THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 260, No. 28, Issue of December 5, pp. 14971-14975, 1985
Printed in U.S.A.

Primary Structure of Bovine Matrix Gla Protein, a New Vitamin K-dependent Bone Protein*

(Received for publication, June 3, 1985)

Paul A. Price and Matthew K. Williamson

From the Department of Biology, B-022, University of California at San Diego, La Jolla, California 92093

The complete amino acid sequence of bovine bone matrix Gla protein (MGP) was determined by automatic sequence analysis of the intact protein and of peptides isolated from tryptic and BNPS-skatole digests. This 79-residue, vitamin K-dependent protein contains a single disulfide bond and 4.8 γ-carboxyglutamate (Gla) residues, one each at positions 37, 41, 48, and 52, and 0.8 Gla and 0.2 Glu at position 2. There is sufficient sequence homology between MGP and bone Gla protein (BGP) to indicate that these two bovine bone proteins arose by gene duplication and subsequent divergent evolution. Although MGP has a very low solubility in water compared to BGP, there is no hydrophobic domain in MGP which could account for its insolubility, and the overall fraction of hydrophobic residues is 32% for MGP compared to 43% for BGP.

MGP is the first vitamin K-dependent protein to be discovered which has several non-γ-carboxylated residues to the NH_2-terminal side of its Gla residues. The presence of NH_2-terminal Glu residues between the putative targeting domain for the γ-carboxylase in the MGP leader sequence and the mid-molecule Gla residues suggests that the γ-carboxylase may have additional, as yet unrecognized, specificity requirements which determine the susceptibility of Glu residues for γ-carboxylation.

Matrix Gla protein (MGP[1]) is the second vitamin K-dependent protein to be discovered in bone (1). It was initially isolated from urea extracts of demineralized bovine bone by virtue of its reversible precipitation upon removal of denaturant. Bone morphogenetic protein, the objective of the urea extraction procedure (2), co-precipitates with MGP. This co-precipitation has led to the suggestion that bone morphogenetic protein may be normally associated with MGP *in vivo* (1, 2).

MGP is physically and chemically distinct from the better characterized bone Gla protein (BGP). While both proteins contain one disulfide bond, MGP has 5 residues of the vitamin K-dependent Ca^{2+} binding amino acid γ-carboxyglutamic acid (Gla) compared to 3 for BGP (1, 3). The apparent molecular weight of MGP on sodium dodecyl sulfate-urea gel electrophoresis is also larger, M_r = 15,000 compared to M_r = 11,000 for BGP (1, 4). Finally bovine MGP contains no hydroxyproline, an amino acid found at position 9 in bovine BGP (1, 3), and antibodies raised against the respective purified bovine proteins do not cross-react (1).

The times at which significant levels of MGP and BGP are first detected in calcifying bone extracts are strikingly different, as is the relative level of the two proteins in tissues which have a calcified collagenous matrix. While MGP is found at comparable levels in bone and dentine (each 0.4 mg/g), BGP levels are high in bone (2 mg/g) and relatively low in dentine (0.4 mg/g) (5). In the rapidly calcifying bone of newborn rats, MGP levels are about 130% of those in adult rats[2] while BGP levels are only 1–2% of those in the adult animal (6). Thus MGP is deposited in bone at an earlier stage of bone development than BGP and is present in calcified collagenous tissues at a more constant weight fraction than is BGP. Together with the evidence for its association with bone morphogenetic protein, these observations suggest that MGP may play a critical role in an early stage of calcified tissue formation.

The primary structure of bovine BGP was established 9 years ago (3) and still remains the only primary structure available for a vitamin K-dependent vertebrate protein which is not involved in blood coagulation. The primary structure of BGP proved to be markedly different from the primary structures of the closely related, Gla-containing proteins involved in coagulation, a result which indicates that BGP arose independently from the coagulation factors during evolution. In order to establish the relationship between MGP and these other classes of vitamin K-dependent vertebrate proteins, we have determined the complete primary structure of this protein.

MATERIALS AND METHODS AND RESULTS[3]

The complete sequence of matrix Gla protein, shown in Fig. 1, was deduced from automatic sequence analysis of the intact protein and of peptides produced from the intact protein by tryptic digestion and by BNPS-skatole cleavage. The protein has 79 amino acid residues in a single polypeptide chain and

* This work was supported in part by United States Public Health Service Grant AM 27029. The protein sequencing facility used in these studies was obtained with the aid of a National Institutes of Health Shared Instrumentation Grant RR01483. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] The abbreviations and trivial names are: MGP, matrix Gla protein; BGP, bone Gla protein, vitamin K-dependent bone protein, osteocalcin; BNPS-skatole, 2-(2-nitrophenylsulfenyl)-3-methyl-3′-bromoindolenine; PTH, phenylthiohydantoin; HPLC, high-performance liquid chromatography; TPCK, tosylphenylalanyl chloromethyl ketone; TFA, trifluoroacetic acid.

[2] Y. Otawara and P. A. Price, manuscript in preparation.

[3] Portions of this paper (including "Materials and Methods," part of "Results," Tables II and III, and Figs. A–C) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are available from the Journal of Biological Chemistry, 9650 Rockville Pike, Bethesda, MD 20814. Request Document No. 85M-1807, cite the authors, and include a check or money order for $4.80 per set of photocopies. Full size photocopies are also included in the microfilm edition of the Journal that is available from Waverly Press.

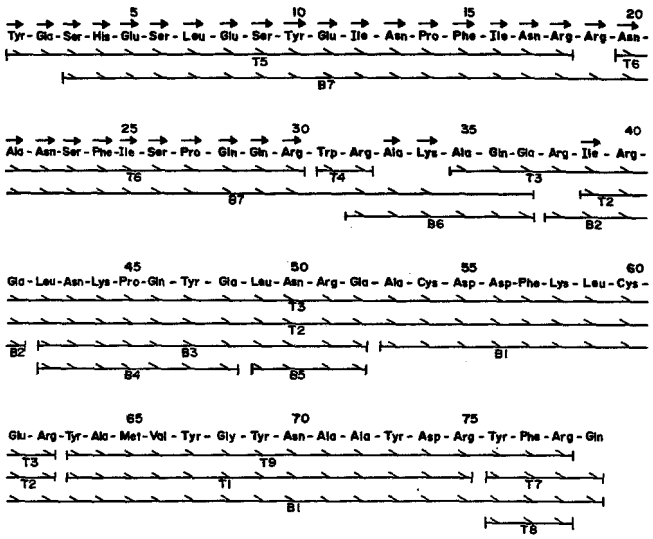

FIG. 1. **Amino acid sequence of bovine matrix Gla protein.** Tryptic peptides are labeled T1 to T9 and peptides generated by BNPS-skatole cleavage are labeled B1 to B7. Residues which were identified by automatic sequence analysis are marked with →.

has a calculated molecular weight of 9961. The single disulfide bond previously identified in MGP (1) joins half-cystine residues at sequence positions 54 and 60.

The assignment of Gla residues to sequence positions 37, 41, 48, and 52 in MGP was established in two ways. From the composition of tryptic peptide T2 it was evident that the peptide contained 3 Gla residues. In two sequenator runs on T2 and in one sequenator run on T3, no PTH derivative of glutamic acid, or of any other amino acid, was detected for positions 37, 41, 48, and 52. This is consistent with the fact that the PTH derivative of Gla is not recovered in automatic sequenator analysis (7). Since the expected yield was obtained for PTH Glu at position 61 and for PTH Gln at positions 36 and 46, the only positions for the 4 Gla residues consistent with the sequenator data are 37, 41, 48, and 52. The second way Gla positions were identified was by BNPS-skatole cleavage. It has been established that reagents which generate electrophilic bromine, such as cyanogen bromide and BNPS-skatole, can produce cleavages at the COOH-terminal side of Gla residues (8, 9). As can be seen in Fig. 1, each putative Gla residue indeed proved to be a BNPS-skatole cleavage site. In addition, the acid hydrolysates of each peptide contained a single residue of γ-hydroxyglutamic acid, the expected product of the cleavage reaction (8). The one anomalous BNPS-skatole peptide, B3, must have arisen from incomplete cleavage at the Gla-48 rather than from the presence of non-γ-carboxylated glutamic acid at this position, since no PTH Glu could be detected at this step in the B3 sequenator run.

The assignment of Gla to position 2 in the sequence of MGP was also supported by the recovery of the expected peptide, B7, in the BNPS-skatole cleavage digest (Fig. 1). Automatic sequenator analysis of the intact protein and of tryptic peptide T5, however, revealed some PTH glutamic acid at this site. The recovery of PTH glutamic acid at position 2 was 13–23% of that predicted from the extrapolated repetitive yield of PTH Glu at positions 5, 8, and 11, which indicates that position 2 has a mixture of approximately 20% Glu and 80% Gla. The amino acid composition of MGP deduced from sequence analysis is in good agreement with the composition determined from acid and alkaline hydrolysates of the purified protein (Table I).

DISCUSSION

The bovine MGP sequence is strikingly homologous to the previously determined sequences of bovine and swordfish BGP (Fig. 2). This sequence homology indicates that MGP and BGP must have arisen by gene duplication and subsequent divergent evolution. Since the sequence homology is greater between bovine MGP and swordfish BGP (14 identical residues) than between the two bovine proteins (10 identical residues), it is likely that the gene duplication event preceded the divergence of swordfish from other vertebrates.

Among the conserved structures in MGP and BGP are 2 Gla residues and the nearby disulfide bond. In BGP, specific chemical modifications of the Gla residues and the disulfide bond have shown that these groups are required for several *in vitro* interactions with hydroxyapatite (10). *In vivo,* the

TABLE I

Amino acid composition of the calf bone matrix Gla protein

Amino acid	Calculated[a]	Found[b]
γ-Carboxyglutamic acid	3.5	4.8
Aspartic acid	9.3	10
Threonine	0.3	0
Serine	4.6	5
Glutamic acid	8.3	9.2
Proline	3.2	3
Glycine	2.3	1
Alanine	6.9	7
Half-cystine[c]	1.6	2
Valine	0.5	1
Methionine	1.0	1
Isoleucine	3.8	4
Leucine	3.5	4
Tyrosine	6.8	8
Phenylalanine	3.6	4
Histidine	0.8	1
Lysine	3.0	3
Tryptophan	ND[d]	1
Arginine	10.2	10
Total residues		79

[a] Composition based on amino acid analysis of the purified protein and a 9961-dalton molecular size.

[b] Values represent data obtained from sequence analysis.

[c] Determined as carboxymethylcysteine after reduction and carboxymethylation.

[d] ND, not determined.

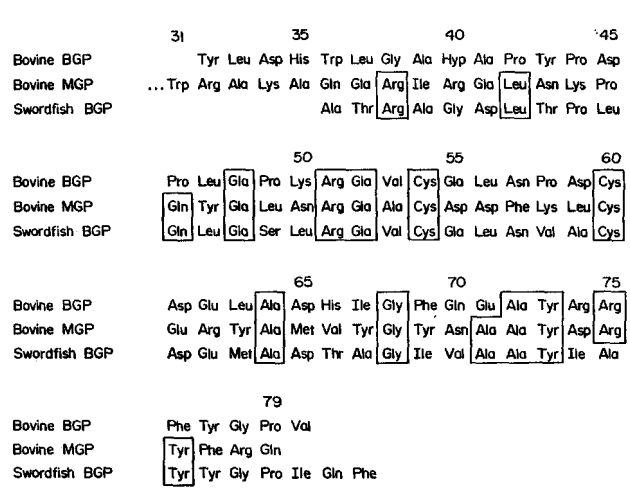

FIG. 2. **Sequence homology between bovine and swordfish bone Gla proteins and bovine matrix Gla protein.**

abnormal non-γ-carboxylated BGP synthesized by animals treated with the vitamin K antagonist warfarin can neither bind to hydroxyapatite strongly (11) nor accumulate in bone (12). It is tempting to speculate that the homologous Gla residues and disulfide bond elements of MGP likewise form a part of the MGP structure which binds to hydroxyapatite, but this remains to be studied.

There are two interesting internal homologies within the Gla-containing region of MGP which could be important to its possible interaction with hydroxyapatite. One is the occurrence of the 4 Gla residues in this region in two homologous pentapeptide domains, Gla-Arg-Ile-Arg-Gla (positions 37–41) and Gla-Leu-Asn-Arg-Gla (positions 48–52). The Chou-Fasman (23) and Garnier (24) algorithms do not consistently predict α-helical structure for these pentapeptide domains. However, if these domains proved to be α-helical, the spacing of Gla residues within them would cause the Gla side chains to lie on the same side of the helix. The other intriguing homology in the Gla-containing region of MGP is the recurring tripeptide Gla-Leu-Asn, which is found at sequence positions 41–43 and 48–50. While it is unclear what significance such internal structural motifs in MGP may have in hydroxyapatite binding, it is worth noting that the second pentapeptide Gla domain in MGP has a counterpart in the Gla-*X*-*X*-Arg-Gla domain of BGP (positions 48–52 in MGP numbering, Fig. 2) and that the Gla-Leu-Asn tripeptide is found at positions 55–57 in both swordfish and bovine BGP (MGP numbering, Fig. 2).

An interesting aspect of the two homologous pentapeptide domains in MGP is the presence of adjacent Arg residues within the Gla domains. Chemical modification of arginine residues inactivates numerous enzymes which act on phosphate, phosphomonoesters, and phosphodiesters, an observation which has led to the suggestion that the side chain of arginine may provide a phosphate anion binding site in such proteins (13). It is worth noting in this context that the guanidino group of arginine can form a bidentate hydrogen-bonded complex with phosphate (Fig. 3) while the ε-amino group of lysine cannot. Such bidentate association between arginine and surface phosphate groups in hydroxyapatite could complement the previously postulated (10–12) bidentate association between the malonate side chain of Gla and surface Ca^{2+} atoms on bone mineral. This additional interaction would increase binding strength and could promote selectivity for given hydroxyapatite crystal faces.

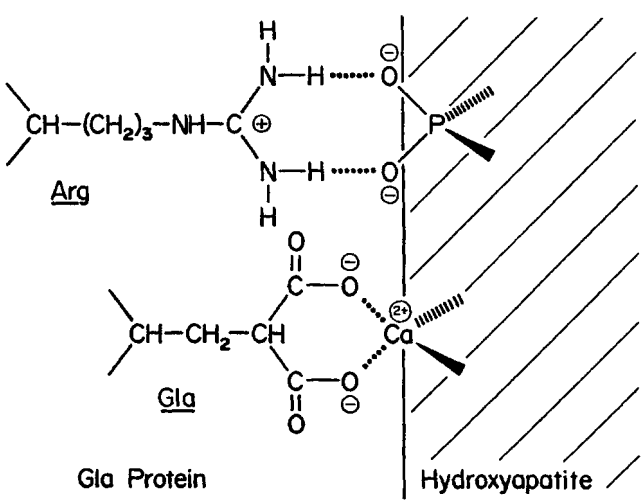

FIG. 3. **Schematic representation for the postulated bidentate complexes between the side chains of arginine and γ-carboxyglutamate in MGP and surface phosphate and calcium ions in hydroxyapatite.** Note that linear hydrogen bonds can be found between both oxygen atoms of a surface phosphate in hydroxyapatite and adjacent N-H hydrogens of the guanidino group of arginine.

The amino acid sequence of MGP provides no explanation for why the protein is insoluble in the absence of denaturants (1). Although only 32% of the amino acids in MGP are hydrophobic, the solubility limit of the protein at neutral pH is less than 0.2 mg/ml. In contrast, BGP has 43% hydrophobic amino acids, yet its water solubility is over 100 mg/ml. There is also no cluster of hydrophobic amino acids within the MGP sequence which could provide a rationale for its water insolubility. Since nonspecific hydrophobic associations do not appear to account for the extreme water insolubility of the protein, one must consider the possibility that insolubility reflects specific aggregation of MGP monomers to form a partially ordered complex *in vitro* which mirrors a possible structure in bone. In this regard, it is interesting to note that MGP was initially isolated by virtue of its co-precipitation with the otherwise water-soluble bone morphogenetic protein upon removal of denaturants (1, 2). The association between the MGP aggregate and the bone morphogenetic protein has been suggested as the mechanism by which the bone morphogenetic protein is attached to the organic matrix of bone (1, 2).

MGP is the first example of a vitamin K-dependent protein which has several glutamic acid residues to the NH_2-terminal side of most of its γ-carboxyglutamic acid residues. In BGP, as well as in such serum proteins as prothrombin, factors VII, IX, and X and proteins C, S, and Z, all glutamic acid residues in the NH_2-terminal first 30–40 residues are γ-carboxylated (3, 14). In contrast, the glutamic acid residues at positions 5, 8, and 11 of MGP are not γ-carboxylated while those at 2, 37, 41, 48, and 52 are. It has been suggested that vitamin K-dependent proteins may contain within their leader sequence a region which targets them for the γ-carboxylase reaction (15). This hypothesis has recently been supported by the discovery that there is in fact a segment of the leader sequence of BGP which is highly homologous to the leader sequences of prothrombin, factors IX and X, and protein CB (25).

There are two possible mechanisms by which the γ-carboxylase could, after first binding to the target portion of the leader sequence, subsequently γ-carboxylate glutamic acid residues in the NH_2-terminal region of a protein. One mechanism would be for the γ-carboxylase to travel down the polypeptide, γ-carboxylating each glutamic acid residue it encountered until it reaches some detachment signal. This mechanism is clearly incompatible with the presence of 3 non-γ-carboxylated glutamic acid residues so close to the NH_2-terminal end of MGP. The other mechanism would be for the γ-carboxylase to have two sites, one site which binds the enzyme to the target domain in the leader sequence, and the other the enzymatic active site which γ-carboxylates all accessible glutamic acid residues in the vitamin K-dependent protein. This latter mechanism provides a simple explanation for the failure of the γ-carboxylase complex to γ-carboxylate glutamic acid residues at the NH_2-terminus of MGP, since secondary and tertiary structure of a given sequence segment could easily limit access of the γ-carboxylase active site to a subset of the NH_2-terminal glutamic acid residues even though the γ-carboxylase complex is anchored to a nearby segment of the leader sequence.

The present investigation represents the first extensive use of BNPS-skatole cleavage at γ-carboxyglutamic acid in the

sequencing of a vitamin K-dependent protein. The conditions used for cleavage at γ-carboxyglutamic acid are identical to those employed for cleavage at tryptophan (20). An advantage of cleavage at Gla residues in protein sequencing is that the product of the cleavage reaction yields, after acid hydrolysis, γ-hydroxyglutamic acid (8). Since this modified amino acid can be quantified during amino acid analysis, one can determine which peptides have arisen as a result of cleavage at Gla and thereby locate Gla residues within the structure of the protein.

Acknowledgment—We would like to thank Dr. Russell Doolittle for providing the secondary structure predictions for MGP.

REFERENCES

1. Price, P. A., Urist, M. R., and Otawara, Y. (1983) *Biochem. Biophys. Res. Commun.* **117,** 765–771
2. Urist, M. R., Huo, Y. K., Brownell, A. H., Hohl, W. M., Buyske, J., Lietze, A., Tempst, P., Hunkapiller, M., and DeLange, R. J. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81,** 371–375
3. Price, P. A., Poser, J. W., and Raman, N. (1976) *Proc. Natl. Acad. Sci. U. S. A.* **73,** 3374–3375
4. Price, P. A., Otsuka, A. S., Poser, J. W., Kristaponis, J., and Raman, N. (1976) *Proc. Natl. Acad. Sci. U. S. A.* **73,** 1447–1451
5. Price, P. A., Williamson, M. K., and Otawara, Y. (1985) in *The Chemistry and Biology of Mineralized Tissues* (Butler, W. T., ed) pp. 159–163, EBSCO Media, Birmingham
6. Price, P. A., Lothringer, J. W., and Nishimoto, S. K. (1980) *J. Biol. Chem.* **255,** 2938–2942
7. Fernlund, P., and Stenflo, J. (1979) in *Vitamin K Metabolism and Vitamin K-dependent Proteins* (Suttie, J. W., ed) pp. 161–165, University Park Press, Baltimore
8. Katayama, K., and Titani, K. (1978) *FEBS Lett.* **95,** 157–160
9. Enfield, D. L., Ericsson, L. H., Fujikawa, K., Walsh, K. A., Neurath, H., and Titani, K. (1980) *Biochemistry* **19,** 659–667
10. Poser, J. W., and Price, P. A. (1979) *J. Biol. Chem.* **254,** 431–436
11. Price, P. A., Williamson, M. K., and Lothringer, J. W. (1981) *J. Biol. Chem.* **256,** 12760–12766
12. Price, P. A., and Williamson, M. K. (1981) *J. Biol. Chem.* **256,** 12754–12759
13. Riordan, T. F. (1979) *Mol. Cell. Biochem.* **26,** 71–92
14. Hojrup, P., Roepstorff, P., and Petersen, T. E. (1982) *Eur. J. Biochem.* **126,** 343–348
15. Nishimoto, S. K., and Price, P. A. (1980) *J. Biol. Chem.* **255,** 6579–6583
16. Price, P. A. (1983) *Methods Enzymol.* **91,** 13–17
17. Hugli, T. E., Vallota, E. H., and Müller-Eberhard, H. J. (1975) *J. Biol. Chem.* **250,** 1472–1478
18. Martin, B., Svendsen, I., and Ottesen, M. (1977) *Carlsberg Res. Commun.* **42,** 99–102
19. Ambler, R. P. (1967) *Methods Enzymol.* **11,** 155–166
20. Fontana, A. (1972) *Methods Enzymol.* **25,** 419–423
21. Hewick, R. M., Hunkapiller, M. W., Hood, L. E., and Dreyer, W. J. (1981) *J. Biol. Chem.* **256,** 7990–7997
22. Hunkapiller, M. W., and Hood, L. E. (1983) *Methods Enzymol.* **91,** 486–493
23. Chou, P. Y., and Fasman, G. D. (1978) *Annu. Rev. Biochem.* **47,** 251–276
24. Garnier, J., Osguthorpe, D. J., and Robson, B. (1978) *J. Mol. Biol.* **120,** 97–120
25. Pan, L. C,. and Price, P. A. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82,** 6109–6113

SUPPLEMENTAL MATERIAL
TO
Primary Structure of Bovine Matrix Gla Protein,
a New Vitamin K-dependent Bone Protein

Paul A. Price and Matthew K. Williamson

MATERIALS AND METHODS

Materials

Calf bone was obtained from Vista Meat Packers, Vista, California. Distilled in glass chemicals for HPLC were obtained from Burdick and Jackson or J.T. Baker. Sequencer reagents were purchased from Applied Biosystems. TPCK-Trypsin was from Worthington Chemicals, BNPS-skatole (2-(2-nitrophenylsulfenyl)-3-methyl-3'-bromoindolenine) was from Pierce, Carboxypeptidase-Y was from CalBiochem/Behring, and Dowex AG50 was from BioRad. All other reagents employed were analytical reagent grade or better. MGP was purified from $CaCl_2$ urea extracts of demineralized, gelatinized calf cortical bone as previously described (1,2).

Methods

Amino acid analysis: Protein or peptide samples were hydrolyzed for 24h at 110°C in 6 M HCl in evacuated, sealed tubes. Analysis for γ-carboxyglutamic acid was accomplished following alkaline hydrolysis in 2.5 M KOH at 110°C for 24h (16). Amino acid analyses were performed on a Beckman 119C amino acid analyzer equipped with a Spectra Physics 4000 data reduction system.

Reduction and Alkylation: Protein was reduced and S-carboxymethylated using the procedure of Hugli, et al (17).

Tryptic digestion: In a typical digestion, 5.6 A280 units of S-carboxymethylated MGP was dissolved in 0.8 ml of 1% NH_4HCO_3, 8M urea. The sample was diluted 1:5 with 1% NH_4HCO_3 and the pH was spot checked with pH indicator paper. Sufficient TPCK-trypsin (a 1% solution in 1% NH_4HCO_3) was added to achieve an enzyme to substrate ratio (by weight) of 1:100 and digestion was carried out at 25°C with magnetic stirring. After 2h an additional aliquot of trypsin, equal to the first, was added and the digestion proceeded an additional 5 h at which time the reaction was terminated by freeze-drying.

Carboxypeptidase Y digestion: Carboxypeptidase Y digestion of 200 nmol of protein was performed as described by Martin, et al. (18) in 0.5% sodium dodecyl sulfate (w/w), 0.1M-2[N-morpholino]ethane sulfonic acid, pH 6.3 for 10 h at 25°C. The ratio of enzyme to substrate used was 1:500 on a molar basis and the reaction was terminated by the addition of Dowex AG 50W-X4 (H^+ form, 200-400 mesh) according to the procedure of Ambler (19).

Chemical Digestion by BNPS-skatole: Digestion of MGP with BNPS-skatole was performed as described by Fontana (20). In a typical digestion 3.6 A280 units of MGP was dissolved with 400 µl of 60% acetic acid. 66 µg of tyrosine, in 60% acetic acid, was added and digestion was initiated by the addition of 8 mg of BNPS-skatole. The vial was wrapped with aluminum foil and stirred at room temperature. After 7.5 h, a second aliquot of tyrosine, equal to the first, was added. The reaction was then allowed to proceed overnight and was terminated the next morning by extracting the BNPS-skatole reagent with diethyl ether. The aqueous phase, which contained the generated peptides, was then freeze-dried.

Peptide isolation: Peptides of MGP were purified using a Waters Associates HPLC system consisting of a WISP autosampler, two model 6000A solvent delivery systems, a 660 solvent programmer, and a 450 variable wavelength detector. Peptides were detected by their absorbance at 220 or 230 nm and peptides were purified using either a Waters semipreparative C-18 (7.8 mm x 30 cm), or an Altex analytical C-3 (4.6 mm x 7.5 cm) reverse phase HPLC column employing 0.1% TFA to 0.1% TFA in CH_3CN:H_2O (60:40) gradients.

Edman Degradation: Automatic Edman degradations were performed on the intact, S-carboxymethylated protein and on purified peptides using an Applied Biosystems Model 470A Gas Phase Sequencer employing the standard "NoVac" program as supplied by the manufacturer (21). In all cases, the end of a peptide was established by the absence of any PTH derivative in 3 or more sequential cycles. PTH-amino acid derivatives were identified by HPLC on a Perkin Elmer HS-3 C-18 column or methylated with 1M methanolic HCL and identified on an IBM Cyano column (22). The HPLC system employed consisted of a Perkin Elmer Series 4 liquid chromatograph, a LC-85B spectrophotometric detector equipped with a 1.4 µl flow cell, an ISS-100 automatic sample injector, a LCI 100 computing integrator, and a model 7500 computer employing Chrom III software.

RESULTS

Sequence Analysis of the Whole Protein

Direct automatic sequence analysis on 10 nmoles of the calf matrix Gla protein provided sequence information on 33 of the first 39 residues of the molecule. The residues which could not be identified were Trp at position 31, Arg at position 32, Ala at position 35, Gln at position 36, Gla at position 37, and Arg at position 38. Nevertheless, the data allow for alignment of several peptides and unambiguously place 3 Glu residues at the N-terminal portion of the molecule.

Isolation and Characterization of Peptides derived from MGP

Tryptic digest: The tryptic digest of S-carboxymethylated MGP was redissolved in 4 ml of 0.1% trifluoroacetic acid and purified by HPLC as shown in Figure A using the analytical C3 reverse phase column. Aliquots from each peak were subjected to amino acid analysis as well as automatic sequence analysis. The peaks which eluted between 111 and 120 minutes were found to be impure and were further purified using a shallower gradient and the same analytical C3 column. This purification resulted in 4 peaks which eluted between 30 and 55 minutes after injection and are identified as T2, T3, T5', and T5 in Figure B. T5' differs from T5 in the presence of 2 Arg residues rather than 1 in its amino acid composition, and, therefore, corresponds to residues 1-19 of the intact MGP molecule (Figure 1). The amino acid compositions for tryptic peptides T1-T8 are given in Table II. The composition of T9 was not obtained because there was insufficient material present in the purifications already on hand, and because its sequence only supported data obtained from the sequence analyses of other tryptic peptides (see Figure 1).

BNPS - skatole digest: The BNPS-skatole digest of S-carboxymethylated MGP was redissolved in 1 ml of 0.1% trifluoroacetic acid and purified using a semi-preparative C-18 reverse phase HPLC column as shown in Figure C. Five peptides were purified to homogeneity in this first step: B2, B3, B4, B5, and B7. The peak which eluted in fractions 48 through 52 was found to be impure and was subsequently fractionated into B7, B6, and B1 by rechromatography on the same C18 column using a shallower gradient (data not shown). Mobile phases were the same as described in Figure C, and the linear gradient was from 55 to 70% B in 2 h. The amino acid compositions supporting the sequences of the BNPS-skatole peptides are given in table III. It may be noted that B2 and B5 have anomalously high levels of tyrosine. This is attributable to the addition of a large excess of tyrosine to act as a scavenger in order to prevent the modification of tyrosine in the intact molecule during treatment with BNPS-skatole.

CPY digestion: Carboxypeptidase Y digestion confirmed that the C-terminal of the molecule was Tyr-Phe-Arg-Gln. After 10 h of digestion at 25°C, the molar ratio of Tyr:Phe:Arg:Gln was 1.0:1.2:1.2:1.0.

Complete Sequence of Bovine Matrix Gla Protein: The complete sequence of the bovine matrix Gla protein, as deduced from the above studies on the whole protein and peptides, is presented in Figure 1. The criteria employed for the identification of Gla at positions 2, 37, 41, 48, and 52 are described under Results. The presence of some glutamic acid at position 2 was established from the presence of PTH Glu at the second sequencer step in the automatic sequencer analysis of the intact protein and of peptide T5. The recovery of PTH Gln at position 2 was 13% (peptide T5) to 23% (intact protein) of that predicted from the extrapolated repetitive yield of PTH Glu at positions 5, 8, and 11.

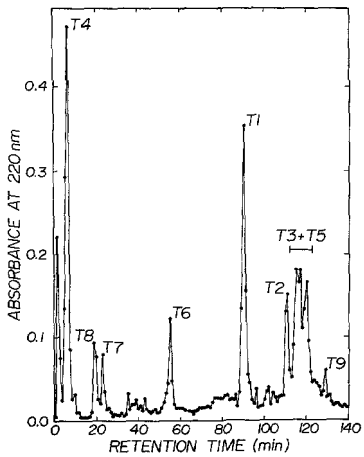

Figure A. High performance liquid chromatography of tryptic digest of S-carboxymethylated MGP on an Altex ultrapore C3 column (4.6 mm x 7.5 cm). Mobile phase, Buffer A - 0.1% TFA; Buffer B -0.1% TFA in CH_3CN:H_2O (60:40, v/v). Flow rate, 1.0 ml/min. Temperature, 25°C. The digested MGP was dissolved in 4 ml of 0.1% TFA and 0.4 ml was applied to the column equilibrated in 100% Buffer A. At injection, a linear gradient from 0-35% B was begun and developed over 2 h.

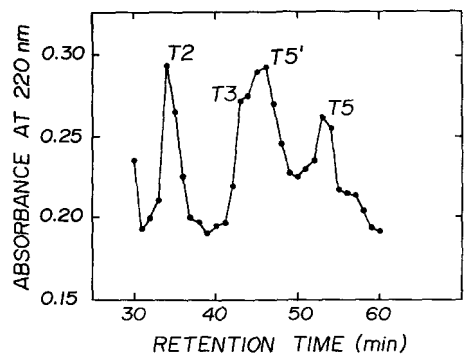

Figure B. High performance liquid chromatography of tryptic digest of S-carboxymethylated MGP on an Altex ultrapore C3 column (4.6 mm x 7.5 cm). Mobile phases, flow rate, and temperature were the same as described in Figure A. 0.4 ml of the digest was applied to the column equilibrated in 100% Buffer A. At injection, a linear gradient from 0-25% B was run over 15 min, followed by a linear gradient from 25-35% B in 2 h.

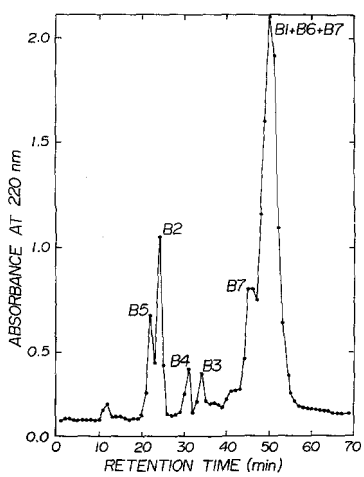

Figure C. High performance liquid chromatography of BNPS-skatole digest of S-carboxymethylated MGP on a Waters semiprep C-18 column (7.8 mm x 30 cm). Mobile phase, Buffer A - 0.1% TFA; Buffer B - 0.1% TFA in CH_3CN:H_2O (60:40, v/v). Flow rate, 1.0 ml/min. Temperature, 25°C. The digested MGP was dissolved in 1 ml of 0.1% TFA and 0.2 ml applied to the column equilibrated in 100% Buffer A. At injection, a linear gradient from 0-100% B was begun and developed over 1 h.

TABLE II

Amino Acid Composition of

Tryptic Peptides from Matrix Gla Protein[a]

Amino Acid	T1	T2	T3	T4	T5	T6	T7	T8
γ-Carboxyglutamic acid[b]		2.8(3)	N.D.(4)		0.2(1)			
S-carboxymethyl cysteine		1.4(2)	1.6(2)					
Aspartic acid	2.1(2)	4.0(4)	5.3(4)		2.0(2)	2.0(2)		
Threonine								
Serine			0.8		2.5(3)	1.6(2)		
Glutamic acid		2.2(2)	7.0[c](3)		3.6(3)	2.1(2)	1.0(1)	
Proline		1.0(1)	N.D.(1)		0.8(1)	1.0(1)		
Glycine	1.1(1)							
Alanine	3.1(3)	1.4(1)	2.2(2)			1.0(1)		
Valine	1.3(1)							
Methionine	0.9(1)							
Isoleucine		1.0(1)	1.2(1)		1.6(2)	1.0(1)		
Leucine		2.9(3)	3.1(3)		1.0(1)			
Tyrosine	3.0(4)	1.1(1)	1.6(1)		1.6(2)		0.7(1)	0.9(1)
Phenylalanine		1.1(1)	1.3(1)		1.1(1)	1.1(1)	1.0(1)	1.0(1)
Histidine			0.7		0.9(1)			
Lysine		2.7(2)	2.4(2)					
Tryptophan				0.8(1)				
Arginine	0.8(1)	2.9(3)	4.2(3)	1.0(1)	1.1(1)	1.4(1)	1.0(1)	1.0(1)
Total Residues	13	24	28	2	18	11	4	3
Amino Terminus	Tyr	Ile	Ala	Trp	Tyr	Asn	Tyr	Tyr
Residue Number	63-75	39-62	35-62	31-32	1-18	20-30	76-79	76-78

[a] The values in the table are molar ratios (those 0.3 or lower are omitted) without correction for destruction during hydrolysis, incomplete hydrolysis, or impurities. Within parentheses are given the number of residues deduced from sequence analysis.

[b] Determined by amino acid analysis of alkaline hydrolysate (16).

[c] Gla + Glu

TABLE III

Amino Acid Composition of

BNPS-Skatole Peptides from Matrix Gla Protein[a]

Amino Acid	B1	B2	B3	B4	B5	B7
DL-γ-Hydroxyglutamic acid[b]		1.3(1)	1.5(2)	1.0(1)	1.6(1)	0.9(1)
Aspartic acid	4.3(4)		2.1(2)	1.1(1)	1.0(1)	3.9(4)
Threonine						
Serine						3.9(5)
Glutamic acid	1.6(2)		1.0(1)	1.0(1)		5.4(6)
Proline			0.9(1)	1.0(1)		1.8(2)
Glycine	1.6(1)					
Alanine	3.7(4)					2.8(3)
Half-cystine	0.8(2)					
Valine	1.0(1)					
Methionine	0.7(1)					
Isoleucine		1.0(1)				2.7(3)
Leucine	1.3(1)		1.5(2)	0.9(1)	0.7(1)	1.1(1)
Tyrosine	3.9(5)	4.5	0.9(1)	0.6(1)	0.7	1.0(1)
Phenylalanine	2.2(2)					1.0(2)
Histidine						0.8(1)
Lysine	0.7(1)		1.1(1)	1.1(1)		1.0(1)
Tryptophan						N.D.(1)
Arginine	4.1(3)	2.1(2)	1.2(1)		0.9(1)	4.4(4)
Total Residues	27	4	11	7	4	35
Amino Terminus	Ala	Arg	Leu	Leu	Leu	Ser
Residue Number	53-79	38-41	42-52	42-48	49-52	3-37

[a] The values in the table are molar ratios (those 0.3 or lower are omitted) without correction for destruction during hydrolysis, incomplete hydrolysis, or impurities. Within parentheses are given the number of residues deduced from sequence analysis.

[b] Quantification is based on the color factor for glutamic acid.