, and Ser 221), not counting those located in Ser-X-Glu/Ser(P) sequences, are located in Ser-X-X-Glu/Ser(P) sequence motifs. Two of these serines were found to be phosphorylated in bovine OPN. Amino acid sequencing of the ethanethiol-treated peptides (peptides 1 and 12c, Fig. 3) showed PTH-*S*-ethylcysteine in cycles corresponding to Ser 7 and Ser 168, respectively. Sequencing and mass spectrometric analysis of peptide 8 (Table 1) demonstrated that Ser 102 was not modified in bovine OPN. Concerning Ser 221, mass spectrometric analysis of the C-terminal part of OPN (peptide 14, Fig. 3 and Table 1), generated by thrombin cleavage at Lys 204–Ala 205, showed a mass divergence of 556.6 Da from the calculated peptide mass, corresponding to seven phosphorylations in the peptide. All of these phosphorylations have been accounted for (Ser 217, Ser 224, Ser 229, Ser 240, Ser 251, Ser 256, and Ser 258), thereby indirectly showing that Ser 221 is not modified.

Localization of phosphothreonines

The method used for conversion of phosphoserines is not applicable to phosphothreonines, which have to be identified by other methods. Bovine OPN contains 17 threonines, of which three (Thr 6, Thr 162, and Thr 192) are located in Thr-X-Glu/Ser(P) sequences. Sequencing of peptides 1 and 13 (Fig. 3) gave expected yields of PTH-threonines in cycles corresponding to Thr 6 and Thr 192, respectively, showing that these threonines are not phosphorylated. Mass spectrometric analysis of the peptides (Table 1) confirmed that these threonines were not modified. Sequencing of peptide 12 (Fig. 3) gave no PTH-threonine in the cycle corresponding to Thr 162, suggesting that this residue is modified. To verify this assumption, the peptide was digested with thermolysin and the resulting peptides were characterized (peptides 12a–c, Fig. 3 and Table 1). The peptide corresponding to residues 156–165 was identified and sequenced (peptide 12b, Fig. 3). Mass spectrometric analysis of the peptide showed a mass of 1,170.8 Da, consistent with a phosphorylation at Thr 162. In addition, hydrolysis of the peptide and subsequent separation of the amino acids on the Amino Quant system (described in the Materials and methods) resulted in unequivocal identification of phosphothreonine in the peptide.

Four threonines in OPN are located in Thr-X-Asp sequences (Thr 50, Thr 69, Thr 92, and Thr 93). Sequence and mass spectrometric analysis of peptides containing these residues showed that none of these threonines are modified in the protein. In the case of Thr 69, no ions were identified by mass spectrometric analysis of peptide 5 (Fig. 3), but sequence analysis of the peptide gave the expected yield of PTH-Thr in the corresponding cycle, thereby showing that Thr 69 is not phosphorylated.

Localization of glycosylations

Three *O*-glycosylated threonines were identified in bovine OPN. Amino acid sequencing of peptides 9 and 10 (Fig. 3) gave no identifiable PTH-amino acids in cycles, corresponding to Thr 115, Thr 124, and Thr 129, respectively. Amino acid analysis showed that no phosphothreonine was present in either of the peptides. Furthermore, the analysis showed that *N*-acetyl-galactosamine was present in both peptides. These data reveal the presence of *O*-linked glycosylations at threonines 115, 124, and 129. Three asparagine residues in bovine OPN (Asn 63, Asn 85, and Asn 193) are located in the putative glycosylation sequence Asn-X-Ser/Thr. Amino acid sequence and mass spectrometric analysis of the involved peptides (peptides 4, 7, and 13, Fig. 3) showed that none of these asparagines are glycosylated in milk OPN.

Discussion

Localization of phosphorylation and glycosylation sites

The complete in vivo phosphorylation pattern of bovine mammary gland OPN, containing 27 phosphoserines and one phosphothreonine, has been determined. This number of phosphorylations is significantly higher than that reported for rat bone OPN (Prince et al., 1987). This difference could be ascribed to species variation, or more likely the existence of tissue-specific isoforms that differ in their phosphorylation pattern. Alternatively, bone matrix OPN might be dephosphorylated as part of an activity regulation mechanism (Price et al., 1994) or due to aging of the protein in the matrix.

The phosphorylation pattern found in bovine milk OPN shows that the primary requirement for serine/threonine phosphorylation is a negatively charged glutamate, aspartate, or phosphoserine in the $n + 2$ or $n + 3$ position. These sequence requirements are identical with the recognition sequences of MGCK (Mercier, 1981) and CKII (Kuenzel et al., 1987), respectively. Only two serines (Ser 158 and Ser 198), located in the MGCK recognition sequence and two serines (Ser 102 and Ser 221) located in the CKII recognition motif are found not to be phosphorylated in milk OPN. Contrary to this, only one threonine (Thr 162) out of seven located in the MGCK recognition sequence is phosphorylated in OPN. Compared to the high degree of phosphorylation of serines, this indicates an MGCK preference for serine residues. This is in agreement with earlier work on the specificity of mammary gland casein kinases, in which proteins with the sequence Thr-X-Glu were found to be poor substrates for the mammary gland enzyme (Bingham & Groves, 1979). Generally, the lack of phosphorylation of residues located in the same kinase recognition motif emphasizes that more information than merely the primary structure is required to identify in vivo phosphorylation sites. During this study we observed no ambiguity in the phosphorylation pattern of milk OPN as phosphorylated residues were never observed unphosphorylated or vice versa. However, it cannot be excluded that minor heterogeneity exists in the phosphorylation of certain individual sites.

The phosphorylation motif requiring an acidic residue in the $n + 2$ position was first observed in the caseins (Mercier, 1981). The same motif has been found for another milk phosphoprotein, component PP3 (Sørensen & Petersen, 1993b, 1993c). Recently, three and eight phosphoserines have been identified in shark, lamb, rat, bovine, and human matrix Gla protein (Price et al., 1994) and spp24 from bovine bone (Hu et al., 1995), respectively. These phosphoserines were all located in Ser-X-Glu/Ser(P) sequence motifs, which illustrates that a kinase that recognizes this acidic motif is not restricted to the mammary gland but is active in the bone mineralization environment as well.

It has been reported that recombinant mouse OPN is able to autophosphorylate tyrosine residue(s) (Ashkar et al., 1993) and protein-bound sulfate has been reported present in rat bone marrow OPN (Nagata et al., 1989). This sulfate was not removed by deglycosylation, indicating the presence of a sulfated tyrosine in mineralized OPN. Bovine OPN contains two tyrosines (Tyr 20 and Tyr 142). Sequencing of peptides 2 and 11 (Fig. 3) gave expected yields of PTH-tyrosine in cycles corresponding to Tyr 20 and Tyr 142, respectively, indicating that the tyrosines of bovine milk OPN are not phosphorylated. However, the lability of the sulfate ester linkage makes the identification of sulfated tyrosines by sequencing difficult, and the presence of sulfated tyrosine in milk OPN cannot be excluded at present.

OPN isolated from rat bone has been reported to possess more than 30 monosaccharides possibly present in one *N*-linked and five to six *O*-linked oligosaccharides (Prince et al., 1987). Bovine OPN contains three putative *N*-glycosylation sites (Asn 63, Asn 85, and Asn 193). Analysis shows that none of these asparagines are glycosylated in milk OPN. In this context it is interesting to note that all of the three putative *N*-glycosylation sequences are phosphorylated at serines located within the sequence motif. This could lead to speculations of blocking of *N*-glycosylation by phosphorylation. However, *N*-glycosylation is a cotranslational modification that takes place immediately after transcription in the endoplasmic reticulum, whereas phosphorylation usually occurs later in the protein maturing process. Hence, the blocking of *N*-glycosylation by phosphorylation seems unlikely. Supporting this, the phosphorylation of OPN synthesized by chicken osteoblasts has been shown to be restricted to the *trans* cisternae of the Golgi stack (Ashkar et al., 1994).

OPN contains a threonine–proline-rich region (residues 114–129) preceding the RGDS sequence. This region contains four evenly distributed threonines that are all adjacent to prolines (Thr 115, Thr 119, Thr 124, and Thr 129). Three of these threonines were found to be *O*-glycosylated (Thr 115, Thr 124, and Thr 129). This *O*-glycosylation of threonines adjacent to prolines has been observed in a number of proteins (Gooley et al., 1991). The nature of the carbohydrate moieties of OPN has not yet been examined. MALDI-MS of native OPN gave a mass of approximately 33,750 Da. Subtraction of the calculated mass of the amino acids and the 28 phosphate groups leaves approximately 2,250 Da for glycosylations corresponding to about 10–12 monosaccharides.

Based on the localization of phosphorylations, glycosylations and transglutaminase-reactive glutamines, a schematic model of bovine OPN has been constructed (Fig. 4). Examination of the model shows that the phosphorylations of OPN are not concentrated in any specific part of the protein but occur in clusters of approximately three phosphate groups, spanned by stretches of unmodified amino acids. This phosphorylation pattern probably plays a role in the interaction with Ca^{2+} ions and hydroxyapatite. Interestingly, the regions containing the transglutaminase-reactive glutamines, the glycosylations, the RGDS-cell adhesion, and the thrombin cleavage site are devoid of phosphorylations in the vicinity.

Thrombin cleavage of OPN

It has previously been shown that mouse, rat, and human OPN contain a thrombin cleavage site six amino acids to the C-terminal side of the RGDS sequence (Senger et al., 1988, 1989a). Our results with bovine OPN show a slightly different cleavage pattern. Reverse-phase separation of a thrombin digest of bovine OPN followed by sequence analysis showed proteolytic cleavage at Arg 147–Ser 148 and Lys 204–Ala 205 (results not shown). The Arg 147–Ser 148 bond is located eight amino acids to the C-terminal side of the RGDS sequence and corresponds to the observed cleavage site in mouse, rat, and human OPN. Thrombin cleavage corresponding to that observed at Lys 204–Ala 205 in bovine OPN has not been reported in other species. The func-

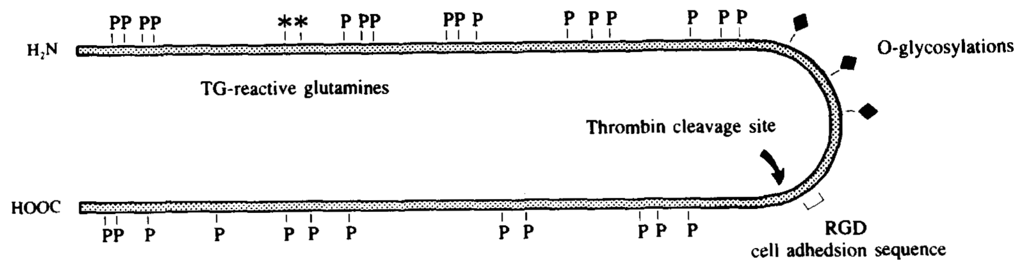

Fig. 4. Posttranslational modifications of bovine OPN. Schematic model of OPN based on the localization of the posttranslational modifications. Ps, diamonds, and asterisks are the same as in Figure 3. Major thrombin cleavage site and RGDS-cell adhesion sequence are indicated.

tional relevance of thrombin cleavage in vivo is not known. However, thrombin cleavage of recombinant OPN, lacking posttranslational modifications, destroyed the RGDS-cell adhesion activity (Xuan et al., 1994). In contrast, OPN from a rat tumor cell line showed enhanced RGDS-dependent cell-binding properties following thrombin cleavage (Senger et al., 1994).

Sites of posttranslational modification are conserved

OPN has been characterized from several species and the cDNA-derived amino acid sequences are known for human, cow, pig, mouse, rat, rabbit, and chicken OPN. An alignment of these sequences including the posttranslational modifications found in the bovine sequence is shown in Figure 5. The alignment reveals that the amino acids found to be modified in the bovine sequence are very well-conserved. At 21 of the 28 positions found to be phosphorylated in bovine OPN, a serine or threonine is found at the homologous position in all other known mammalian sequences. Moreover, the acidic amino acid motif that confers the kinase recognition site is preserved among the shown mammalian sequences in 23 of 28 cases. This high degree of conservation of serines and threonines, even in otherwise less well-conserved regions of the protein, indicates that the phosphorylation of OPN at specific sites is essential for the function of the protein. Furthermore, the conservation indicates that a phosphorylation pattern similar to that found in the bovine protein could be expected in OPN from other species.

Analysis of the threonine–proline-rich region containing the *O*-glycosylations (residues 114–129) revealed that the threonines corresponding to Thr 115, Thr 124, and Thr 129 in the bovine sequence are conserved in all known mammalian sequences, whereas Thr 119 is present only in the bovine sequence. Furthermore, the neighboring proline residues are also conserved at most positions. Therefore, this part of the sequence may represent a region with no other evolutionary restraint than the presence of threonines adjacent to prolines representing recognition sites for glycosyltransferases.

Functional significance of the phosphorylations

Immunohistochemical studies have localized OPN to the mineralization front in developing bone, which suggests that OPN is important in the process of matrix mineralization (McKee et al., 1990). OPN could exert its influence on the mineralization process either by functioning as a promoter of mineralization by nucleating hydroxyapatite crystal growth or by acting as an inhibitor/moderator of calcification. In vitro experiments have shown that OPN concentrations greater than 25 mg/L (~0.7 μM) inhibited both hydroxyapatite formation and crystal growth in a dose-dependent manner (Boskey et al., 1993). A number of studies have addressed the problem of identifying the molecular features in OPN responsible for this inhibitory activity. By removing 84% of the covalently bound phosphate with alkaline phosphatase, the hydroxyapatite-inhibiting activity was reduced more than 40-fold (Hunter et al., 1994). Furthermore, treatment with glycine ethyl ester in the presence of carbodiimide, thereby modifying 86% of the carboxylate groups in OPN, reduced the hydroxyapatite-inhibiting activity by sixfold. Thus, the hydroxyapatite-inhibiting activity involves both phosphate and carboxylate groups, although the phosphate groups seem to be quantitatively more important than the carboxylate groups. Other studies have shown that OPN is effectively dephosphorylated by tartrate-resistant acid phosphatase, resulting in a form that could no longer support osteoclast binding (Ek-Rylander et al., 1994).

An OPN concentration of approximately 0.1 μM is sufficient for inhibition of crystal growth in urine (Shiraga et al., 1992). This concentration is too low for inhibition to be on the basis of chelation of calcium. It therefore seems likely that crystal growth is inhibited by an interaction of OPN with the crystals. However, a recent study disputes these interpretations and reports that OPN may actually promote urinary stone formation (Kohri et al., 1993). From 1 L of bovine milk it is possible to isolate 20–40 mg OPN, corresponding to a minimum concentration of 0.6 μM. The function of OPN in milk is not clear but a function as an inhibitor of crystallization, similar to that in urine, is a possibility. Actually, it has been speculated that OPN acts as a general inhibitor of phase separation in physiological fluids of high supersaturation (Hunter et al., 1994).

Recently, OPN mRNA as well as the mature protein have been detected in human atherosclerotic plaques, where the expression was found to be associated closely with macrophages at sites of calcification (Giachelli et al., 1993; Shanahan et al., 1994). It is speculated that OPN, in cooperation with other acidic proteins expressed by macrophages, provides a milieu that favors calcium crystallization resulting in mineralization of the plaques, eventually leading to plaque rupture and subsequent thrombosis (Shanahan et al., 1994).

In summary, these findings suggest that OPN in calcification processes acts as a regulator of crystal growth, capable of both initiating and inhibiting calcium-salt crystallization, depending on the environment and state of the protein. To fulfill this role,

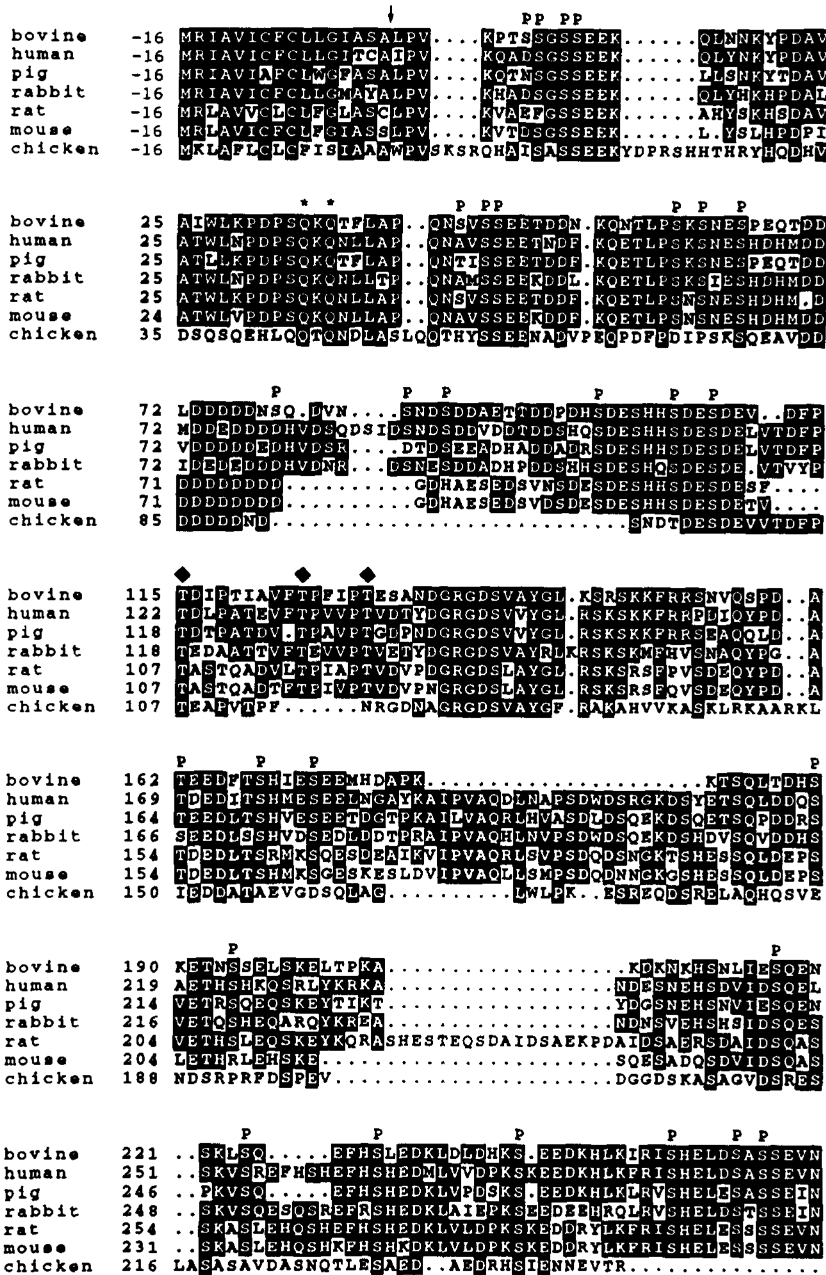

Fig. 5. Alignment of OPN sequences. The cDNA-derived amino acid sequences of bovine (Kerr et al., 1991), human (Kiefer et al., 1989), pig (Wrana et al., 1989), rabbit (Tezuka et al., 1992), rat (Oldberg et al., 1986), mouse (Smith & Denhardt, 1987), and chicken (Castagnola et al., 1991) OPN were aligned using the PileUp program from the GCG package (Genetics Computer Group, Madison, Wisconsin). Identical amino acids are shown in black and introduced gaps are indicated by dots. The arrow indicates signal peptide cleavage site. Ps, diamonds, and asterisks are the same as in Figure 3.

the phosphorylations of OPN are essential in the function of the protein.

The S-ethylcysteine derivatization technique

β-Elimination of peptide-bound phosphoserine followed by the addition of ethanethiol converts phosphoserine to *S*-ethylcysteine. However, location of the phosphoserine as the N- or C-terminal residue in the peptide results in the formation of ethylamine or pyruvate (Meyer et al., 1991).

Mass spectrometric analysis of peptide 4 (Fig. 3) showed that Ser 60, Ser 62, and Ser 65 are phosphorylated (Table 1). Ethanethiol treatment of the peptide gave *S*-ethylcysteine at residues corresponding to Ser 60 and Ser 62 but failed to convert Ser 65 to *S*-ethylcysteine. This lack of conversion is probably caused by Pro 66, which could convey structural hindrance during the elimination process (Sørensen et al., 1995). Similar to this, two phosphoserines in bovine κ-casein, both followed by prolines, failed to yield *S*-ethylcysteine when subjected to the ethanethiol treatment (Rasmussen, Sørensen, Peterson, Nielson, & Thomsen, unpubl. data). These results strongly indicate that β-elimination of phosphoserines followed by prolines is not possible.

Materials and methods

Materials

Bovine OPN was purified from bovine milk as described (Sørensen & Petersen, 1993a). Thermolysin and OPA were from Sigma

(St. Louis, Missouri). Endoproteinase Lys-C was from Boehringer Mannheim (Mannheim, Germany). Vydac C_{18} and C_4 (10 μM) were from The Separations Group (Hesperia, California). Capcell C_{18} (5 μM) was from Shiseida (Tokyo, Japan). Ethanethiol, tetrahydrofuran, and methanol were from Rathburn (Walkerburn, Scotland). Reagents used for sequencing were purchased from Applied Biosystems (Foster City, California). All other reagents were of analytical reagent grade. Thrombin was purified from bovine blood as described (Magnusson, 1965).

Generation of peptides

OPN was digested with endoproteinase Lys-C (EC 3.4.21.40) using an enzyme:substrate ratio of 1:250 (w/w) in 0.1 M ammonium bicarbonate, pH 8.0, at 37 °C for 18 h. Fractions E5 (residues 151–181) and E6 from the endoproteinase Lys-C digest were further digested with thermolysin (EC 3.4.24.2) at a concentration of 10 μg/mL in 0.1 M pyridine-acetate, pH 6.5, 5 mM $CaCl_2$, at 55 °C for 8 h. Thrombin digestion was performed using an enzyme:OPN ratio of 1:100 (w/w) in 0.1 M ammonium bicarbonate, pH 8.0, at 37 °C for 1 h.

Phosphoamino acid analysis

Peptides containing phosphoamino acids were detected by short-time hydrolysis and subsequent amino acid analysis. Peptides were hydrolyzed under vacuum at 110 °C for 2–3 h in the presence of 6 M HCl, 0.05% phenol, 1% thioglycolic acid. The amino acids were analyzed essentially as described by Barkholt and Jensen (1989).

Identification of phosphothreonine

To detect phosphothreonine in peptides, a system optimized for separation of acidic amino acids was employed. Peptides were hydrolyzed under vacuum at 110 °C for 7 h in the presence of 6 M HCl, 0.05% phenol, 1% thioglycolic acid. The hydrolysates were dried and solubilized in water prior to derivatization with OPA. Derivatization and analysis were performed on an Amino Quant system (Hewlett-Packard). The Amino Quant was programmed to mix a 5.0-μL sample with 5.0 μL 1 M K_3BO_3 and 1.0 μL OPA solution (stock solution: 50 mg OPA, 4 mL CH_3OH, 0.5 mL 1 M K_3BO_3, pH 10.4, and 50 μL ethanethiol). The sample was allowed to react with the derivation mixture for 2 min and subsequently separated on a Capcell C_{18} (5 μm) column (4 × 200 mm). The column was eluted with a stepwise linear gradient (0–5 min, 0–35% B; 5–13 min, 35–60% B; 13–16 min, 60–70% B) (buffer A: 14.65 mM H_3PO_4, 0.275% [w/w] tetrahydrofuran, pH 7.20; buffer B: 75% [w/w] CH_3OH). The column was operated at a flow rate of 0.6 mL/min at 40 °C. The OPA-amino acids were detected on an on-line Hewlett-Packard 1046A programmable fluorescence detector. At these conditions phosphothreonine was baseline separated from all other amino acids and eluted at 15.8 min.

Isolation of peptides

Peptides from the digests were separated by reverse-phase HPLC on a Vydac C_{18} column. The separation was carried out on a Pharmacia LKB system (Pharmacia, Uppsala, Sweden), consisting of a 2248 LC gradient pump connected to a 2252 LC controller and a 2510 Uvicord SD detector equipped with a 226-nm filter and a flow cell with a 2.5-mm pathlength. The peptides were separated in 0.1% TFA and eluted with a gradient of acetonitrile developed over 50 min (0–5 min, 0% B; 5–40 min, 0–50% B; 40–50 min, 50–95% B) at a flow rate of 0.85 mL/min. Peptides were detected in the effluent by measuring the absorbance at 226 nm. The peptides were repurified by reverse-phase HPLC under the same conditions as described above, except that this time 0.05% heptafluorobutyric acid was used as the ion-pairing agent. The thrombin digest was separated by reverse-phase HPLC on a Vydac C_4 column. The peptides were separated in 0.1% TFA and eluted with a gradient of 2-propanol. Elution and detection conditions were as described above.

Modification of phosphoserine with ethanethiol

Peptides containing *O*-linked phosphate were treated with ethanethiol to convert phosphoserine into *S*-ethylcysteine. The conversion was performed as a β-elimination followed by the addition of ethanethiol, essentially as described by Meyer et al. (1986). The dried peptide was incubated for 1 h at 50 °C under nitrogen with 50 μL of freshly prepared derivatization mixture (stock solution: 80 μL ethanol, 60 μL ethanethiol, 65 μL 5M NaOH, 400 μL H_2O). The sample was then cooled to room temperature and neutralized by addition of 10 μL glacial acetic acid. Derivatized samples were vacuum-dried and frozen. Prior to sequencing, the samples were dissolved with 50 μL 0.1% TFA and applied to Biobrene-treated glass fiber filters.

Amino acid sequencing

Amino acid sequence analysis was performed on an Applied Biosystems model 477A sequencer equipped with a 120A HPLC for on-line identification of the PTH derivatives. PTH-*S*-ethylcysteine was baseline separated from other PTH-amino acids and eluted just before the diphenylthiourea peak in the system used.

Mass spectrometric analysis

Peptides were subjected to mass spectrometric analysis by PDMS or MALDI-MS. PDMS was carried out using a BioIon 20K plasma desorption time-of-flight instrument (BioIon AB, Uppsala, Sweden). Samples (10–50 pmol) were dissolved in 0.1% TFA and applied to nitrocellulose-covered targets, spin-dried, and micro-rinsed as described (Nielsen & Roepstorff, 1988). MALDI-MS was performed on a BioIon prototype 1 instrument. Samples (2–5 pmol) were dissolved in 0.6 μL 0.1% TFA and mixed with 0.6 μL α-cyano-4-hydroxycinnamic acid (15 g/L) on the target. Spectra were recorded using a nitrogen ultraviolet laser at 337 nm and an acceleration voltage of 24 kV. Thirty to one hundred mass spectra were averaged and calibrated externally using human insulin. Theoretical peptide masses were calculated using the GPMAW program (Lighthouse Data, Odense, Denmark).

Acknowledgments

We thank Lise Møller for excellent technical assistance and MD Foods Research and Development Center, Brabrand, Denmark for milk samples. This work is part of the FØTEK program supported by the Danish Government and the Danish Dairy Board.

References

Ashkar S, Glimcher MJ, Saavedra RA. 1993. Mouse osteopontin expressed in *E. coli* exhibits autophosphorylating activity of tyrosine residues. *Biochem Biophys Res Commun 194*:274–279.

Ashkar S, Schaffer JL, Salih E, Gerstenfeld LC, Glimcher MJ. 1994. Phosphorylation of osteopontin by Golgi kinases. In: *Osteopontin: Role in cell signalling and adhesion*. New York Academy of Sciences, Abstract book. P1.

Barkholt V, Jensen AL. 1989. Amino acid analysis: Determination of cysteine plus half-cystine in proteins after hydrochloric acid hydrolysis with a disulfide compound as additive. *Anal Biochem 177*:318–322.

Bingham EW, Groves ML. 1979. Properties of casein kinase from lactating bovine mammary gland. *J Biol Chem 254*:4510–4515.

Boskey A, Maresca M, Ullrich W, Doty SB, Butler WT, Prince CW. 1993. Osteopontin-hydroxyapatite interactions in vitro: Inhibition of hydroxyapatite formation and growth in a gelatin-gel. *Bone Miner 22*:147–159.

Castagnola P, Bet P, Quarto R, Gennari M, Cancedda R. 1991. cDNA cloning and gene expression of chicken osteopontin. *J Biol Chem 266*:270–274.

Craig AM, Bowden GT, Chambers AF, Spearman MA, Greenberg AH, Wright JA, McLeod M, Denhardt DT. 1990. Secreted phosphoprotein mRNA is induced during multi-stage carcinogenesis in mouse skin and correlates with the metastatic potential of murine fibroblasts. *Int J Cancer 46*:133–137.

Denhardt DT, Guo X. 1993. Osteopontin: A protein with diverse functions. *FASEB J 7*:1475–1482.

Ek-Rylander B, Flores M, Wendel M, Heinegård D, Andersson G. 1994. Dephosphorylation of osteopontin and bone sialoprotein by osteoclastic tartrate-resistant acid phosphatase. *J Biol Chem 269*:14853–14856.

Fet V, Dickinson ME, Hogan BLM. 1989. Localization of the mouse gene for secreted phosphoprotein I (Spp-1) (2ar, osteopontin, bone sialoprotein 1, 44-kDa bone phosphoprotein, tumor-secreted phosphoprotein) to chromosome 5, closely linked to *Ric* (*Rickettsia resistance*). *Genomics 5*:375–377.

Fisher LW, Hawkins GR, Tuross N, Termine JD. 1987. Purification and partial characterization of small proteoglycans I and II, bone sialoproteins I and II, and osteonectin from the mineral compartment of developing human bone. *J Biol Chem 262*:9702–9708.

Franzén A, Heinegård D. 1985. Isolation and characterization of two sialoproteins present only in bone calcified matrix. *Biochem J 232*:715–724.

Giachelli CM, Bae N, Almeida M, Denhardt DT, Alpers CE, Schwartz SM. 1993. Osteopontin is elevated during neointima formation in rat arteries and is a novel component of human atherosclerotic plaques. *J Clin Invest 92*:1686–1696.

Gooley AA, Classon BJ, Marschalek R, Williams KL. 1991. Glycosylation sites identified by detection of glycosylated amino acids released from Edman degradation: The identification of Xaa-Pro-Xaa-Xaa as a motif for Thr-*O*-glycosylation. *Biochem Biophys Res Commun 178*:1194–1201.

Hu B, Coulson L, Moyer B, Price PA. 1995. Isolation and molecular cloning of a novel bone phosphoprotein related in sequence to the cystatin family of thiol protease inhibitors. *J Biol Chem 270*:431–436.

Hunter GK, Kyle CL, Goldberg HA. 1994. Modulation of crystal formation by bone phosphoproteins: Structural specificity of the osteopontin-mediated inhibition of hydroxyapatite formation. *Biochem J 300*:723–728.

Ikeda T, Shirasawa T, Esaki Y, Yoshiki S, Hirokawa K. 1993. Osteopontin mRNA is expressed by smooth muscle-derived foam cells in human atherosclerotic lesions of the aorta. *J Clin Invest 92*:2814–2820.

Kerr JM, Fisher LW, Termine JD, Young MF. 1991. The cDNA cloning and RNA distribution of bovine osteopontin. *Gene 108*:237–243.

Kiefer MC, Bauer DM, Barr PJ. 1989. The cDNA and derived amino acid sequence for human osteopontin. *Nucleic Acids Res 17*:3306.

Kohri K, Nomura S, Kitamura Y, Nagata T, Yoshioka K, Iguchi M, Yamate T, Umekawa T, Suzuki Y, Sinohara H, Kurita T. 1993. Structure and expression of the mRNA encoding urinary stone protein (osteopontin). *J Biol Chem 268*:15180–15184.

Kohri K, Suzuki Y, Yoshida K, Yamamoto K, Amasaki N, Yamate T, Umekawa T, Iguchi M, Sinohara H, Kurita T. 1992. Molecular cloning and sequencing of cDNA encoding urinary stone protein, which is identical to osteopontin. *Biochem Biophys Res Commun 184*:859–864.

Kuenzel EA, Mulligan JA, Sommercorn J, Krebs EG. 1987. Substrate specificity determinants for casein kinase II as deduced from studies with synthetic peptides. *J Biol Chem 262*:9136–9140.

Magnusson S. 1965. Preparation of highly purified bovine thrombin (E.C. 3.4.4.13) and determination of its N-terminal amino acid residues. *Ark Kemi 24*:349–358.

Mark MP, Prince CW, Gay S, Austin RL, Bhown M, Finkelman RD, Butler WT. 1987a. A comparative immunocytochemical study on the subcellular distributions of 44 kDa bone phosphoprotein and bone γ-carboxyglutamic acid (Gla)-containing protein in osteoblasts. *J Bone Miner Res 2*:337–346.

Mark MP, Prince CW, Oosawa T, Gay S, Bronkers ALJJ, Butler WT. 1987b. Immunohistochemical demonstration of a 44-kDa phosphoprotein in developing rat bones. *J Histochem Cytochem 35*:707–715.

McKee MD, Nanci A, Landis WJ, Gotoh Y, Gerstenfeld LC, Glimcher MJ. 1990. Developmental appearance and ultrastructural immunolocalization of a major 66 kDa phosphoprotein in embryonic and post-natal chicken bone. *Anat Rec 228*:77–92.

Mercier JC 1981. Phosphorylation of caseins, present evidence for an amino acid triplet code posttranslationally recognized by specific kinases. *Biochimie (Paris) 63*:1–17.

Meyer HE, Hoffmann-Posorske E, Heilmeyer MG Jr. 1991. Determination and location of phosphoserine in proteins and peptides by conversion to *S*-ethylcysteine. *Methods Enzymol 201*:169–185.

Meyer HE, Hoffmann-Posorske E, Korte H, Heilmeyer MG Jr. 1986. Sequence analysis of phosphoserine-containing peptides. Modification for picomolar sensitivity. *FEBS Lett 204*:61–66.

Nagata T, Todescan R, Goldberg HA, Zhang Q, Sodek J. 1989. Sulphatation of secreted phosphoprotein I (SSPI, osteopontin) is associated with mineralized tissue formation. *Biochem Biophys Res Commun 165*:234–240.

Nielsen PF, Roepstorff P. 1988. Optimization of sample preparation for plasma desorption mass spectrometry of peptides and proteins using a nitrocellulose matrix. *Biomed Environ Mass Spectrom 17*:355–362.

Oldberg Å, Franzén A, Heinegård D. 1986. Cloning and sequence analysis of rat bone sialoprotein (osteopontin) cDNA reveals an Arg-Gly-Asp cell-binding sequence. *Proc Natl Acad Sci USA 83*:8819–8823.

Price PA, Rice JS, Williamson MK. 1994. Conserved phosphorylation of serines in the Ser-X-Glu/Ser(P) sequences of the vitamin K-dependent matrix Gla protein from shark, lamb, rat, cow and human. *Protein Sci 3*:822–830.

Prince CW, Oosawa T, Butler WT, Tomana M, Bhown AS, Bhown M, Schrohenloher RE. 1987. Isolation, characterization, and biosynthesis of a phosphorylated glycoprotein from rat bone. *J Biol Chem 262*:2900–2907.

Roach PJ. 1991. Multisite and hierarchal protein phosphorylation. *J Biol Chem*:14139–14142.

Saavedra RA. 1994. The roles of autophosphorylation and phosphorylation in the life of osteopontin. *BioEssays 16*:913–918.

Senger DR, Perruzzi CA, Gracey CF, Papadopoulos A, Tenen DG. 1988. Secreted phosphoproteins associated with neoplastic transformation: Close homology with plasma proteins cleaved during blood coagulation. *Cancer Res 48*:5770–5774.

Senger DR, Perruzzi CA, Papadopoulos A. 1989a. Elevated expression of secreted phosphoprotein I (osteopontin, 2ar) as a consequence of neoplastic transformation. *Anticancer Res 9*:1291–1300.

Senger DR, Perruzzi CA, Papadopoulos A, Tenen DG. 1989b. Purification of a human milk protein closely similar to tumor-secreted phosphoproteins and osteopontin. *Biochim Biophys Acta 996*:43–48.

Senger DR, Perruzzi CA, Papadopoulos-Sergiou A, Van De Water L. 1994. Adhesive properties of osteopontin: Regulation by a naturally occurring thrombin-cleavage in close proximity to the GRGDS cell-binding domain. *Mol Biol Cell 5*:565–574.

Shanahan CM, Cary NRB, Metcalfe JC, Weissberg PL. 1994. High expression of genes for calcification-regulating proteins in human atherosclerotic plaques. *J Clin Invest 93*:2393–2402.

Shiraga H, Min W, VanDusen WJ, Clayman MD, Miner D, Terrell CH, Sherbotie JR, Forman JW, Przysiecki C, Neilson EG, Hoyer JR. 1992. Inhibition of calcium oxalate crystal growth in vitro by uropontin: Another member of the aspartic acid-rich protein superfamily. *Proc Natl Acad Sci USA 89*:426–430.

Smith JH, Denhardt DT. 1987. Molecular cloning of a tumor promoter-inducible mRNA found in JB6 mouse epidermal cells: Induction is stable at high, but not at low, cell densities. *J Cell Biochem 34*:13–22.

Sørensen ES, Petersen TE. 1993a. Purification and characterization of three proteins isolated from the proteose peptone fraction of bovine milk. *J Dairy Res 60*:189–197.

Sørensen ES, Petersen TE. 1993b. Phosphoproteins. *Protein Sci 2*(Suppl 1):111.

Sørensen ES, Petersen TE. 1993c. Phosphorylation, glycosylation and amino acid sequence of component PP3 from the proteose peptone fraction of bovine milk. *J Dairy Res 60*:535–542.

Sørensen ES, Petersen TE. 1994. Identification of two phosphorylation motifs in bovine osteopontin. *Biochem Biophys Res Commun 198*:200–205.

Sørensen ES, Rasmussen LK, Højrup P, Petersen TE. 1995. Localization

of in vivo phosphorylation sites in multiphosphorylated proteins: Application of *S*-ethylcysteine derivatization and mass spectrometric analysis. In: Atassi MZ, Appella E, eds. *Methods in protein structure analysis*, pp 217–225. New York: Plenum Press. Forthcoming.
Sørensen ES, Rasmussen LK, Møller L, Jensen PH, Højrup P, Petersen TE. 1994. Localization of transglutaminase-reactive glutamine residues in bovine osteopontin. *Biochem J 304*:13–16.
Tezuka K, Sato T, Kamioka H, Nijweide PJ, Tanaka K, Matsuo T, Ohta M, Kurihara N, Hakeda Y, Kumegawa M. 1992. Identification of osteopontin in isolated rat osteoclasts. *Biochem Biophys Res Commun 186*:911–917.
Wrana JL, Zhang Q, Sodek J. 1989. Full length sequence of porcine secreted phosphoprotein-I (SPP-I, osteopontin). *Nucleic Acids Res 17*:10119.
Xuan JW, Hota C, Chambers AF. 1994. Recombinant GST–human osteopontin fusion protein is functional in RGD-dependent cell adhesion. *J Cell Biochem 54*:247–255.