Characterization of a Novel Dentin Matrix Acidic Phosphoprotein

IMPLICATIONS FOR INDUCTION OF BIOMINERALIZATION*

Anne George, Boris Sabsay, Philip A. L. Simonian, and Arthur Veis†

From the Division of Oral Biology, Northwestern University, Chicago, Illinois 60611

Acidic phosphorylated proteins have been shown to be prominent constituents of the extracellular matrix of bone and dentin. The acidic phosphoproteins of bone contain more glutamic acid than aspartic acid and a lower serine content than either. On the other hand, the major dentin acidic phosphoproteins, phosphophoryns, have been defined as aspartic acid- and serine-rich proteins, with a lesser content of glutamic acid. Both sets of phosphoproteins have been implicated as key participants in regulating mineralization, but it has been difficult to unify their mechanisms of action.

We have now identified, by cDNA cloning, a new serine-rich acidic protein of the dentin matrix, AG1, with a composition intermediate between the bone acidic proteins and dentin phosphophoryns. AG1 has numerous acidic consensus sites for phosphorylation by both casein kinases I and II. Immunochemical and organ culture biosynthetic studies show that AG1 is present in phosphorylated form at low levels in the dentin matrix. If fully phosphorylated, AG1 would bear a net charge of -176/molecule of 473 residues. AG1 contains single RGD integrin binding and N-glycosylation sequences. The overall picture that emerges is that of a matrix-associated acidic phosphoprotein, with a potentially high calcium ion binding capacity, present at levels compatible with a regulatory role in dentin mineralization.

Mineralization, the biogenic formation of mineral deposits, is one of the most widespread processes in nature (1). All aspects of this mineral deposition are important, but our attention has been focused on the processes in eukaryotes, where the mineral deposition at extracellular sites is mediated and regulated by the proteins and other components that the cells secrete to form the matrix. Although the specific components of the extracellular matrix may differ, it is generally considered that matrix-mediated mineralization follows a basic strategic plan in which the cells first form the structural components of the extracellular matrix and then deliver regulatory macromolecules that modify the properties of the matrix and lead to the location and stereospecific induction or nucleation of mineral deposition (2). Many candidates have been proposed as the regulatory nucleating proteins in various systems, but no definitive data have yet verified this role for any one protein.

Bone and dentin are the principal mineralized tissues of vertebrates. In these tissues, type I collagen fibrils serve as the extracellular matrix structural component, and a carbonate apatite constitutes the mineral phase. Several acidic proteins have been considered as the potential mineral inductive regulatory macromolecules (3). We have studied dentin for many years as a relatively simple model system and have determined that it contains a set of acidic phosphoproteins, called phosphophoryns, one or more of which could be involved in regulating the nucleation and crystal growth processes in dentin (4). The phosphophoryns, characterized by a very high content of aspartic acid and phosphorylated serine residues (5, 6), have proven to be extremely difficult to purify and sequence by standard protein chemical means (7). However, knowing the amino acid sequence of any of these molecules was considered essential to further understanding their specific roles in the dentin extracellular matrix. Thus, we have taken the approach of molecular biology and begun to clone the phosphophoryn genes to determine their sequences.

We report here the complete deduced amino acid sequence of the first of the unique dentin acidic phosphoproteins to be cloned and our deductions concerning phosphorylation sites and conformation, all of which have direct relevance to the potential role of this group of phosphoproteins in the mineralization process. We also demonstrate that this protein is present, uniquely, in dentin and that it is phosphorylated in vivo.

EXPERIMENTAL PROCEDURES

Extraction of RNA—Total RNA was isolated from the odontoblast-pulp fibroblast complex of incisors of 3-week-old Sprague-Dawley rats using the guanidine isothiocyanate method of Chirgwin et al. (8) as modified by Han et al. (9). Freshly extracted incisors, including the intact pulps, were cleaned of all adherent peridontal tissues and frozen in liquid nitrogen. Less than 2 min elapsed between extraction and freezing. The frozen incisors were then pulverized in 3-g batches in a liquid nitrogen freezer-mill (Spex Industries Inc.). Each batch of pulverized incisors was extracted in 20 ml of cold 5 M guanidine isothiocyanate by homogenization for 1 min in a Polytron homogenizer. The homogenate was centrifuged at 12,000 rpm for 5 min in a Beckman JA-20 rotor at 4 °C. The pellet, containing the nucleic acids, was resuspended in 30 ml of guanidine isothiocyanate, sheared again for 10 s in a Polytron homogenizer, and recentrifuged. The intact RNA and sheared DNA fragments were in the supernatant.

The RNA was precipitated by the addition of 350 μl of 1 M acetic acid and 7.5 ml of ethanol (−20 °C). After incubation at −20 °C for 3 h, the RNA precipitate was collected by centrifugation at 6000 rpm for 10 min at −10 °C. The pellet was dissolved in 10 ml of cold guanidine hydrochloride (6 M guanidine HCl, 25 mM EDTA, 10 mM β-mercaptoethanol, pH 7.0). The RNA was purified by two additional cycles of repetitive precipitation from cold guanidine hydrochloride.
Dentin Matrix Phosphoprotein 12625

with ethanol as described above. The reagents, tubes, and other equipment used in all of the procedures described above were treated with diethyl pyrocarbonate to inhibit ribonuclease activity.

**Isolation of mRNA and Preparation of Double-stranded DNA**—Poly(A+) RNA was obtained by chromatography on an oligo(dT)-cellulose column (9). The cDNA was synthesized from poly(A+) RNA using a CDNA synthesis kit following the manufacturer’s protocol (Amerham Corp.). The CDNA was methylated with EcoRI methylase and ligated to EcoRI linkers. The resultant cDNA was cut with EcoRI and passed over a Sephadex G-50 column. The fractions without free linkers were selected.

Cloning of Inserts into λgt11—The selected Sephadex G-50 fractions were ligated into a λgt11 vector that had been cut with EcoRI and dephosphorylated. After ligation, the clones were packaged, and *Escherichia coli* strain Y1090 cells were infected. The library was subsequently amplified to yield a titer of 1 × 10⁹ plaque-forming units/ml.

**Screening of Rat Incisor cDNA Library**—The infected *E. coli* cells were grown for 3.5 h at 37°C on 14-cm LB plates. These were overlaid with nitrocellulose membranes (Schleicher & Schuell) that had been soaked in 10 mM isopropl-1-thio-β-D-galactopyranoside and air-dried. The *E. coli* cells were grown overnight at 37°C. The membranes were washed with 0.05% Tween/Tris-buffered saline and blocked for 8 h with 3% bovine serum albumin and 1% normal goat serum at 4°C. The library was screened with a 1:1000 dilution of a polyclonal anti-riPP antibody. After a 1-h preincubation at 37°C in 5% CO₂, 95% air, 150 pCi of 32P₀₄ was added. Incubation was continued for 24–72 h. At the end of the labeling period, the teeth were demineralized with 0.5 M EDTA at pH 7.5 in the presence of 150 mM N-ethylmaleimide (15). After a 1-h preincubation at 37°C in 5% CO₂, 95% air, 150 μCi of 32P₀₄ was added. Incubation was continued for 24–72 h. At the end of the labeling period, the teeth were cooled to 4°C, washed in 15% NaCl, and split open to expose the pulp. The tooth shards were sonicated and then thoroughly extracted with 4.0 M guanidine hydrochloride to remove the constituents of the cellular and other nonmineralized tissue components. The proteins of the mineralized matrix were then extracted during demineralization with 0.5 M EDTA at pH 7.5 in the presence of a protease inhibitor mixture (30 mM α-aminoo-n-caproic acid, 5 mM benzamide, 2 mM phenylmethylsulfonyl fluoride, 0.5 mM N-ethylmaleimide). The phosphophoryns were isolated and purified by chromatography on DEAE-cellulose columns (6, 7). The solubilized proteins were analyzed by gel electrophoresis followed by silver staining or autoradiography, depending on the information desired.

**RESULTS**

The initial λgt11 cDNA library, expressed in *E. coli* Y1090 cells, yielded 39 positive plaques from ~10⁶ plaque-forming

---

1 The abbreviations used are: riPP, rat incisor phosphophoryn(s); bp, base pair(s); RACE, rapid amplification of cDNA ends.
Dentin Matrix Phosphoprotein

The cleavage site of the putative signal sequence is indicated by the arrow at residue 16. The single RGD sequence and the single putative N-glycosylation site are boxed. Each of the Ser residues that would be the most likely candidate for phosphorylation by either casein kinase I or II are shaded.
there was an intervening Arg, Lys, the 3′untranslated region consists of 1000 nucleotides and terminal to the residue and the presence of Asp or Glu at position -3 or +3 from the Ser residue. A residue was not considered a substrate if these have not been designated here.

A residue was not considered a substrate if these have not been designated here.

The criteria for selecting these sites were the embedding of the serine in an acidic sequence NH₂ or COOH-terminal to the residue and the presence of Asp or Glu at position -3 or +3 from the Ser residue. A residue was not considered a substrate if there was an intervening Arg, Lys, or His residue (19, 20). Although a number of the Thr residues are likely candidates for phosphorylation, these have not been designated here.

units in the initial screen with the polyclonal anti-riPP antibody, and these were verified by plaque purification. On the assumption that riPP contained some sequences of contiguous Asp residues, these plaques were screened with the poly-Asp nucleotide probe described above. This screening yielded a few very strongly positive plaques, which were then the focus of further study. It is important to note that there were other positive plaques, but we have not yet explored these. The selected strongly positive clones were cross-screened and found to fit into two groups with related (but different sized) inserts.

The first group of related clones to be sequenced proved to be virtually identical to rat bone osteopontin. There was a 95% sequence identity to nucleotides 1364–1457 of the bone protein according to the National Institutes of Health and EMBL nucleic acid data bases. The expressed protein osteopontin has not been shown to be a major constituent of matrix proteins, the 107 Ser residues are the predominant acid composition is shown in Table I.

### TABLE I

| Amino acid | Residues | mol % |
|------------|----------|-------|
| Ser        | 107      | 21.88 |
| Glu        | 72       | 14.72 |
| Asp        | 62       | 12.68 |
| Gly        | 36       | 7.36  |
| Gln        | 33       | 6.75  |
| Arg        | 29       | 5.93  |
| Thr        | 26       | 5.32  |
| Ala        | 20       | 4.09  |
| Pro        | 19       | 3.89  |
| Asn        | 17       | 3.48  |
| Leu        | 15       | 3.07  |
| Lys        | 11       | 2.25  |
| His        | 8        | 1.64  |
| Val        | 8        | 1.64  |
| Phe        | 7        | 1.43  |
| Ile        | 5        | 1.02  |
| Met        | 5        | 1.02  |
| Tyr        | 5        | 1.02  |
| Trp        | 2        | 0.41  |
| Cys        | 2        | 0.41  |
| Total      | 489      |       |

calculated from the composition data, the molecular weight of the secreted AG1 is 59,000.

To examine the tissue distribution of AG1 mRNA and to verify its presence in the odontoblast-pulp cells, an 800-bp nucleotide sequence from the 5′-end of the AG1 cDNA (Fig. 1) was used to carry out Northern analyses of mRNA extracted from rat skin, liver, brain, calvaria, tibia, and dentin. In this case, the mRNA from the odontoblasts was separated from the pulp cells by carefully extracting the pulp with an endodontic file. This leaves a clean, albeit incomplete layer of odontoblasts attached to dentin by their odontoblastic processes. Other pulp cells do not adhere to dentin. On the other hand, some odontoblasts break off at the process and remain attached to the pulp so that the pulp cell fraction is “contaminated” with odontoblasts. Fig. 3 shows that the odontoblasts contain the AG1 mRNA, whereas with equivalent poly(A⁺) RNA loading, no AG1 mRNA was detected in skin, liver, brain, or calvaria. Trace amounts of AG1 mRNA were detectable (data not shown) in tibia mRNA following prolonged exposure of the blot, but AG1 was certainly not prominent. These data are unequivocal in showing that AG1 mRNA is essentially odontoblast-specific.

Two experiments were carried out to demonstrate the presence of the expressed protein in the tissue. First, pAG1 was expressed at high density in the λgt11/E. coli system and blotted onto nitrocellulose. The filters with bound recombinant AG1 were then exposed to the polyclonal anti-riPP antibody used for its initial detection. The filters were extensively washed in binding buffer to eliminate nonbound components of the mixture of antibodies. The filters were then treated as equivalent to affinity columns, and the affinity-bound antibody was eluted and concentrated. An acrylamide gel electrophoresis run was carried out with the total DEAE-purified phosphophoryn extracted from fresh rat incisors in standard fashion. It is important to note that as in the organ culture dentin demineralization, the teeth were washed extensively with guanidine hydrochloride to remove nonmineral phase proteins prior to demineralization. Western analysis was carried out (Fig. 4, lane 3), with the result that a single band at Mr = 61,000 was detected by the filter-bound antibody, monospecific to recombinant AG1. The initial polyclonal anti-riPP antibody revealed several bands (lane 4). In addition to binding to the principal α- and β-riPPs, as expected, a band corresponding to the putative AG1 was evident. However, to see AG1 with the polyclonal anti-riPP antibody, total riPP had to be loaded on the gel at a 10-fold higher concentration than usual, and this revealed a number of stained components not usually observed (lane 4). In contrast to the other bands, only the monospecific anti-recombinant AG1 antibody yielded a single sharp band. These data show that a single protein reactive to the antibody specific to recombinant AG1 is present in the mineralized compartment of the dentin matrix. The apparent molecular weight of organ culture-produced AG1 (Mr = 61,000) is higher than the calculated core protein molecular weight. As indicated below, this is attributed to post-translational modification.

We have recently reported that the dentin phosphophoryns are substrates for phosphorylation by casein kinases I and II and that membrane-bound forms of these kinases are present in the endoplasmic reticulum/Golgi compartments of osteoblast-like cells (18). Using the most conservative rules for II is indicated in boldface type.
Dentin Matrix Phosphoprotein

**Fig. 3.** Northern blot analysis determining tissue distribution of AG1 message. Poly(A+) RNA was isolated from each tissue and probed with the 800-bp 5'-fragment of pAG1 as described for Fig. 1. Lane 1, rat skin; lane 2, rat liver; lane 3, rat brain; lane 4, rat calvaria; lane 5, rat odontoblasts; lane 6, nucleotide size standards (in kilobases).

**Fig. 4.** Expressed AG1 protein is present in dentin matrix and is phosphorylated. Lane 1, SDS-polyacrylamide gel electrophoresis of globular protein molecular weight standards. Numbers indicate the relative molecular weights in thousands. Lane 2, standard preparation of purified rat incisor phosphophoryns loaded at 20 µg and stained with Stains-All. The α- and β-phosphophoryns are evident. Lane 3, Western blot of the purified phosphophoryns from lane 2 (10 µg) reacted with the selected monospecific anti-recombinant AG1 antibody. Lane 4, Western blot of the purified phosphophoryns from lane 2 reacted with the total polyclonal anti-riPP antibody. In this lane, the amount of phosphophoryns loaded was >75 µg. At lower concentrations, the AG1 band could not be seen. The origin of the intense band at ~70 kDa is not known, but is seen under these high loading concentrations. Lane 5, autoradiograph of the purified phosphophoryn components extracted from rat incisors cultured in vitro and labeled with 32P for 72 h following the procedure of DiMuzio et al. (15). The incisors were washed with 4.0 M guanidine HCl to remove cellular components prior to demineralization of dentin and phosphophoryn extraction. Protease inhibitors were included during extraction (7). A band corresponding to AG1 is prominent in the immunoblots and autoradiograph.

This observation suggested that a second way of revealing the presence of AG1 might be to label incisor dentin in culture with 32PO4. After 72 h in organ culture, the mineralized matrix-associated proteins were extracted, and the phosphophoryn fraction was isolated and purified by DEAE chromatography. The 32P-labeled cultures yielded matrix-associated mineral phase bound radiolabeled phosphophoryns. Gel electrophoresis of the protein showed the presence of a radiolabeled band corresponding precisely to the position of the specific anti-AG1 antibody (Fig. 4, lane 5). In fact, in the organ culture system, AG1 was more heavily labeled than the higher molecular weight phosphophoryns. As noted above, the apparent molecular weight of this band is 61,000. The addition of the minimum 55 phosphate groups would increase core mass by 4,350 from 53,000, a value in excellent agreement with the observed data. These data substantiate that pAG1 represents and encodes a real protein present in dentin in a phosphorylated state.

**DISCUSSION**

The dentin extracellular matrix consists of ~90% collagen, with the rest composed of phosphoproteins and acidic glycoproteins (21, 22). The principal phosphoproteins discussed in the literature are the serine- and aspartic acid-rich phosphophoryns (5). All of the above data indicate that rat dentin...
contains another previously unreported unique, acidic, and potentially highly phosphorylated extracellular matrix protein, designated here as AG1. Relative to the phosphophoryns, AG1 has a high content of Glu residues and some similarities in amino acid composition to osteopontin and bone sialoprotein. However, AG1 also has features that have been proposed for phosphophoryn and seems to be exceptionally suited to play some regulatory role within the matrix of mineralized bone and dentin.

AG1 is remarkably hydrophilic. As shown in Fig. 5, a plot of the hydrophobic/hydrophilic balance using the hydropathy measure of Kyte and Doolittle (23) shows that virtually only the putative signal peptide domain is hydrophobic. The carboxyl-terminal region, which contains the single cysteine residue, is also hydrophilic. Conformational analysis algorithms developed by the University of Wisconsin Genetics Computer Group program using Chou-Fasman methods predict that the AG1 backbone (i.e. before any potential post-translational modification such as phosphorylation or glycosylation) might have a few $\beta$-turn regions, but there are no regions of ordered long-range $\beta$-sheet or $\alpha$-helix (Fig. 6).

In spite of the fact that the cDNA was detected as reactive with the putative poly-Asp oligonucleotide probe, no extended poly-Asp sequence is present. There are, however, acidic patches of Asp and Glu, such as at residues 84–89 (EEDD) and 254–257 (EEDD). Furthermore, there are several regions with 3 consecutive acidic residues and many more with 2.

Table II shows that exactly half of the Glu residues appear as consecutive EE sequences, and except at positions 105–106, these are in Ser- and Thr-rich domains. Only two of the EE pairs are near or adjacent to a basic residue (positions 281–282 and 441–442). Similarly, many of the Asp residues are in Ser-, Thr-, and Glu-rich sequences.

The acidic nature of the molecule is markedly enhanced when the potential for phosphorylation of strategically placed Ser and Thr is considered. We have shown that casein kinase II and membrane-associated casein kinase II-like kinases in osteoblasts can readily phosphorylate riPP (18) and that membrane-associated odontoblast casein kinase II-like kinases can rephosphorylate dephosphorylated riPP. Thus, it is likely that endoplasmic reticulum/Golgi-associated casein kinase II could have access to nascent AG1 during its chain synthesis and secretion. As noted above, Ser or Thr residues in acidic sequences with Glu, Asp, or Ser(P) at position +3 from the carboxyl-terminal side are preferred casein kinase II substrates. Additional upstream Asp or Glu at position +2 may also be sufficient to direct Ser or Thr phosphorylation by casein kinase II in model peptides and physiological protein substrates. Substrate recognition sites for casein kinase I include Ser(P), Asp, or Glu at position -3 from the amino-terminal direction from the substrate Thr and Ser. Ser(P) is the most active phosphorylation determinant in model peptides (27). The more acidic the neighboring environment, the better the substrate sequence. Using these minimal criteria and the deduced sequences of Fig. 2, it appears

2 C. Sfeir and A. Veis, unpublished data.
that 55 of the 107 serine residues in the secreted AG1 could be readily phosphorylated. If this were the case and with the Ser(P) acid dissociation constant at pH 6.8 (28), AG1 would carry a charge of -88 from the phosphate groups alone and a net charge of -175/molecule at physiological pH. Such a molecule should have a very high capacity for binding divalent cations such as calcium, as already determined for the phosphophoryns (29).

Inspection of the amino acid sequence determined for recombinant AG1 shows several sequence domains that contain an abundance of phosphorylation consensus sites for casein kinases I and II. Consider the sequence from residues 442 to 454 (SNSTGTSSESEED). Each S and T residue in the sequence is a potential substrate for casein kinase I1. It is possible that the hypercharged domain. Further analysis of this type is not possible to examine the interactions of the AG1 protein with both collagen and mineral for assessing its potential role in the regulation of mineralization of dentin as well as to determine its localization within the matrix and other aspects of its function and production.

Acknowledgments—we thank Drs. J. H. Han and P. Tylzanowski for help in preparing the initial odontoblast cDNA library and Drs. L. J. Fisher and M. Young (National Institute of Dental Research) for valuable guidance in the library screening. We are indebted to Dr. Paul Price for suggestions concerning the isolation of the monospecific anti-AG1 antibody.

REFERENCES

1. Lowenstam, H. A., and Weiner, S. (1989) On Biomineralization, Oxford University Press, Oxford
2. Veis, A., and Sabby, B. (1983) in Biomineralization and Biological Metal Accumulation (Westbrook, P., and de Jong, E. W., eda.) pp. 191–203, D. Reidel Publishing Co., Dordrecht, The Netherlands
3. Gorski, J. P. (1992) Calkin. Tissue Res. 40, 391–396
4. Veis, A. (1989) in Biomineralization: Chemical and Biochemical Perspectives (Mann, S., Webb, J., and Williams, R. J., eda.) pp. 199–222, VCH Verlagsgesellschaft mbH, Weinheim, Germany
5. Dimuzio, M. T., and Veis, A. (1978) Calkin. Tissue Res. 25, 169–178
6. Stetler-Stevenson, W. G., and Veis, A. (1979) Biochemistry 22, 4328–4335
7. Sabby, B., Stetler-Stevenson, W. G., Lechner, J. H., and Veis, A. (1991) Biochim. J. 276, 690–707
8. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Butler, W. J. (1979) Biochemistry 18, 5294–5299
9. Han, J. H., Stratton, C., and Rutter, W. J. (1987) Biochemistry 26, 1617–1625
10. Rahima, M. R., Tsay, T.-G., Andujar, M., and Veis, A. (1988) J. Histochem. Cytochem. 36, 153–157
11. Sanger, F., Coulson, A. R., Borell, B. G., and Smith, A. J. H. (1980) J. Mol. Biol. 143, 161–178
12. Takah, S., and Richardson, C. C. (1989) J. Biol. Chem. 264, 6447–6458
13. Barnes, W. M., Bevan, M., and Sou, P. H. (1983) Methods Enzymol. 101, 98–122
14. Froehman, M. A., Dush, M. K., and Martin, G. R. (1988) Proc. Natl. Acad. Sci. U.S.A. 85, 9898–9902
15. DiMaio, M. T., Blow, M., and Butler, W. T. (1988) Calkin. Tissue Int. 37, 242–249
16. Butler, W. T. (1991) J. Biol. Buccale 19, 83–89
17. van Heijl, G. (1980) J. Mol. Biol. 144, 99–106
18. Wu, C. B., Pelech, S. L., and Veis, A. (1992) J. Biol. Chem. 267, 16588–16594
19. Roach, F. J., and Flotow, H. (1991) J. Biol. Chem. 266, 3724–3727
20. Meggio, F., Marchiori, F., Borin, G., Chessa, G., and Pinna, L. A. (1984) J. Biol. Chem. 260, 14576–14579
21. Schlueter, R. J., and Veis, A. (1984) Biochemistry 3, 1657–1665
22. Carmichael, D. J., Veis, A., and Wang, E. T. (1971) Calcif Tissue Res. 7, 331–344
23. Kyte, J., and Doolittle, R. F. (1982) J. Mol. Biol. 157, 105–132
24. Jameson, B. A., and Wolf, H. (1988) Comput. Appl. Biosci. 4, 181–186
25. Meggio, F., Ferich, J. W., Meyer, H. E., Hoffmann-Poesske, E., Lennon, D. P. W., John, B. R., and Pinna, L. A. (1989) Eur. J. Biochem. 186, 461–464
26. Walton, G. M., Spiesz, J., and Gill, G. N. (1985) J. Biol. Chem. 260, 4745–4750
27. Flotow, H., Gravess, P. R., Wang, A., Fiol, C. J., Roesske, R. W., and Roach, P. J. (1990) J. Biol. Chem. 265, 14294–14299
28. Lee, S., and Veis, A. (1970)  1970 Biochemistry 16, 2971–2979
29. Stetler-Stevenson, W. G., and Veis, A. (1987) Calkin. Tissue Int. 40, 97–102
30. Meir, U. T., and Biolo, G. (1992) Cell 76, 127–138