Communication

Sequence Determination of an Extremely Acidic Rat Dentin Phosphoprotein*

(Received for publication, May 17, 1996, and in revised form, July 9, 1996)

Helena H. Ritchie‡ and Lee-Ho Wang*†

From the ‡Department of Pediatrics, The University of Iowa, Iowa City, Iowa 52242 and the †Division of Hematology, the Department of Internal Medicine, the University of Texas Medical School, Houston, Texas 77030

The mineralization process associated with the conversion of predentin to dentin is believed to be initiated and controlled by a set of acidic regulatory noncollagenous proteins (NCPs) which include phosphophoryn, the major NCP in dentin. Phosphophoryn binds tightly to collagen and is believed to initiate the formation of apatite crystals which play a central role in the mineralization process. During the process of analyzing the 3′ end of an odontoblast-specific cDNA which codes for dentin sialoprotein (Ritchie, H. H., Hou, H., Veis, A., and Butler, W. T. (1994) J. Biol. Chem. 269, 3698–3702), we discovered a 801-base pair open reading frame. This downstream open reading frame encodes a putative leader sequence and a very acidic mature protein sequence having a deduced amino acid composition containing high percentages of both Ser (43%) and Asp (31%) residues which closely coincides with the amino acid composition of phosphophoryns from human, bovine, rat, and rabbit (i.e. Asp (30–40%) and Ser (38–50%)). This newly identified cDNA therefore encodes a protein with characteristics similar to phosphophoryn. Here we present the cDNA sequence, the deduced amino acid sequence, and the prospective Ser residue-specific casein kinase I and II phosphorylation sites for this putative phosphoprotein.

The calcification process that accompanies the transition of predentin to dentin is poorly understood, due in part to the difficulties in isolating and characterizing unique sets of extracellular matrix molecules that contribute to this complex process (1–4). Phosphophoryn, the most abundant noncollagenous protein in dentin, is secreted by odontoblasts through odontoblastic processes and appears at the mineralization front within a short time after labeling with [33P]phosphate (5, 6). Phosphophoryn is known to bind large amounts of calcium with a relatively high affinity (7) and to then form an insoluble aggregate in the presence of Mg2+ and Ca2+ (8). Because of its affinity for calcium, phosphophoryn may concentrate these ions and participate in the formation of apatite crystals. For example, Linde and co-workers (9) have demonstrated that when phosphophoryn is immobilized on a stable support and incubated in physiological solutions of calcium and phosphate, phosphophoryn induced the formation of hydroxyapatite (HAP).1 Studies by the same group (10) and by Boskey et al. (11) also suggested a dual role for phosphophoryn as both an initiator of HAP formation at low phosphophoryn concentrations and as an inhibitor of HAP formation at higher phosphophoryn concentrations.

Phosphophoryn is also believed to have a specific affinity for collagen (2, 12, 13) which comprises as much as 80% of the protein in dentin. Furthermore, phosphophoryn was found to be specifically associated with the “e” band of collagen (14). This site-specific protein-protein interaction, coupled with phosphophoryn’s ability to initiate or inhibit HAP formation when calcium is present, has lead to the currently accepted view that phosphophoryn plays a central role in the mineralization process by virtue of its ability to target mineralization to selected sites as well as to control mineralization, temporally, with organ development.

While only several NH2-terminal residues and small proteolytically obtained segments of internal amino acid sequences of phosphophoryn are currently known, the complete amino acid sequence of phosphophoryn has not as yet been reported. During the process of analyzing the 3′ end of dentin sialoprotein (DSP) cDNA, another recently cloned dentin-specific protein (15), an open reading frame with a size of 801 bp was revealed. This open reading frame was found to encode a putative leader sequence and a deduced very acidic mature protein sequence with an amino acid composition comprised primarily of Ser (43%) and Asp (31%) residues which coincides with the amino acid composition of phosphophoryns from human, bovine, rat, and rabbit (i.e. Asp (~30–40%) and Ser (~38–50%)) (2, 16–21). Here we present the first reported cDNA sequence, the deduced amino acid sequence, and the postulated Ser residue-specific casein kinase I and II phosphorylation sites for this putative phosphoprotein.

* This work was supported by National Institutes of Health Grant DE11442–01 (to H. H. R.). 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.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank®/EBI Data Bank with accession number(s) U63111. §To whom correspondence should be addressed: BB135, Dept. of Pediatrics, the University of Iowa, Iowa City, IA 52242. Tel.: 319-353-2055; Fax: 319-335-4489; E-mail: dritchie@tic-pot.ctic.uiowa.edu.

1 The abbreviations used are: HAP, hydroxyapatite; DSP, dentin sialoprotein; NCP, noncollagenous protein; bp, base pair(s); kb, kilobase(s); PCR, polymerase chain reaction.

EXPERIMENTAL PROCEDURES

RNA Preparation and Reverse Transcription-PCR—The total RNA was extracted from adult rat incisors using RNAzol® (Biotech Laboratories, Inc., Houston, TX). A cDNA pool was synthesized from total RNA using an oligo(dT) primer and reverse transcriptase. This cDNA pool was then denatured at 95°C for 5 min and amplified with the primer set comprising an oligoprimer corresponding to rat DSP cDNA nucleotide sequence 1054–1069 (15) and a poly(dT) primer. PCR was then performed as follows: denaturation (1 min at 94°C), reannealing (1 min at 56°C), and amplification (3 min at 65°C), for 40 cycles.

DNA Sequencing—The PCR products recognized by the 3′ end rat DSP probe were subcloned into TA vectors using standard techniques (22). Following company procedures, Erase-A-Base Kit (Promega, WI) was used to generate unidirectional deletions of the 2-kb insert for DNA sequencing. DNA was sequenced according to Sanger et al. (23).

Northern Blot Analysis—The total RNA from rat incisor was electrophoresed using a 1.2% agarose gel containing 2.2 M formaldehyde (22). RNA was transferred onto a nitrocellulose paper and hybridized overnight with a 32P-labeled putative rat phosphophoryn probe at 42°C in 50% formamide and 6× SSC. The filter was washed with 2× SSC and 0.1× SSC.

This is an open access article under the CC BY license.

21695
RESULTS

The Deduced Phosphophoryn cDNA Sequence and the Deduced Amino Acid Sequence—Using an oligonucleotide corresponding to DSP cDNA nucleotide sequence 1549–1565 (15) and a poly(dT) primer to amplify the cDNA pool from rat incisors, a 2-kb PCR fragment recognized by 32P-labeled 3' end rat DSP probe was obtained and subcloned into a TA vector for DNA sequencing by the Erase-A-Base technique. An open reading frame, located immediately downstream from the 3' coding region of DSP (i.e. identical to the reported DSP sequence) (15) was identified in this 2-kb insert. This open reading frame contained 801 nucleotides, representing 267 amino acids, including the 27-amino acid putative leader sequence and a 240-amino acid mature protein sequence (Fig. 1). The deduced 27-residue leader sequence is Met-Gly-His-Ser-Arg-Ile-Gly-Ser-Ser-Asn-Ser-Asp-Gly-Ser-Met-Gln-Gly. The completed DNA sequence for this clone was found to contain a translation start site (ATG; code for putative initiation codon Met) located at nucleotide position 43. Bracket enclose the 27-amino acid signal peptide sequence. Underlined 4-amino acid sequence is identical with the protein microsequence data for rat (25). Parentheses enclose a variety of proteins involved in different cellular functions known to phosphorylate serine and/or threonine residues in casein kinase I and II sites since these ubiquitous enzymes are case in kinase I and II isolates from osteoblast-like cells can catalyze phosphorylation of nascent dentin phosphophoryns (28). Because Ser residues are the predominant amino acids in our putative phosphophoryn sequence, we examined the potential phosphorylation sites in the consensus sequence (Asp/Glu)-X-Ser (29). Phosphophoryn contains 29 putative primary casein kinase I sites (Fig. 1). Interestingly, in some cases, the phosphorylation of the serine residue downstream from this newly phosphorylated serine (Ser(P)-X) coincides with the amino acid composition of authentic phosphophoryns from rat, rabbit, bovine, and human (29, 30). Phosphorylation Sites—Table I compares the amino acid compositions of deduced rat phosphophoryn versus native rat, rabbit, bovine, and human phosphophoryns.

Table I. Phosphophoryn protein is secreted by odontoblasts in a highly phosphorylated form (2, 6, 26, 27). In vitro studies have demonstrated that membrane-bound forms of casein kinases I and II isolated from osteoblast-like cells can catalyze phosphorylation of nascent dentin phosphophoryns (28). Because Ser residues are the predominant amino acids in our putative phosphophoryn sequence, we examined the potential casein kinase I and II sites since these ubiquitous enzymes are known to phosphorylate serine and/or threonine residues in a variety of proteins involved in different cellular functions (29, 30).

Table I). Phosphophoryn protein is secreted by odontoblasts in a highly phosphorylated form (2, 6, 26, 27). In vitro studies have demonstrated that membrane-bound forms of casein kinases I and II isolated from osteoblast-like cells can catalyze phosphorylation of nascent dentin phosphophoryns (28). Because Ser residues are the predominant amino acids in our putative phosphophoryn sequence, we examined the potential casein kinase I and II sites since these ubiquitous enzymes are known to phosphorylate serine and/or threonine residues in a variety of proteins involved in different cellular functions (29, 30).
Rat Dentin Phosphophoryn

Fig. 2. A, the potential casein kinase I (*1) and II (*2) phosphorylation sites for the putative phosphophoryn. *1a, primary target site (i.e. (Asp/Glu)-X-Ser) for casein kinase I (29, 36); *2b–*2r, secondary, tertiary (and so forth) sites for casein kinase II (29, 39–41); *2a, primary target site (i.e. (Ser/Thr)-X-(Ser/Thr)) for casein kinase II (29, 39–41); *2b–*2r, secondary, tertiary (and so forth) sites for casein kinase II (i.e. (Ser/Thr)-X-(Ser/Thr)). B, the acidic residues and potential acidic regions generated by casein kinases I and II are depicted in boxes. Many of these acidic regions consist of (DD), (pSpS), (pSD), units (where n = 1, 2, or 3). Several significant acidic domains containing for example 21 residues (96–116), immediately followed by a series of 2 residues (i.e. mainly PSD) repeats (118–155) and 24 residues (176–199) are particularly evident.

X-Ser). When this secondary target site is phosphorylated, it then determines the tertiary Ser site for casein kinase I and so forth. For example, Ser\(^{103}\) is a primary target for casein kinase I. Based on this phosphorylation mechanism, once Ser\(^{103}\) is phosphorylated, it can trigger phosphorylation of the following 17 target Ser sites ending at Ser\(^{154}\) (1a through 1r) (see Fig. 2). When all those subsequent secondary, tertiary ... etc. sites are included, a total of 66 potential phosphorylation sites may be available for casein kinase I.

Casein Kinase II Sites—This rat phosphophoryn contains 23 potential primary casein kinase II sites ((Ser/Thr)-X-(Glu/Asp)). Many of these sites overlap with the sites for casein kinase I, and 13 sites are specific for casein kinase II. It is plausible that the overlapping sites ensure the phosphorylation of specific phosphophoryn domains that may be crucial to collagen and/or calcium binding activities during dentinogenesis (2, 4, 7, 9, 11–13) (see Fig. 2). For example, Ser\(^{134}\), once phosphorylated, becomes the primary target site of casein kinase II and enables the phosphorylation of Ser\(^{154}\). Following this mechanism, the casein kinase II could then phosphorylate the following 17 serines (spaced every two amino acids upstream from Ser\(^{134}\) ending at Ser\(^{154}\) (2a–2r in Fig. 2). Overlapping kinase I and II activities within this Ser-rich domain (103–154) could therefore provide a "safeguard" mechanism to ensure the phosphorylation of this particular domain. There are, in total, 55 potential casein kinase II sites of which 37 sites overlap with casein kinase I.

Overall, 78% of the Ser residues (81 out of 103) could potentially become phosphorylated by casein kinases I and/or II. This number is consistent with the reported 85–87% of phosphorylated Ser residues in native phosphophoryn (2, 21). With this number of phosphoserines, phosphophoryn would carry an additional charge of –130 from phosphate groups alone (i.e. –1.6/ Ser(P) or Thr(P)). Furthermore, if 3 of the 9 Thr residues are phosphorylated by casein kinase II, the charge from the combined phosphoserine and phosphothreonine residues would be –134. Therefore, phosphophoryn would carry an overall net charge of –213 at physiological pH. Such a molecule would have a very high capacity for binding divalent cations such as calcium and magnesium, as reported for phosphophoryn (2, 7, 8, 21).

Potential N-Glycosylation Site—Seven potential N-glycosylation sites (Asn-Xaa-(Ser/Thr)) (31, 32) are present in this novel dentin protein at amino acid positions 31, 37, 69, 170, 202, and 261 (Fig. 1). Because of the overlap of potential casein kinase sites and N-glycosylation sites, only positions 31 and 37 are likely to be glycosylated. The other sites would more likely be subjected to phosphorylation of Ser residues.

Northern Blot Analysis—We examined newborn rat tooth germs for putative phosphophoryn mRNA expression. To eliminate the DSP DNA sequence, we constructed a cDNA probe for \(^{32}P\) random primer labeling containing only the phosphophoryn DNA sequence (i.e. from nucleotide position 208). Northern blot analysis indicated that multiple 4.6-kb transcripts were detected in the newborn tooth germs (Fig. 3). Therefore, transcripts for the putative phosphophoryn are expressed in the rat tooth germs.

DISCUSSION

During the process of analyzing the 3' end of an odontoblast-specific cDNA which codes for dentin sialoprotein (15), we discovered an open reading frame with a size of 801 bp. Our newly discovered cDNA was found to encode a novel rat dentin protein whose characteristics are in solid agreement with the following reported features for phosphophoryn: (i) the predicted NH\(_2\)-terminal amino acid sequence (i.e. Asp-Asp-Pro-Asn) is identical to that derived from protein microsequencing for one form of rat phosphophoryn (25), (ii) this is the NH\(_2\)-terminal amino acid sequence (i.e. Asp-Asp-Pro-Asn-Ser-Ser) obtained from our putative phosphophoryn (Fig. 1) agrees with that reported by Reynolds and co-workers (33) (i.e. Asp-Ser-Pro-Asp-Ser-Ser) for bovine phosphophoryn, (iii) amino acid sequences of Asp and Ser-Ser-Ser were found interspersed in our putative rat phosphophoryn. Additionally, a sequence of Asp-Asp-Asp-Pro-Asn) is Seven potential casein kinase II sites (35; 2a–2r) of which 37 sites overlap with casein kinase II. Multiple transcripts for putative phosphophoryn were detected near 4.6 kb in the newborn rat tooth germs.
Ser-Ser-Ser was identified in our putative protein. These se-
quence combinations of Asp-Ser, Asp-Ser-Ser, and Asp-Ser-
Ser were also observed in the NH₂-terminal 50 residues of
bovine phosphophoryn (33), and (iv) its deduced amino acid
composition contained high percentages of Ser (43%) and Asp
(31%) residues which coincided with the amino acid composi-
tion of phosphophoryns from human, bovine, rat, and rabbit
(i.e. Asp (~30–40%) and Ser (~38–50%)) (2, 16–21).

Phosphophoryn is the most acidic protein so far discovered.
The interspersed arrangement of Ser and Asp residues enables
phosphophoryn to be an excellent substrate for casein kinases
I and II phosphorylation action. The 78% of Ser residues in this
protein which can potentially be phosphorylated coincide with
the reported 85–87% of Ser(P) in authentic phosphophoryn
(21). As discussed previously (see “Results”), many of these
phosphorylation reactions occur at secondary and tertiary Ser
sites and therefore result from a phosphorylation cascade-type
mechanism involving casein kinases I and II operating over
similar domains but in opposite directions to ensure complete
phosphorylation within these specificSer-richdomains.

Roach and co-workers (34–36) have reported that threonine
can also serve as a substrate for casein kinase I. Based on this
information, all 9 threonine residues and an additional 10
serine residues could also be phosphorylated. This mechanism
could enable the sequential phosphorylation of a 47-residue
acidic patch extending from Ser96 to Ser142. In this case, 88% of
the Ser residues (91 out of 103) could potentially become phos-
phorylated by casein kinases I and II. The 78% of Ser residues in
this protein which can potentially be phosphorylated coincide
with the reported 85–87% of Ser(P) in authentic phosphophoryn
(21). As discussed previously (see “Results”), many of these
phosphorylation reactions occur at secondary and tertiary Ser
sites and therefore result from a phosphorylation cascade-type
mechanism involving casein kinases I and II operating over
similar domains but in opposite directions to ensure complete
phosphorylation within these specific Ser-rich domains.

By using a cDNA probe containing only the phosphophoryn
dNA sequence beginning from position 208, we determined by
Northern blot analysis whether this putative phosphophoryn
cDNA was indeed transcribed in tooth germ. Multiple tran-
scripts, sized around 4.6 kb, were detected (Fig. 3). The pres-
ence of multiple phosphophoryn transcripts may be due to more
than one phosphophoryn gene or due to alternative splicing.
However, it is equally likely that these multiple transcripts are
due to the use of multiple polyadenylation signals, similar to
many other mRNAs encoding extracellular matrix proteins.
Further experiments are needed to determine the origin of
these multiple transcripts. However, our Northern blot
strongly suggests that the putative phosphophoryn mRNA was
present in the rat tooth germ total RNA pool. Therefore, the
DNA sequence for this novel protein is unlikely to be an arti-
fact. Taken together, we strongly feel that rat tooth genes do
actively synthesize the mRNA of this putative phosphophoryn.

The presence of both DSP and phosphophoryn DNA se-
quences in the PCR product could be due to an artifact gener-
ated during reverse transcription-PCR or subsequent cloning
processes. However, the possibility of a bicistronic gene could
not be excluded. Further work, such as the examination of
these two genes at the genomic level, is needed to investigate
these possibilities.

Acknowledgment—We thank Dr. David G. Ritchie for helpful discus-
sion and contributions during the preparation of this manuscript.

REFERENCES
1. Veis, A. (1978) in Ions in Macromolecular and Biological Systems (Everett, D.
H., and Vincent, B., eds) pp. 259–272, Sciencitania, Bristol, UK
2. Veis, A. (1985) in Chemistry and Biology of Mineralized Tissues (Butler, W. T.,
ed) pp. 170–184, EBSCO Media, Birmingham, AL
3. Linde, A., and Goldberg, M. (1993) Crit. Rev. Oral Biol. Med. 4, 679–728
4. Butler, W. T., and Ritchie, H. H. (1995) Int. Dev. Biol. 39, 169–179
5. Weinberg, M., and Leblond, C. P. (1973) J. Cell Biol. 56, 838–845
6. Rahima, M., Tsay, T.-G., Andujar, M., and Veis, A. (1988) J. Histoch. 
Cytochem. 36, 153–157
7. Marsh, M. E. (1989) Biochemistry 28, 346–352
8. Marsh, M. E. (1989) Biochemistry 28, 339–345
9. Linde, A., Lusi, A., and Crenshaw, M. A. (1989) Calcif. Tissue Int. 44,
286–295
10. Lusi, A., Crenshaw, A., and Linde, A. (1988) Arch. Oral Biol. 33, 685–691
11. Boskey, A. L., Muresca, S., Doyt, S., Sabsay, B., and Veis, A. (1990) Bone
Miner. 11, 55–65
12. Stetler-Stevenson, W., and Veis, A. (1986) Calcif. Tissue Int. 38, 135–141
13. Fujisawa, R., and Kuboki, Y. (1992) Calcif. Tissue Int. 51, 436–442
14. Traub, W., Jodkaikin, A., Arad, T., Veis, A., and Sabsay, B. (1992) Matrix 12,
197–203
15. Ritchie, H. H., Hau, V., and Butler, W. T. (1994) J. Biol. Chem. 269,
3698–3702
16. Richardson, W. S., Beegle, W. F., Butler, W. T., and Munksgaard, E. C. (1977) 
J. Dent. Res. 56, 233–237
17. DiMuzio, M., and Veis, A. (1978) Calcif. Tissue Res. 225, 169–178
18. Termine, J., Belcourt, A., Miyamoto, M. S., Conn, K. M., and Nylen, M. U.
(1980) J. Biol. Chem. 255, 9769–9772
19. Stetler-Stevenson, W., and Veis, A. (1983) Biochemistry 22, 4326–4335
20. Takagi, Y., and Veis, A. (1984) Calcif. Tissue Int. 36, 259–265
21. Linde, A. (1985) in Chemistry and Biology of Mineralized Tissues (Butler, W. T., ed) pp. 344–355, EBSCO Media, Birmingham, AL
22. Sambrook, J., Fritsch, E. F., Maniatis, T. (1989) Molecular Cloning: A Labora-
tory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Har-
bor, NY
23. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A.
74, 5463–5467
24. Kozak, M. (1986) Cell 44, 283–292
25. Butler, W. T., Bhowm, M., DiMuzio, M. T., Cottman, W. H., and Linde, A. (1983) 
Arch. Biochem. Biophys. 225, 176–186
26. DiMuzio, M. T., and Veis, A. (1978) J. Biol. Chem. 253, 6845–6852
27. Munksgaard, E., Richardson, W., and Butler, W. T. (1978) Arch. Oral Biol. 23,
583–585
28. Wu, C., Pelek, S., and Veis, A. (1992) J. Biol. Chem. 267, 16588–16594
29. Pinna, L. (1990) Biochim. Biophys. Acta 1054, 267–284
30. Tuazon, P., and Traugh, J. (1991) Adv. Second Messenger Phosphoprotein Res. 
23, 123–164
31. Marshall, R. (1972) Annu. Rev. Biochem. 41, 673–702
32. Bause, E. (1983) Biochem. J. 209, 331–336
33. Crossley, M. A., Huq, N. L., Kirsbaum, L., and Reynolds, E. C. (1996) J. Dent.
Res. 75, 154
34. Flook, H., and Roach, P. J. (1989) J. Biol. Chem. 264, 9126–9128
35. Flook, H., Graves, P. R., Wang, A., Fid, C. J., Roeske, R. W., and Roach, P. J.
(1990) J. Biol. Chem. 265, 14264–14269
36. Flook, H., and Roach, P. J. (1991) J. Biol. Chem. 266, 3724–3727
37. Evans, J., Chiu, T., and Chan, S. (1994) Biochimica 34, 1359–1375
38. Evans, J., and Chan, S. (1994) Biopolymers 34, 507–527
39. Meggio, F., Marchiori, F., Bosin, G., Chessa, G., and Pinna, L. A. (1994) Biochim.
Biophys. Acta 1259, 14576–14579
40. Kuenzel, E., Mulligan, J. A., Summecorn, J., and Krebs, E. G. (1987) J. Biol.
Chem. 262, 9136–9140
41. Meggio, F., Perich, J., Meyer, H. E., Hoffmann-Porsske, E., Lennon, D. P.
W., Jhons, R. B., and Pinna, L. A. (1989) Eur. J. Biochem. 186, 459–464