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

PAUL A. PRICE, JEFFREY S. RICE, AND MATTHEW K. WILLIAMSON

Department of Biology, University of California at San Diego, La Jolla, California 92093-0322

Abstract

The present studies demonstrate that matrix Gla protein (MGP), a 10-kDa vitamin K-dependent protein, is phosphorylated at 3 serine residues near its N-terminus. Phosphoserine was identified at residues 3, 6, and 9 of bovine, human, rat, and lamb MGP by N-terminal protein sequencing. All 3 modified serines are in tandemly repeated Ser-X-Glu sequences. Two of the serines phosphorylated in shark MGP, residues 2 and 5, also have glutamate residues in the $n + 2$ position in tandemly repeated Ser-X-Glu sequences, whereas the third, shark residue 3, would acquire an acidic phosphoserine in the $n + 2$ position upon phosphorylation of serine 5. The recognition motif found for MGP phosphorylation, Ser-X-Glu/Ser(P), has been seen previously in milk caseins, salivary proteins, and a number of regulatory peptides.

A review of the literature has revealed an intriguing dichotomy in the extent of serine phosphorylation among secreted proteins that are phosphorylated at Ser-X-Glu/Ser(P) sequences. Those phosphoproteins secreted into milk or saliva are fully phosphorylated at each target serine, whereas phosphoproteins secreted into the extracellular environment of cells are partially phosphorylated at target serine residues, as we show here for MGP and others have shown for regulatory peptides and the insulin-like growth factor binding protein 1. We propose that the extent of serine phosphorylation regulates the activity of proteins secreted into the extracellular environment of cells, and that partial phosphorylation can therefore be explained by the need to ensure that the phosphoprotein be poised to gain or lose activity with regulated changes in phosphorylation status.

The presence of 3 phosphoserine residues in all MGPs tested, which span shark to man, strongly supports the conclusion that phosphorylation of 3 serine residues is critical to the as yet unknown function of MGP. We have previously speculated that MGP is a locally acting regulator of cell growth and/or differentiation in the wide variety of cells that secrete the protein. Evidence for this hypothesis includes the strong induction of MGP expression by retinoic acid, a potent morphogen, and the independent discovery of MGP by differential cDNA screening as a gene that is overexpressed in transformed cells, in cells undergoing apoptosis, and in cells undergoing differentiation changes in culture. If MGP indeed acts on target cells near the point of its secretion to achieve changes in growth and/or differentiation, then regulated changes in the extent of MGP phosphorylation could provide a rapid and sensitive mechanism for controlling its activity.

In the course of these studies we have made 2 improvements in the use of ethanethiol derivatization to identify phosphoserine residues during N-terminal protein sequencing, the demonstration that this reaction can be carried out on proteins bound to a polyvinylidene difluoride membrane and the observation that maximal conversion of phosphoserine to S-ethylcysteine requires 4 h at 60 °C rather than the previously published 1 h at 50 °C.

Keywords: matrix Gla protein; N-terminal sequencing; secreted phosphoproteins; serine phosphoprotein; vitamin K dependency

Matrix Gla protein is a 10-kDa secreted protein that contains 5 residues of the vitamin K-dependent Ca^{2+} binding amino acid, γ-carboxyglutamic acid (Price et al., 1983; Price & Williamson, 1985). MGP is unique among the presently known vitamin K-dependent proteins in that it is synthesized by an exceptionally broad array of tissues and cells, whereas the vitamin K-dependent proteins involved in blood coagulation are synthesized primarily in the liver, and the bone Gla protein is synthesized in bone and dentin. MGP mRNA has been detected by Northern blot analysis in all vertebrate tissues tested to date, with the highest levels in heart, lung, kidney, and cartilage (Fraser & Price, 1988; Hale et al., 1988). MGP is also secreted by many cells in culture, including osteoblasts, chondrocytes, cardiac myocytes, vascular endothelial cells, breast cells, fibroblasts, pneumocytes, and kidney cells (Fraser & Price, 1988; Hale et al., 1988; Chen et al., 1990; Cancela & Price, 1992; Rannels et al., 1993). In spite of the broad tissue distribution of MGP synthesis, only 3 tissues have been shown to accumulate significant levels of MGP in an extracellular matrix: bone, cartilage, and calcified cartilage (Price et al., 1983; Hale et al., 1988; Rice et al., 1994).

Although the function of MGP is presently unknown, there is evidence to suggest that the protein could regulate an aspect of cell growth and differentiation. MGP gene transcription is strongly induced by retinoic acid in all normal human cell types tested (Cancela & Price, 1992), and the human MGP gene promoter contains a perfect direct repeat that is nearly identical to the retinoic acid response element in the human retinoic acid receptor β gene promoter (Cancela et al., 1990). These observations suggest that MGP could mediate some of the known actions of retinoic acid on cell growth and differentiation (Cancela & Price, 1992). Following the original isolation and sequencing of MGP from bone (Price et al., 1983; Price & Williamson, 1985), MGP has been independently discovered by differential cDNA screening as a gene that is overexpressed by breast cancer cells (Chen et al., 1990), by prostate epithelial cells undergoing apoptosis (Briehl & Miesfeld, 1991), and by vascular smooth muscle cells undergoing dedifferentiation in cell culture (Shanahan et al., 1993). Although each of these 3 independent discoveries of MGP is based on increased MGP production in cells undergoing fundamentally different transitions, the fact that MGP is one of the few genes overexpressed in each instance does suggest that MGP expression is driven by transitions in cell growth and differentiation, and that the protein may in fact play a role in these processes.

Previous studies have established the complete structure of bovine and shark MGP by protein sequencing (Price & Williamson, 1985; Rice et al., 1994), and the structures of human, rat, and mouse MGP by cDNA sequencing (Price et al., 1987; Cancela et al., 1990; Ikeda et al., 1991). The only posttranslational modifications identified in MGP to date by protein sequencing are the location of the vitamin K-dependent γ-carboxyglutamate residues in shark and bovine MGP, and the identification of a proteolytic processing event that can occur at the C-terminus of MGP and that removes the residues Arg-Arg-Gly-Ala to yield a 79-residue variant of the protein (Hale et al., 1991). There is evidence, however, that another posttranslational modification may be present in MGP because residue 5 in shark MGP could not be identified as 1 of the 20 common amino acids during protein sequencing (Rice et al., 1994). The initial objective of the present studies was to identify residue 5 in shark MGP. We have identified residue 5 as phosphoserine and have also found that 2 other nearby serines in shark MGP are partially phosphorylated. An examination of MGP isolated from 4 mammalian species has revealed that the partial phosphorylation of 3 serine residues in MGP is evolutionarily conserved.

Results

Identification of phosphoserine in MGP from calcified tissues

In order to identify the amino acid residue at position 5 in the shark MGP sequence, a residue that could not be identified as one of the 20 common amino acids during protein sequencing (Rice et al., 1994), the 23-residue CNBr peptide, which corresponds to the N-terminus of shark MGP, was subjected to mass spectroscopic analysis by Dr. Lowell Ericsson at the University of Washington. This analysis indicated that phosphoserine could be present in the peptide. To confirm the identification of phosphoserine at position 5 in the shark MGP sequence, the intact protein was treated with ethanethiol in order to convert the putative phosphoserine to S-ethylcysteine (Meyer et al., 1986, 1991) and then subjected to N-terminal sequence analysis. The HPLC separation of the PTH derivative released at the fifth sequencer cycle revealed the presence of a single peak that co-eluted with S-ethylcysteine and was present at the level predicted on the assumption that residue 5 was initially phosphoserine. No serine could be detected in the HPLC separation, indicating that phosphorylation at this position is nearly complete in the shark protein.

This sequence analysis unexpectedly revealed the presence of S-ethylcysteine at positions 2 and 3 of shark MGP. These positions had previously been identified as serine in the amino acid sequencing of shark MGP (Rice et al., 1994), and indeed the HPLC analysis of the second and third cycles of the sequential Edman degradation reaction revealed the presence of PTH serine at a level about 15% of that obtained for the PTH derivative of S-ethylcysteine. Low recovery of serine is not an unusual occurrence in the N-terminal sequencing of proteins because of the formation of PTH-dehydroalanine as a result of esterification of the serylhydroxyl group during the cleavage reaction and subsequent loss of the trifluoroacetyl group during conversion (Hunkapiller, 1988). Therefore, a lower than expected repetitive yield of serine would not, in the absence of other information, be taken as evidence that a fraction of the serine in this position is phosphorylated. The presence of phosphoserine in shark MGP was confirmed by acid hydrolysis and amino acid analysis (data not shown).

Inspection of the presently available MGP sequences revealed that 2 of the serines that are phosphorylated in shark MGP, residues 2 and 5, are conserved serines in all known MGP sequences (Rice et al., 1994). To see if these conserved serine residues are functional sites of phosphorylation, MGP was purified from bovine, lamb, and human cortical bone. After derivatization with ethanethiol, N-terminal protein sequencing revealed the presence of phosphoserine at the 2 predicted serines of these proteins, residues 3 and 6 (mammalian numbering), as well as at residue 9 (Table 1). The level of serine phosphorylation ranged from 65 to 92%, depending on the sequence position and the species (Table 1).

Table 1. *Percent phosphorylation of serine in matrix Gla protein from several species*[a]

Source of MGP	% Phosphorylation at indicated serine residue		
	Residue 3	Residue 6	Residue 9
Bovine bone			
Unfractionated	85	65	76
1–83 Form	85	69	79
1–79 Form	78	60	64
Human bone	87	83	92
Lamb bone	88	83	86
Rat kidney cell culture medium	84	82	67
	Residue 2	Residue 3	Residue 5
Shark calcified cartilage	87	84	99

[a] MGP was purified, transferred to a PVDF membrane, and derivatized with ethanethiol for 4 h at 60 °C as described in the Materials and methods. The percent phosphorylation was determined from the amount of PTH-*S*-ethylcysteine divided by PTH-*S*-ethylcysteine + PTH-serine at the indicated residues in MGP.

We have reported previously the presence of 2 forms of MGP in bovine bone, cartilage, and plasma, one of which is 79 residues in length with the C-terminus FRQ and the other 83 residues in length with the C-terminus FRQRRGA. In order to assess the possibility that the C-terminal proteolytic cleavage(s) that convert the 83 to the 79 form may be affected by the extent of MGP phosphorylation, the 2 MGP forms were separated electrophoretically and transferred to PVDF (Hale et al., 1991), derivatized with ethanethiol, and subjected to N-terminal protein sequencing. As seen in Table 1, the 79-residue form of MGP is less phosphorylated than the 83-residue form at all 3 sites of phosphorylation. The difference in the extent of phosphorylation is small, however, and may not be significant. These results indicate that the C-terminal proteolytic processing of MGP is not strongly correlated with the extent of its phosphorylation.

One possible explanation for the partial phosphorylation of serines in the SXE sequences of MGP is nonenzymatic dephosphorylation under the acidic and basic conditions used to extract and purify the protein. To test stability under the acidic conditions used to extract MGP from calcified tissues, purified bovine MGP was placed in 10% formic acid for 72 h at 25 °C. To test stability under the alkaline conditions used to purify MGP by gel filtration, purified bovine MGP was placed in 100 mM Tris, pH 9, with 6 M guanidine HCl for 48 h at 25 °C. Sequence analysis revealed essentially identical phosphorylation of serines 3, 6, and 9 in the acid, and base-treated protein as in the untreated control, and it is therefore unlikely that the conditions used in the isolation of MGP cause nonenzymatic dephosphorylation of the protein.

Identification of phosphoserine in MGP from conditioned tissue culture medium

Although MGP is a relatively abundant constituent of calcified tissues, many soft tissues that do not accumulate significant levels of the protein nevertheless contain levels of MGP mRNA that are much higher than found in calcified tissues. The highest levels of MGP mRNA are in fact found in heart, lung, kidney, and cartilage (Fraser & Price, 1988; Hale et al., 1988). To evaluate the possible phosphorylation of MGP synthesized by representative soft tissue cells, we undertook the isolation and analysis of MGP from the conditioned medium of the NRK 52E cell line, a rat cell that is known to secrete high levels of MGP into culture medium (Fraser & Price, 1988). Conditioned serum-free medium was harvested from confluent NRK cells, dialyzed against 50 mM HCl, and freeze-dried. The dry protein, which contained 160 μg of rat MGP by radioimmunoassay, was then separated from higher molecular weight conditioned medium proteins by gel filtration over a Sephacryl S-100-HR column. As seen in Figure 1, there is an absorbance peak at the elution position expected for MGP. Radioimmunoassay of the effluent fractions revealed that over 80% of the MGP antigen applied to the column was recovered in this peak, and that no other MGP antigenic component could be detected. This putative MGP component has an apparent molecular weight of approximately 14 kDa on SDS-PAGE (Fig. 1), in agreement with the 14-kDa value previously obtained for the MGP antigen in the conditioned medium of NRK cells by Western blot (Fraser & Price, 1988). N-terminal protein sequencing of this putative MGP component revealed only a single amino acid sequence, a sequence that proved to be identical to that deduced for rat MGP from the structure of the rat MGP cDNA (Price et al., 1987). After derivatization of NRK medium MGP with ethanethiol, a strong
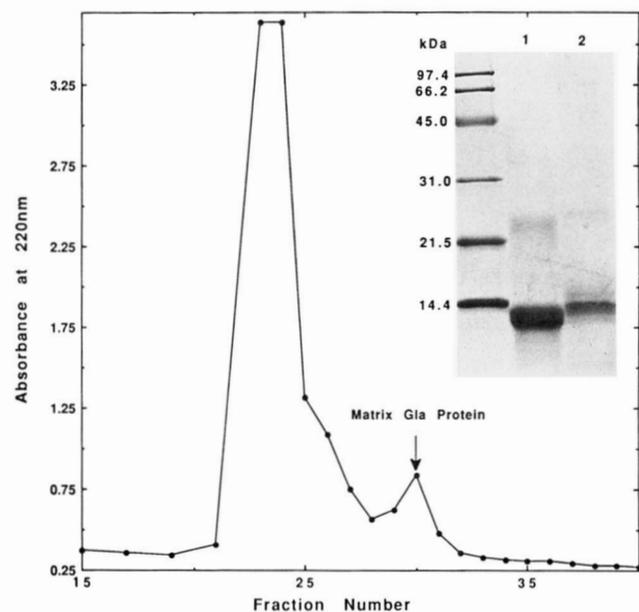
Fig. 1. Purification of MGP from the conditioned medium of NRK cells. Conditioned serum-free medium from confluent NRK cells was dialyzed against 50 mM HCl and freeze dried. The dried proteins were dissolved in 100 mM Tris, pH 9.0, 6 M guanidine HCl, and loaded onto a 0.9 × 150-cm Sephacryl S-100 HR column equilibrated in the same buffer. Fraction size, 1 mL. The arrow denotes the elution position of bovine MGP on this column. **Inset:** SDS-PAGE of purified MGP secreted by NRK cells; 10 μg of protein from fraction 30 were electrophoresed on an 18% polyacrylamide gel and stained with Coomassie blue (lane 2). Purified bovine MGP is in lane 1.

S-ethylcysteine signal was obtained at residues 3, 6, and 9 (Table 1).

Improvements in the quantitative determination of serine phosphorylation in proteins

To facilitate the transfer of MGP from the 6 M guanidine buffer used in its purification to the alkaline ethanethiol solution required for conversion of phosphoserine to *S*-ethylcysteine, we transferred the protein from the guanidine-containing buffer to a PVDF membrane using a ProSpin device. Subsequent experiments demonstrated an equivalent percentage of *S*-ethylcysteine at residues 3, 6, and 9 for ethanethiol reactions carried out with MGP in solution and with MGP adsorbed to PVDF. We therefore conclude that the hydrophobic interactions that bind MGP to the PVDF membrane do not interfere with the reaction of ethanethiol with the charged, hydrophilic phosphoserine residues in the protein.

In the initial experiments using the published ethanethiol reaction conditions, 50 °C for 1 h, we noted that the recovery of phosphoserine plus serine at residue 6 was only 15% of that which would be predicted based on the repetitive yield of the other PTH amino acids in MGP. Additional experiments were therefore undertaken to evaluate the possibility that incomplete conversion of phosphoserine to *S*-ethylcysteine may have occurred. As can be seen in Figure 2, increasing the reaction temperature to 60 °C improved the recovery of phosphoserine at residue 6 from 4 to 22%. Complete conversion of phosphoserine residue 6 to *S*-ethylcysteine however required 4 h reaction at 60 °C (Fig. 2). The incomplete conversion of phosphoserine to *S*-ethylcysteine under the published reaction conditions of 1 h at 50 °C was also observed for residues 3 and 9 in MGP (Fig. 2) and was found for MGP in solution or bound to a PVDF membrane.

It should be noted that the values for percent phosphoserine at given residues in MGP, which are shown in Table 1 and Figure 2, could be overestimations, because a reduced recovery of serine can sometimes occur in protein sequencing (Hunkapiller, 1988). We do not believe that the extent of overestimation can be large, however, because the sum of serine plus phosphoserine is 75% of that expected on the basis of repetitive yield calculations for residue 3, 55% of that expected for residue 6, and 85% of that expected for residue 9 (the data are for bovine MGP after 4 h reaction with ethanethiol at 60 °C). The minimum estimate of percent phosphoserine can be calculated if the serine recovery at a given residue is adjusted upward so as to give a serine plus phosphoserine amount which is equal to that expected on the basis of repetitive yield. For bovine MGP with a 4-h reaction with ethanethiol at 60 °C (cf. Fig. 2), the minimum percent phosphoserine would be 64% for residue 3, 36% for residue 6, and 65% for residue 9.

Discussion

Recognition sequence for MGP phosphorylation

The pattern of serine phosphorylation found in the 5 MGP sequences studied here suggests that the primary requirement for serine phosphorylation is a negatively charged glutamate or phosphoserine residue in the $n + 2$ position (Fig. 3). All 3 serines phosphorylated in the mammalian MGPs have glutamate residues in the $n + 2$ position in tandemly repeated Ser-X-Glu sequences. Two of the serines phosphorylated in shark MGP, residues 2 and 5, also have glutamate residues in the $n + 2$ position in tandemly repeated SXE sequences, whereas the third, shark residue 3, would acquire an acidic phosphoserine in the $n + 2$ position upon phosphorylation of serine 5.

Although the location of phosphoserine residues has been established in relatively few extracellular phosphoproteins, the pri-
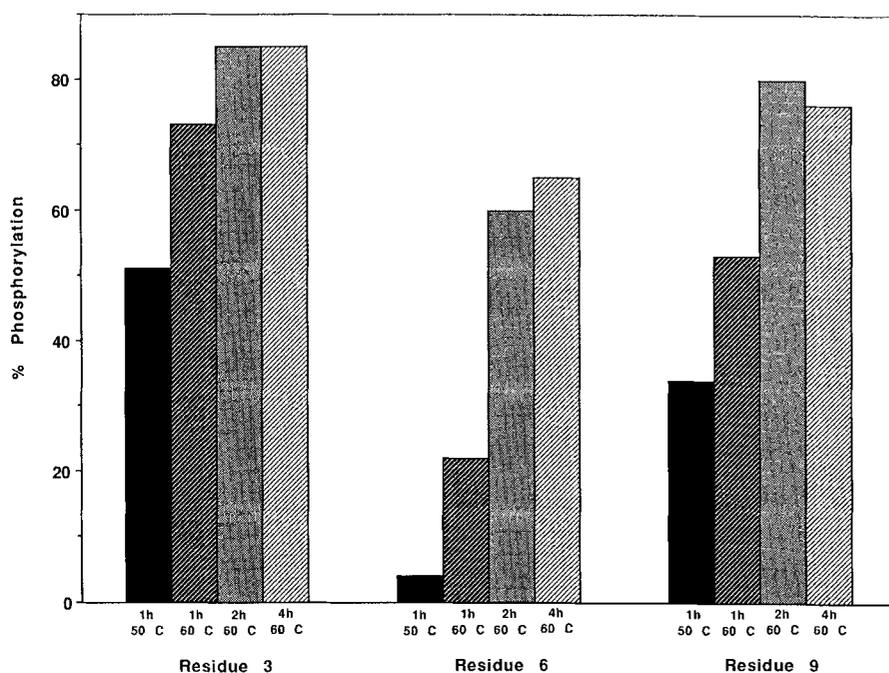
Fig. 2. Analytical determination of the percent phosphorylation of serine residues 3, 6, and 9 in bovine MGP. MGP was adsorbed to PVDF and reacted with ethanethiol reagent for 1 h at 50 °C, or for 1, 2, or 4 h at 60 °C. The PVDF membranes were then loaded into the protein sequencer and MGP was sequenced through 10 amino acids. The percent phosphorylation was determined from the amount of PTH-*S*-ethylcysteine divided by PTH-*S*-ethylcysteine + PTH-serine at the indicated residues of bovine MGP.

		1			5						10		
bMGP:	Tyr	Gla	Ser(P)	His	Glu	Ser(P)	Leu	Glu	Ser(P)	Tyr	Glu
hMGP:	Tyr	Gla	Ser(P)	His	Glu	Ser(P)	Met	Glu	Ser(P)	Tyr	Glu
lMGP:	Tyr	Gla	Ser(P)	His	Glu	Ser(P)	Leu	Glu	Ser(P)	Tyr	Glu
rMGP:	Tyr	Gla	Ser(P)	His	Glu	Ser(P)	Met	Glu	Ser(P)	Tyr	Glu
sMGP:		Asp	Ser(P)	Ser(P)	Glu	Ser(P)	Asn	Glu	Ile	Glu

Fig. 3. The Ser-S-Glu/Ser(P) recognition motif for the phosphorylation of serine in MGP. Ser-X-Glu/Ser(P) recognition motifs are boxed. Sequences (from top to bottom): bMGP, bovine MGP (Price & Williamson, 1985); hMGP, human MGP (Cancela et al., 1990); lMGP, lamb MGP; rMGP, rat MGP (Price et al., 1987); sMGP, shark MGP (Rice et al., 1994).

mary requirement for serine phosphorylation in the majority of phosphoproteins whose structure is known is also a glutamate or a phosphoserine residue in the $n + 2$ position. This pattern of serine phosphorylation was first noted in the milk caseins (Mercier, 1981), and all 24 sites of phosphorylation that have been found in the 4 bovine caseins (α_{s1}, α_{s2}, β, and κ) are in fact found at serines in SXE sequences (16 instances) or in sequences that become Ser-X-Ser(P) after the phosphorylation of SXE serines (8 instances) (Fiat & Jollès, 1989). A Ser-X-Glu/Ser(P) recognition sequence motif has also been found for another phosphoprotein isolated from milk, PP3 (Sørensen & Petersen, 1993).

A number of phosphoproteins have been isolated from saliva, each of which is phosphorylated at serines that are in SXE sequences, or in sequences that become Ser-X-Ser(P) after phosphorylation of SXE serines. Histidine-rich polypeptide and proline-rich polypeptide each contain 1 phosphoserine in an SXE sequence (Wong & Bennick, 1980; Oppenheim et al., 1986), statherin contains 2 phosphoserines in SXE sequences (Schlesinger & Hay, 1977), and cystatin S contains 1 phosphoserine in an SXE sequence and 1 in a Ser-X-Ser(P) sequence (Isemura et al., 1991). It is of interest to note that all phosphoproteins isolated from milk and saliva are fully phosphorylated at each of the target serine residues.

A Ser-X-Glu/Ser(P) recognition sequence motif has also been found for several regulatory peptides and related proteins. Adrenocorticotropin (Eipper & Mains, 1982; Bennett et al., 1983), progastrin (Dockray et al., 1987; Varro et al., 1988; Desmond et al., 1989), and interleukin 6 (May & Sehgal, 1992) have 1 phosphoserine, and in each case the phosphoserine is in an SXE sequence. Proenkephalin A has 3 phosphoserines: 2 in SXE sequences and 1 in a Ser-X-Ser(P) sequence (Watkinson et al., 1989). Finally the insulin-like growth factor binding protein 1 contains 3 phosphoserines, each of which lies in an SXE sequence (Jones et al., 1993).

In contrast to the phosphoproteins in milk and saliva, which are fully phosphorylated at each target serine residue, the phosphoproteins that are secreted into the extracellular environment of cells are partially phosphorylated in all cases where the extent of phosphorylation has been investigated. Partial phosphorylation of target serines has been noted for rat and human adrenocorticotropin (Eipper & Mains, 1982; Bennett et al., 1983); for human, porcine, and dog progastrin (Dockray et al., 1987; Varro et al., 1988; Desmond et al., 1989); for bovine proenkephalin A (Watkinson et al., 1989); and for human IGFBP-1 (Jones et al., 1993). Partial phosphorylation of 2 serines in SXE sequences has also been noted in human fibrinogen (Seydewitz et al., 1985) and was observed here in human, bovine, lamb, rat, and shark MGP. To our knowledge, among the phosphoproteins that are phosphorylated at serines in the SXE se-

quence motif and secreted into the extracellular environment of cells, only IL6 is not known to be partially phosphorylated. It should be noted, however, that the methods used to identify the site of phosphorylation in IL6 would not have revealed partial phosphorylation if it were present (May & Sehgal, 1992).

Specific phosphorylation of serines in Ser-X-Glu/Ser(P) sequences could account for the high phosphoserine content of phosphoproteins in which the sequence position of phosphoserine has not yet been established. Particularly noteworthy in this regard is osteopontin, a protein which, like MGP, accumulates in the extracellular matrix of bone. The amino acid composition of bone osteopontin indicates the presence of 12 phosphoserines (Prince et al., 1987), and the 287-residue sequence of osteopontin deduced from its cDNA structure reveals the presence of 23 serines in SXE sequences and 4 serines that become Ser-X-Ser(P) if all SXE serines are phosphorylated (Oldberg et al., 1986; Kiefer et al., 1989; Wrana et al., 1989). If there are in fact 27 serines that can be phosphorylated in osteopontin, then the presence of only 12 phosphoserines in the amino acid composition of osteopontin suggests that this protein may also be partially phosphorylated. It is also of interest that the 520-residue amino acid sequence deduced from the cDNA structure of an acidic dentine phosphoprotein, also called phosphoryn, reveals the presence of 36 serines in SXE sequences and 21 serines, which become Ser-X-Ser(P) if all SXE serines are phosphorylated (George et al., 1993).

The 3 sites of serine phosphorylation in MGP are contained within exon 2 of the human MGP gene (residues 3–12 in human MGP). Exon 2 is the only exon in the human MGP gene that could not previously be related to a functional domain of the protein (Cancela et al., 1990). Exon 1 (residues −19 to +1) contains the transmembrane signal sequence, exon 3 (residues +14 to +37) contains the γ-carboxylase recognition sequence, and exon 4 (residues +39 to +84) contains most of the γ-carboxyglutamic acid residues as well as the invariant Gla-X-X-X-Gla-X-Cys sequence common to all known, Gla-containing vertebrate proteins (Price et al., 1987). The 10 amino acid residues in exon 2 of human MGP and the corresponding sequences of bovine, rat, and lamb MGP have a conserved high density of 11 negative charges (assuming all serines are phosphorylated and fully ionized). Although this region of shark MGP has significant differences in sequence compared to the mammalian MGPs (Rice et al., 1994; Fig. 3), it is worth noting that shark MGP also has an equivalent high density of 11 negative charges in the corresponding span of 10 residues.

The SXE-specific protein kinase

A protein kinase has been isolated from the Golgi fraction of breast tissue from lactating guinea pigs (Moore et al., 1985) and

cows (MacKinlay et al., 1977), which, in studies carried out using peptide substrates, proved to be specific for serine residues in Ser-X-Glu/Ser(P) sequences (Meggio et al., 1988). This enzyme appears to be the best candidate for the protein kinase that phosphorylates the proteins that are secreted into milk, namely the caseins and PP3. The specificity of this Golgi-associated protein kinase is quite different from that found for casein kinases I and II (Meggio et al., 1988), which are cytosolic enzymes that are not known to phosphorylate casein or other secreted proteins in vivo. The Golgi-associated protein kinase of lactating breast tissue is a membrane-bound protein of 74 kDa apparent size (Moore et al., 1985).

The Golgi-associated protein kinase of lactating breast tissue could be the kinase that phosphorylates the proteins that are secreted by salivary gland cells into saliva because these proteins share with the milk proteins the property of apparently quantitative phosphorylation at each target serine in a Ser-X-Glu/Ser(P) sequence. It seems less likely, however, that phosphoproteins secreted into the extracellular environment of cells, such as the regulatory peptides and MGP, are phosphorylated by the same enzyme that phosphorylates milk and saliva proteins. Antibodies specific for the Golgi-associated protein kinase of lactating breast tissue strongly label the Golgi complex of lactating breast tissue, as expected, but fail to label the Golgi in a variety of other tissues including liver, spleen, and kidney (Moore et al., 1985). The possible existence of a different protein kinase isozyme that phosphorylates proteins secreted into the extracellular environment of cells is also supported by the observation that such secreted proteins are only partially phosphorylated at target serine residues. As discussed further below, partial phosphorylation of serine residues in extracellular proteins may provide a mechanism to rapidly alter the activity of the target phosphoproteins by changing the activity of the secretory pathway protein kinase, and this additional level of regulation could require conformational changes in this kinase that have no counterpart in the Golgi-associated protein kinase of lactating breast tissue.

The SXE-specific protein kinase, which phosphorylates proteins secreted into the extracellular environment, appears to be distributed widely in vertebrate tissues. In the present study we have found an identical pattern of Ser-X-Glu/Ser(P)-specific phosphorylation in MGP isolated from the extracellular matrix of bone and calcified cartilage and in MGP isolated from the conditioned medium of normal rat kidney cells. Bone, calcified cartilage, and kidney cells must each therefore contain an SXE-specific protein kinase. The evolutionary conservation of serine phosphorylation in MGP indicates that phosphorylation is probably critical to MGP function, and that SXE-specific phosphorylation is therefore likely to be found in MGP expressed by other tissues in addition to bone, cartilage, and kidney. Other tissues known to express MGP, and therefore likely to contain the SXE-specific protein kinase, include heart, lung, uncalcified cartilage, breast, spleen, brain, prostate, testis, and liver (Fraser & Price, 1988; Chen et al., 1990; Briehl & Miesfeld, 1991). Normal cells known to secrete MGP in culture include fibroblasts, articular cartilage chondrocytes, osteoblasts, vascular endothelial cells, breast cells, cardiac myocytes, and lung cells (Fraser & Price, 1988; Chen et al., 1990; Cancela & Price, 1992; Rannels et al., 1993), and transformed cells known to express MGP include chondrosarcomas, osteosarcomas, and breast cancer cells (Hale et al., 1988; Chen et al., 1990; Cancela & Price,

1992). The phosphorylation of regulatory peptides and fibrinogen at Ser-X-Glu/Ser(P) sequences indicates that the protein kinase must also be in the pituitary, gastric mucosa, adrenal medulla, and liver, in addition to other cells and tissues. Clearly many and possibly most vertebrate tissues must have a secretory pathway protein kinase that can phosphorylate secreted proteins at serines in Ser-X-Glu/Ser(P) sequences.

A model for the regulation of secreted protein activity by phosphorylation of serine

As noted above, in reviewing the literature on extracellular phosphoproteins, we have found an intriguing dichotomy in the extent of serine phosphorylation among secreted proteins that are phosphorylated at Ser-X-Glu/Ser(P) sequences. Phosphoproteins secreted into milk or saliva are fully phosphorylated at each target serine, whereas phosphoproteins secreted into the extracellular environment of cells are partially phosphorylated at target serine residues. It seems likely that the partial phosphorylation of serine in these purified proteins accurately reflects the in vivo extent of phosphorylation and is not caused by enzymatic dephosphorylation during purification. This conclusion is supported by the fact that widely different procedures have been used to purify these proteins, many of which are not conducive to the action of phosphatases. In addition, if dephosphorylation were a frequent problem in phosphoprotein purification, it is unclear why the milk and salivary phosphoproteins are invariably fully phosphorylated at each target serine residue. If partial serine phosphorylation indeed accurately reflects the in vivo status of these proteins, it seems likely that partial rather than complete phosphorylation provides a clue as to the function of this modification. The other possibility, that phosphoserines in these proteins have no function and so there is no selective advantage to favor complete phosphorylation, appears less likely because partial phosphorylation has been observed in so many proteins. It should also be noted that the partial phosphorylation of MGP at 3 target serine residues is sufficiently important to the function of MGP to be conserved in species as diverse as shark and man.

We propose that the extent of serine phosphorylation regulates the activity of proteins secreted into the extracellular environment of cells, and that partial phosphorylation could therefore be explained by the need to ensure that the phosphoprotein be poised to gain or lose activity with changes in phosphorylation status. There are several possible ways in which the extent of serine phosphorylation could be regulated, which include regulation of the activity of kinases or phosphatases that could act on the phosphoprotein either prior to or after secretion. In order to gain further insight into how changes in the extent of serine phosphorylation could regulate biological activity, we have derived an equation[1] that predicts the effect of changes

---

[1] In order to derive a simple relationship that will illustrate the potential of the SXE-specific protein kinase to regulate the activity of secreted phosphoproteins, we have assumed that the extent of phosphorylation of each target serine in a protein can be described by the integrated first order equation $[S]/[S]_0 = e^{-kt}$, where $[S]_0$ is the initial substrate concentration, $[S]$ is the substrate concentration remaining at time t, and k is the rate constant. The fractional extent of phosphorylation at a target serine will be given by $([S]_0 - [S])/[S]_0) = 1 - e^{-kt}$. If we assume that only the fully phosphorylated form of a protein with n target serines will be active, the percentage of maximal phosphoprotein activity will then be given by the equation $(1 - e^{-kt})^n \cdot 100\%$.

in the activity of the secretory pathway protein kinase on the activity of secreted phosphoproteins with different numbers of target serine residues. As can be seen in Figure 4, the dependence of phosphoprotein activity on the activity of the kinase predicted by this equation is hyperbolic if only 1 serine must be phosphorylated for activity and sigmoid if 2 or more sites must be phosphorylated for activity. A sigmoid dependence of secreted phosphoprotein activity on the activity of the kinase provides a larger phosphoprotein activity response for a given change in the activity of the secretory pathway protein kinase than is the case for a hyperbolic dependence.

The best example of a phosphoprotein whose activity appears to be regulated by the extent of phosphorylation may be IGFBP-1. IGFBP-1 is partially phosphorylated at 3 serine residues and the extent of its phosphorylation depends upon the tissue or cellular source of the protein. IGFBP-1, which is mostly nonphosphorylated, potentiates the effects of IGF-1 on cell growth (Elgin et al., 1987), whereas IGFBP-1, which is predominantly phosphorylated, inhibits IGF-1-stimulated cell growth (Burch et al., 1990). Phosphorylation also increases the affinity of IGFBP-1 for IGF-1 (Jones et al., 1991). If the inhibitory activity of IGFBP-1 is associated with the fully phosphorylated protein and the stimulatory activity with the nonphosphorylated form, the sigmoid dependence of activity predicted by our model for proteins with 3 phosphorylation sites indicates that the transition from one activity of IGFBP-1 to the other could be achieved with modest changes either in the activity of Ser-X-Glu/Ser(P)-specific protein kinases or in the activity of phosphatases, which act on proteins phosphorylated at these sites.

Analytical determination of phosphoserine during protein sequencing

We describe here 2 improvements in the application of the ethanethiol reaction to the identification of phosphoserine during protein sequencing. The first improvement is the demonstration that phosphoserine can be converted to S-ethylcysteine in proteins adsorbed to PVDF membranes. Adsorption of a protein to a PVDF membrane (e.g., with a ProSpin device) provides a

convenient method to remove the buffers used to purify the protein prior to its derivatization with ethanethiol. Transfer to a PVDF membrane is also a widely used method to prepare proteins for protein sequencing after electrophoretic separations. An additional advantage of the PVDF membrane procedure is that the membrane can be loaded directly into the protein sequencer, thereby avoiding possible problems in sequence analysis due to contaminants in the ethanethiol reaction mixture.

The second improvement in the ethanethiol procedure is the demonstration that the more rigorous reaction conditions of 4 h at 60 °C are needed to achieve maximal conversion of some phosphoserine residues to S-ethylcysteine. This improvement could potentially allow the identification of phosphoserine residues, such as residue 6 in bovine MGP, that do not react completely under the published conditions of 1 h at 50 °C. The quantitative conversion of phosphoserine to S-ethylcysteine also permits a more accurate estimation of the extent of serine phosphorylation at each site in a protein, as we report here for MGP.

Functional significance of MGP phosphorylation

Although the function of the vitamin K-dependent MGP is presently unknown, the strong induction of the protein by retinoic acid in all normal human cells studied has led to the suggestion that MGP may mediate an aspect of the action of retinoic acid in cell growth and differentiation by acting on target cells near the point of its secretion (Cancela & Price, 1992). A role for MGP in the regulation of cell growth and differentiation is also suggested by the observation that MGP is overexpressed in breast cancer cells (Chen et al., 1990), in prostate epithelial cells undergoing apoptosis (Briehl & Miesfeld, 1991), and in vascular smooth muscle cells undergoing dedifferentiation in cell culture (Shanahan et al., 1993). If MGP indeed regulates cellular activities, the model we have presented for the regulation of secreted protein activity by phosphorylation indicates that changes in MGP phosphorylation status could provide a rapid and sensitive MGP activity response to changes in cells or in the extracellular environment.
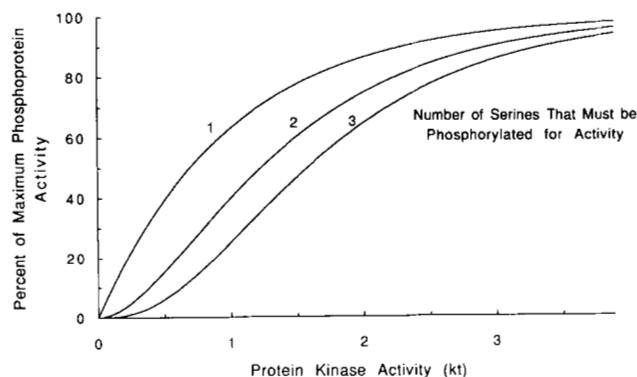
Materials and methods

Materials

NRK 52E cells were purchased from the American Type Culture Collection. Coon's F-12 medium, newborn calf serum, antibiotics, and trypsin were purchased from Irvine Scientific. Human femurs were obtained from the University of California–San Diego Body Donation Program, lamb and calf tibias were purchased from the slaughter house, and soupfin shark (*Galeorhinus galeus*) vertebrae were generously provided by Ghio Seafood Products (San Diego, California). ProSpin sample preparation cartridges and sequencer reagents were purchased from Applied Biosystems (Foster City, California). All other reagents were analytical reagent grade or better.

Fig. 4. Predicted dependence of secreted phosphoprotein activity on the activity of the SXE-specific secretory pathway protein kinase. These curves were generated from the equation: percent maximal phosphoprotein activity = $(1 - e^{-kt})^n \cdot 100\%$, where n = the number of target serine residues in the protein that must be fully phosphorylated for activity, k is the rate constant for the kinase, and t is the time during which the kinase can act on the phosphoprotein. (See footnote 1 for derivation.)

Purification of MGP

The purification of MGP from the calcified vertebrae of the soupfin shark has been described recently (Rice et al., 1994). An essentially identical procedure was used here to purify MGP from calf, lamb, and human bone. The water-insoluble fraction

of the non-collagenous proteins in the acid-demineralization extract of bone was first prepared as described (Hale et al., 1991) from midshaft region cortical bone of lamb and calf tibias and of adult human femurs. MGP was then purified from this insoluble fraction by gel filtration over a 2 × 150-cm column of Sephacryl S-200-HR equilibrated with 100 mM Tris, pH 9, with 6 M guanidine HCl (Rice et al., 1994). The MGP-containing fractions were pooled and either stored frozen or dialyzed against 50 mM HCl and dried.

MGP was also purified from 100 mL of serum-free conditioned medium from NRK 52E cells by dialysis against 50 mM HCl, freeze drying, and separation by gel filtration over a 0.9 × 150-cm column of Sephacryl S-100-HR equilibrated with 100 mM Tris, pH 9, with 6 M guanidine HCl. The MGP-containing fractions were pooled, dialyzed against 50 mM HCl, and frozen prior to adsorption onto PVDF.

Cell culture

NRK 52E cells were grown to confluency in Coon's F-12 medium containing 10% newborn calf serum as described (Fraser & Price, 1988). After 5 days culture at confluence with daily medium changes, the 100-mm culture plates were washed with phosphate-buffered saline and placed in serum-free medium as described (Johansen et al., 1992). Twenty-four hours later, the serum-free conditioned medium was removed and stored at −20 °C prior to purification of MGP. Radioimmunoassay (Otawara & Price, 1986) of this conditioned medium demonstrated that the concentration of rat MGP was 2 μg/mL.

Modification of phosphoserine with ethanethiol

Phosphoserine residues were converted to S-ethylcysteine by reaction with ethanethiol as described (Meyer et al., 1986, 1991). The dried protein was dissolved with 50 or 100 μL of derivatization reagent (stock solution: 80 μL ethanol, 65 μL 5 M NaOH, 60 μL ethanethiol, and 400 μL H₂O), flushed with N₂, capped, and the reaction was allowed to proceed for 1 h at 50 °C. The sample was then cooled to room temperature and neutralized by the addition of 0.2 mL glacial acetic acid per milliliter of derivatization reagent. Derivatized samples were loaded immediately onto a Biobrene-treated glass fiber filter for protein sequencing.

In an alternative procedure, MGP samples in the purification buffer (6 M guanidine HCl, 100 mM Tris, pH 9.0) were diluted 1:2 with H₂O and adsorbed onto PVDF membranes using a ProSpin device; MGP samples in 0.05 M HCl were adsorbed without dilution. The PVDF membranes were then rinsed with 20% methanol in order to remove salts or contaminants and stored at −70 °C until use. Phosphoserine residues in proteins adsorbed to PVDF were converted to S-ethylcysteine using the ethanethiol derivatization reagent. The PVDF membrane was placed in a 1.5-mL screw-cap polypropylene tube along with 100 μL of derivatization reagent (see above), flushed with N₂, capped, and the reaction was allowed to proceed for varying lengths of time at 50 or 60 °C (Fig. 2) or for 4 h at 60 °C (Table 1).

N-terminal protein sequencing

Automatic Edman degradations were performed on MGP samples using an Applied Biosystems 470A gas phase sequencer

equipped with a model 120 on-line HPLC and employing the normal program supplied by the manufacturer. Prior to the start of the sequencer program, the Biobrene-treated glass fiber filter or the PVDF membrane was dried for 30 min with argon gas in the assembled sequencer cartridge in order to remove ethanethiol reagent. PTH amino acid derivatives were separated using a 2.1-mm × 22-cm C-18 reverse-phase HPLC column (Applied Biosystems) and the gradient conditions recommended by Applied Biosystems. PTH-S-ethylcysteine elutes just ahead of DPTU under the conditions employed. The amount of PTH-S-ethylcysteine was determined using the calibration factor for PTH-methionine (Meyer et al., 1986).

Acid hydrolysis and amino acid analysis

Samples tested for the presence of phosphoserine were hydrolyzed for 4 h at 100 °C in 6 N HCl (Murthy & Iqbal, 1991). Hydrolyzed samples were dried and derivatized with 9-fluoroenylmethyl chloroformate using a Varian 9095 autosampler and analyzed using a Perkin-Elmer series 4 HPLC equipped with a Varian Amino Tag C₁₈ column (4.6 mm × 15 cm) and a Perkin-Elmer LC85 detector.

References

Bennett HPJ, Brubaker PL, Seger MA, Solomon S. 1983. Human phosphoserine 31 corticotropin 1–39. Isolation and characterization. *J Biol Chem* 258:8108–8112.

Briehl MM, Miesfeld RL. 1991. Isolation and characterization of transcripts induced by androgen withdrawal and apoptotic cell death in the rat ventral prostate. *Mol Endocrinol* 5:1381–1388.

Burch WW, Correa J, Shively JE, Powell DR. 1990. The 25-kilodalton insulin-like growth factor (IGF)-binding protein inhibits both basal and IGF-I-mediated growth of chick embryo pelvic cartilage in vitro. *Endocrinol Metab* 70:173–180.

Cancela L, Hsieh CL, Francke U, Price PA. 1990. Molecular structure, chromosome assignment, and promoter organization of the human matrix Gla protein gene. *J Biol Chem* 265:15040–15048.

Cancela ML, Price PA. 1992. Retinoic acid induces matrix Gla protein gene expression in human bone cells. *Endocrinology* 130:102–108.

Chen L, O'Bryan JP, Smith HS, Liu E. 1990. Overexpression of matrix Gla protein mRNA in malignant human breast cells: Isolation by differential cDNA hybridization. *Oncogene* 5:1391–1395.

Desmond H, Varro A, Young J, Gregory H, Nemeth J, Dockray GJ. 1989. The constitution and properties of phosphorylated and unphosphorylated C-terminal fragments of progastrin from dog and ferret antrum. *Regul Pept* 25:223–233.

Dockray GJ, Varro A, Desmond H, Young J, Gregory H, Gregory RA. 1987. Post-translational processing of the porcine gastrin precursor by phosphorylation of the COOH-terminal fragment. *J Biol Chem* 262:8643–8647.

Eipper BA, Mains RE. 1982. Phosphorylation of pro-adrenocorticotropin/endorphin-derived peptides. *J Biol Chem* 257:4907–4915.

Elgin RG, Busby WH, Clemmons DR. 1987. An insulin-like growth factor (IGF) binding protein enhances the biologic response to IGF-I. *Proc Natl Acad Sci USA* 84:3254–3258.

Fiat AM, Jollès P. 1989. Caseins of various origins and biologically active casein peptides and oligosaccharides: Structural and physiological aspects. *Mol Cell Biochem* 87:5–30.

Fraser JD, Price PA. 1988. Lung, heart, and kidney express high levels of mRNA for the vitamin K-dependent matrix Gla protein. Implications for the possible functions of matrix Gla protein and for the tissue distribution of the γ-carboxylase. *J Biol Chem 263*:11033-11036.

George A, Sabsay B, Simonian PAL, Veis A. 1993. Characterization of a novel dentin matrix acidic phosphoprotein. *J Biol Chem 268*:12624-12630.

Hale JE, Fraser JD, Price PA. 1988. The identification of matrix Gla protein in cartilage. *J Biol Chem 263*:5820-5824.

Hale JE, Williamson MK, Price PA. 1991. Carboxy terminal proteolytic processing of matrix Gla protein. *J Biol Chem 266*:21145-21149.

Hunkapiller MW. Gas phase sequence analysis of proteins/peptides. 1988. In: Bhown AS, ed. *Protein/peptide sequence analysis: Current methodologies*. Boca Raton, Florida: CRC Press, Inc. pp 87-117.

Ikeda T, Yamaguchi A, Icho T, Tsuchida N, Yoshiki S. 1991. cDNA and deduced amino acid sequence of mouse matrix Gla protein: One of five glutamic acid residues potentially modified to gla is not conserved in the mouse sequence. *J Bone Min Res 6*:1013-1017.

Isemura S, Saitoh E, Sanada K, Minakata K. 1991. Identification of full-sized forms of salivary (S-type) cystatins (cystatin SN, cystatin SA, cystatin S, and two phosphorylated forms of cystatin S) in human whole saliva and determination of phosphorylation sites of cystatin S. *J Biochem 110*:648-654.

Johansen JS, Williamson MK, Rice JS, Price PA. 1992. The identification of proteins secreted by human osteoblastic cells in culture. *J Bone Min Res 7*:501-512.

Jones JI, Busby WH Jr, Wright G, Smith CE, Kimack NM, Clemmons DR. 1993. Identification of the sites of phosphorylation in insulin-like growth factor binding protein-1. *J Biol Chem 268*:1125-1131.

Jones JI, D'Ercole AJ, Camacho-Hubner C, Clemmons DR. 1991. Phosphorylation of insulin-like growth factor (IGF)-binding protein 1 in cell culture and in vivo: Effects on affinity for IGF-I. *Proc Natl Acad Sci USA 88*:7481-7485.

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

MacKinlay AG, West DW, Manson W. 1977. Specific casein phosphorylation by a casein kinase from lactating bovine mammary gland. *Eur J Biochem 76*:233-243.

May LT, Sehgal PB. 1992. Phosphorylation of interleukin-6 at serine[54]: An early event in the secretory pathway in human fibroblasts. *Biochem Biophys Res Commun 185*:524-530.

Meggio F, Boulton AP, Marchiori F, Borin G, Lennon DPW, Calderan A, Pinna LA. 1988. Substrate-specificity determinants for a membrane-bound casein kinase of lactating mammary gland. *Eur J Biochem 177*:281-284.

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

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

Meyer HE, Hoffman-Posorske E, Korte H, Heilmeyer LMG Jr. 1986. Sequence analysis of phosphoserine-containing peptides. *FEBS Lett 204*:61-66.

Moore A, Boulton AP, Heid HW, Jarasch ED, Craig RK. 1985. Purifica-

tion and tissue-specific expression of casein kinase from the lactating guinea-pig mammary gland. *Eur J Biochem 132*:729-737.

Murthy LR, Iqbal K. 1991. Measurement of picomoles of phosphoamino acids by high-performance liquid chromatography. *Anal Biochem 193*:299-305.

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.

Oppenheim FG, Yang YC, Diamond RD, Hyslop D, Offner GD, Troxler RF. 1986. The primary structure and functional characterization of the neutral histidine-rich polypeptide from human parotid secretion. *J Biol Chem 261*:1177-1182.

Otawara Y, Price PA. 1986. Developmental appearance of matrix Gla protein during calcification in the rat. *J Biol Chem 261*:10828-10832.

Price PA, Fraser JD, Metz-Virca G. 1987. Molecular cloning of matrix Gla protein: Implications for substrate recognition by the vitamin K-dependent γ-carboxylase. *Proc Natl Acad Sci USA 84*:8335-8339.

Price PA, Urist MR, Otawara Y. 1983. Matrix Gla protein, a new γ-carboxyglutamic acid-containing protein which is associated with the organic matrix of bone. *Biochem Biophys Res Commun 117*:765-771.

Price PA, Williamson MK. 1985. Primary structure of bovine matrix Gla protein, a new vitamin K-dependent bone protein. *J Biol Chem 260*:14971-14975.

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.

Rannels SR, Cancela ML, Wolpert EB, Price PA. 1993. Matrix Gla protein mRNA expression in cultured type II pneumocytes. *Am J Physiol 265*:L270-L278.

Rice JS, Williamson MK, Price PA. 1994. Isolation and sequence of the vitamin K-dependent matrix Gla protein from the calcified cartilage of the soupfin shark. *J Bone Min Res 9*:567-576.

Schlesinger DH, Hay DI. 1977. Complete covalent structure of statherin, a tyrosine-rich acidic peptide which inhibits calcium phosphate precipitation from human parotid saliva. *J Biol Chem 252*:1689-1695.

Seydewitz HH, Kaiser C, Witt I. The location of a second in vivo phosphorylation site in the A α-chain of human fibrinogen. 1985. In: Henschen A, ed. *Fibrinogen — Structural variants and interactions*. Berlin: Walter de Gruyter & Co. pp 121-132.

Shanahan CM, Weissberg PL, Metcalfe JC. 1993. Isolation of gene markers of differentiated and proliferating vascular smooth muscle cells. *Circulation Res 73*:193-204.

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

Varro A, Desmond H, Pauwels S, Gregory H, Young J, Dockray GJ. 1988. The human gastrin precursor. *Biochem J 256*:951-957.

Watkinson A, Young J, Varro A, Dockray GJ. 1989. The isolation and chemical characterization of phosphorylated enkephalin-containing peptides from bovine adrenal medulla. *J Biol Chem 264*:3061-3065.

Wong RSC, Bennick A. 1980. The primary structure of a salivary calcium-binding proline-rich phosphoprotein (protein C), a possible precursor of a related salivary protein A. *J Biol Chem 255*:5943-5948.

Wrana JL, Zhang Q, Sodek J. 1989. Full length cDNA sequence of porcine secreted phosphoprotein-I (SPP-I, osteopontin). *Nucleic Acids Res 17*:10119.