THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 10, Issue of March 7, pp. 7949–7955, 2003
Printed in U.S.A.

Delineation of the Hydroxyapatite-nucleating Domains of Bone Sialoprotein*

Received for publication, November 22, 2002
Published, JBC Papers in Press, December 18, 2002, DOI 10.1074/jbc.M211915200

Coralee E. Tye‡, Kevin R. Rattray‡, Kevin J. Warner‡, Jonathan A. R. Gordon‡, Jaro Sodek§, Graeme K. Hunter‡, and Harvey A. Goldberg‡¶

From the ‡Canadian Institutes of Health Research Group in Skeletal Development and Remodeling, Division of Oral Biology and Department of Biochemistry, School of Dentistry, University of Western Ontario, London, Ontario N6A 5C1, Canada and the §Canadian Institutes of Health Research Group in Matrix Dynamics, Faculty of Dentistry, University of Toronto, Toronto, Ontario M5S 3E2, Canada

Bone sialoprotein (BSP) is a highly modified, anionic phosphoprotein that is expressed almost exclusively in mineralizing connective tissues and has been shown to be a potent nucleator of hydroxyapatite (HA). Two polyglutamic acid (poly[E]) regions, predicted to be in an α-helical conformation and located in the amino-terminal half of the molecule, are believed to be responsible for this activity. Using a prokaryotic expression system, full-length rat BSP was expressed and tested for HA nucleating activity in a steady-state agarose gel system. The unmodified protein is less potent than native bone BSP, indicating a role for the post-translational modifications in HA nucleation. Site-directed mutagenesis of the poly[E] regions in full-length BSP was performed, replacing the poly[E] with either polyaspartic acid (poly[D]) or polyalanine (poly[A]) to examine role of charge and conformation, respectively, in HA nucleation. Replacement of single domains with either poly[A] or poly[D] did not alter nucleating activity nor did replacement of both domains with poly[D]. Replacement of both domains with poly[A], however, significantly decreased nucleating activity. In addition, two recombinant peptides, each encompassing one of the two poly[E] domains, were expressed and tested for nucleating activity. Whereas the peptide encompassing the second poly[E] domain was capable of nucleating HA, the first domain peptide showed no activity. The conformation of the wild-type and mutated proteins and peptides were studied by circular dichroism and small angle x-ray scattering, and no secondary structure was evident. These results demonstrate that a sequence of at least eight contiguous glutamic acid residues is required for the nucleation of HA by BSP and that this nucleating "site" is not α-helical in conformation.

Mineralization of the extracellular matrix in bone, dentin, and cementum is a complex, poorly understood process that is believed to involve both hydroxyapatite-nucleating and -modulating noncollagenous proteins. In bone, it has been postulated that type I collagen acts as a structural matrix, whereas HA nucleation is mediated by an anionic phosphoprotein (1–3). Of the noncollagenous proteins, bone sialoprotein (BSP)[1] is the most likely candidate. The highly anionic nature of BSP and its spatio-temporal pattern of expression have led investigators to propose a role of this protein in the mineralization of bone (1–4).

Mammalian BSPs contain an average of 327 amino acids, which includes a 16-residue signal sequence. The protein has a molecular mass of ~33–34 kDa. However, post-translational modifications, including both *N*- and *O*-linked glycosylation, tyrosine sulfation, and serine and threonine phosphorylation, constitute 50% of the total mature protein weight of ~75 kDa. Analysis of the mammalian BSP cDNAs reveals a 45% level of sequence identity, plus an additional 10–23% in conservative replacements. However, identity of up to 90% is observed in and around two polyglutamic acid sequences (poly[E]); an Arg-Gly-Asp (RGD) cell-binding motif; sites of phosphorylation, sulfation, and glycosylation; and sequences near the amino and carboxyl termini, which are rich in tyrosine residues (5).

Normally BSP expression is limited almost exclusively to mineralized connective tissues, and its expression is localized to areas of bone formation. By *in situ* hybridization, it has been shown that BSP expression occurs in osteoblasts actively engaged in bone formation and is found at low or undetectable levels in other regions of mineralized tissue (6–11). Transfection of BSP into nonmineralizing MC3T3-E1 osteoblast subclones was shown to restore their ability to form mineral deposits (12). Overexpression of BSP in a genetically engineered osteosarcoma cell line, K8, also resulted in an increase in mineral formation *in vitro* (13). It has also been observed that transgenic mice overexpressing BSP demonstrate more rapid healing of artificially induced wound sites in bone (14). On the basis of this information, BSP is likely to be involved with early mineral deposition in bone.

Using a steady-state agarose gel system, BSP was found to nucleate HA (15, 16). Synthetic homopolymers of glutamic acid have also been shown to nucleate HA, indicating that the two poly[E] regions found within BSP may play a role in this function (17). Trypsin digestion of native porcine BSP produced a series of peptides, two of which contain one of the poly[E] sequences. Each of these peptides was found to possess nucle-

* Circular dichroism experiments were carried out at the University of Western Ontario Biomolecular Interactions and Conformations Facility, which is supported by a Multi-user Equipment and Maintenance Grant from the Canadian Institutes of Health Research (CIHR). This work was supported by the CIHR and the Canadian Arthritis Network Centers of Excellence. 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.

¶ To whom correspondence should be addressed: School of Dentistry, Dental Sciences Bldg., University of Western Ontario, London, Ontario N6A 5C1, Canada. Tel.: 519-661-2182; Fax: 519-850-2316; E-mail: hagoldbe@uwo.ca.

[1] The abbreviations used are: BSP, bone sialoprotein; rBSP, recombinant BSP; HA, hydroxyapatite; poly[E], polyglutamic acid; poly[D], polyaspartic acid; poly[A], polyalanine; FPLC, fast protein liquid chromatography.

RAT BSP AMINO ACID SEQUENCE

```
MVSMKNFHRRIKAEDSEENGVFKYRPRYFLYKHAYFYPPL (40)
                        1→     ←1
KRFPVQGGSDSSEENGDGDSSEEEGEEEETSNEEENNEDS (80)

EGNEDQEAEAENATLSGVTASYGVETTADAGKLELAALQL (120)
                        2→        ←2
PKKAGDAEGKAPKMKESDEEEEEEEEEENENEEAEVDENE (160)

QVVNGTSTNSTEVDGGNGPSGGDNGEEAEEASVTEAGAEG (200)

TTAGVRELTSYGTTAVLLNGFQQTTPPPEAYGTTSPPARK (240)

SSTVEYGEEYEQIGNEYNTAYETYDENNGEPRGDTYRAYE (280)
                    3→        ←3
DEYSYYKGHGYEGYEGQDYYYHQLVPRGSHHHHH       (314)
```

FIG. 1. **Rat bone sialoprotein amino acid sequence.** *Solid underline*, first domain peptide rBSP-(43–101); *broken underline*, second domain peptide rBSP-(134–206); *1 → ← 1*, first poly[E] domain; *2 → ← 2*, second poly[E] domain; *3 → ← 3*, thrombin-cleavable His tag.

	62 – 69	139 – 148
rBSP	N..........EEEGEEEE.............EEEEEEEEEE...................C	
rBSP-pE1A	N..........AAAGAAAA.............EEEEEEEEEE...................C	
rBSP-pE2A	N..........EEEGEEEE.............AAAAAAAAAA...................C	
rBSP-pE1,2A	N..........AAAGAAAA.............AAAAAAAAAA...................C	
rBSP-pE1D	N..........DDDGDDDD.............EEEEEEEEEE...................C	
rBSP-pE2D	N..........EEEGEEEE.............DDDDDDDDDD...................C	
rBSP-pE1,2D	N..........DDDGDDDD.............DDDDDDDDDD...................C	

FIG. 2. **Full-length rat BSP mutations.** Six mutants of full-length rBSP were created by removing the wild-type polyglutamic acid (poly[E]) sequence(s) and replacing them with polyalanine (poly[A]) or polyaspartic acid (poly[D]) sequences. rBSP-pE1A and rBSP-pE1D have the first poly[E] sequence replaced by either poly[A] or a poly[D] while maintaining a glycine residue at amino acid position 65. rBSP-pE2A and rBSP-pE2D have the second poly[E] sequence replaced with poly[A] or poly[D], respectively. rBSP-pE1,2A and rBSP-pE1,2D had both poly[E] sequences altered to poly[A] and poly[D], respectively, while maintaining the glycine residue at position 65.

ating activity, whereas the tryptic peptides without a poly[E] sequence were unable to nucleate HA (18). Analysis of recombinant peptides of the two nucleating domains in porcine BSP revealed that only the peptide containing the first poly[E] sequence of porcine BSP was able to nucleate HA (19).

Knowledge of the secondary structure of BSP is limited. Based on the primary sequence BSP was believed initially to maintain an open, extended structure with regions that had the ability to form both α-helical and β-sheet structure (20). However, nuclear magnetic resonance (NMR) studies have indicated a loose open structure for a 55-residue peptide containing the RGD cell attachment sequence (21), whereas more recent studies on full-length, fully modified BSP by one-dimensional NMR also showed an unstructured, flexible conformation in solution, with no α-helical and β-sheet structure present (22). In contrast, Wuttke *et al.* (23) describe BSP as a globule linked to a thread-like structure, with 5% α-helix, 32% β-sheet, 17% β-turn, and 46% random coil. They propose that the carboxyl-terminal part of BSP, which is devoid of glycans, is globular in nature, whereas the highly glycosylated amino-terminal part of the protein is thread-like (23).

The purpose of the present study was to utilize a prokaryotic full-length rat BSP expression system to further investigate the domain(s) responsible for the HA nucleating activity of BSP. Site-directed mutagenesis of the poly[E] domains of full-length BSP was performed, and the importance of charge and conformation to the nucleating activity of BSP was examined. The conformation of the wild-type and mutated proteins was studied by circular dichroism and small angle x-ray scattering. In addition, two recombinant peptides, each encompassing one of the two poly[E] domains, were expressed and tested for nucleating activity.

EXPERIMENTAL PROCEDURES

Full-Length rBSP Plasmid Construction—The first 51 base pairs of rat cDNA encoding the signal sequence as well as the first amino acid (Phe) were removed and replaced by two vector derived amino acids (Met and Val). In addition, two contiguous arginine codons at positions 9 and 10, which are low usage in *Escherichia coli*, were silently mutated to two high usage codons using the overlapping primer extension method (24). For purification purposes, a thrombin-cleavable pentahistidine tag (Pro-Arg-Gly-Ser-His-His-His-His-His) was added to the carboxyl terminus of the cDNA. The resulting rBSP-His cDNA was then subcloned into the pET28a expression vector (Novagen). The amino acid sequence of the final construct is shown in Fig. 1.

rBSP Peptide Plasmid Construction—Two partial-length BSP polypeptides, corresponding to amino acids 3–100 and 101–314 of the rBSP-pET28 construct, were cloned by the introduction of novel restriction sites into the rat BSP sequence by Overlap Extension PCR (24) using oligonucleotide primers (Invitrogen). The PCR product was initially subcloned into pGEM-T plasmid (Promega) and subsequently subcloned into pET28a giving rise to pET28a-rBSP (3–100) and pET28a-rBSP (101–314).

Recombinant DNA procedures were carried out using methods described by Sambrook *et al.* (25). The coding sequence of all plasmids was confirmed by DNA sequencing.

Site-directed Mutagenesis of rBSP—The poly[E] domains of rBSP were mutated by introducing restriction sites as above. The restriction sites *Ava*I and *Aat*II were engineered around the first poly[E] domain (residues 62–69), and the restriction sites *Nar*I and *Acc*I were engineered around the second domain (residues 139–148). Six mutants were created by removing the wild-type polyglutamic acid sequence(s) with *Ava*I and *Aat*II or *Nar*I and *Acc*I and replacing it with an oligonucleotide containing the desired mutation. rBSP-pE1A and rBSP-pE1D replaced the first poly[E] domain with either a polyalanine (poly[A]) domain or a polyaspartic acid (poly[D]) domain while maintaining a glycine residue at amino acid position 65. rBSP-pE2A and rBSP-pE2D replaced the second poly[E] domain with poly[A] or poly[D], respectively. rBSP-pE1,2A and rBSP-pE1,2D had both poly[E] domains replaced by poly[A] and poly[D], respectively, while maintaining the glycine residue at position 65. The coding sequence of all mutants was confirmed by DNA sequencing. A diagram depicting the mutations is shown in Fig. 2.

Protein Expression and Purification—Native rat bone BSP was purified from the long bones of adult rats as described previously (26).

For recombinant proteins, *E. coli* strain BL21(DE3) cells were transformed with the expression plasmids described above and grown in PO_4-buffered Super Broth (SB) supplemented with 15 μg/ml kanamycin and 0.4% glucose to an A_{600} of 0.6–0.9. After induction with 2 mM isopropyl-β-D-thiogalactopyranoside, cultures were grown for a further 4 h. The bacterial cells were then fractionated by sonication in denaturing binding buffer (5 mM imidazole, 0.5 M NaCl, 0.02 M Tris/HCl, 6 M urea, pH 7.9). The protein extract was loaded onto Poly-prep chromatography columns (Bio-Rad) packed with 1.5 ml His-bind resin (Novagen) that had previously been charged with 50 mM $NiSO_4$. Proteins were eluted by competition with 0.5 M imidazole-containing elution buffer. Nickel affinity elution fractions were then pooled, and the full-length proteins were immediately subjected to fast protein liquid chromatography (FPLC) purification.

The two expressed and His-bond affinity-purified rBSP peptides, rBSP-(43–101) and rBSP-(134–206), were incubated with trypsin (L-1-tosylamido-2-phenylethyl chloromethyl ketone-treated) at an enzyme: substrate ratio of 1:50 in trypsin digestion buffer (50 mM Tris-HCl, 5 mM $CaCl_2$, pH 7.8) for 90 min at 37 °C. The reaction was terminated by the addition of an equal volume of Mono-Q buffer A (50 mM Tris-HCl, 7 M urea, pH 7.4). The trypsin-digested peptides were then purified by FPLC. All buffers and columns used for FPLC purification followed established protocols (26). Proteins were purified using a Q-Sepharose Fast Flow column (Fast Q) followed by size exclusion purification with a Superdex 200PG column (1.6 × 60 cm) (Amersham Biosciences). Chromatography buffers contained either 7 M urea for ion exchange

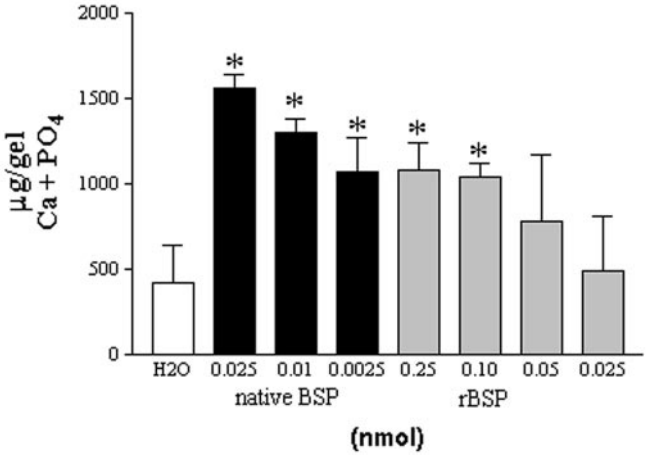

FIG. 3. **Effect of native and recombinant BSP on HA nucleation *in vitro*.** Prokaryotic recombinant BSP and native BSP extracted from rat long bones were tested for HA nucleating activity in steady-state agarose gels at concentrations of calcium (7.1 mM) and phosphate (4.3 mM) below the threshold for spontaneous precipitation. A negative control containing no protein was included. Native BSP was shown to nucleate HA at concentrations of 0.0025–0.025 nmol, whereas rBSP was shown to require a minimum concentration of between 0.05–0.10 nmol of protein to nucleate HA. *, statistical significance compared with negative control value determined by one-way analysis of variance, $p < 0.05$.

purification or 4 M urea for size exclusion purification. Protein containing fractions were analyzed using 12.5% or high density gels on the Phastgel system (Amersham Biosciences) and stained with Stains-all and silver nitrate as described previously (27). The corresponding Stains-all cyan-positive protein fractions were pooled and dialyzed using Spectra/Por 3 (cut-off, 12–14 kDa) dialysis membrane (Spectrum) for the full-length proteins and Spectra/Por 3 (cut-off, 3.5 kDa) dialysis membrane for the peptides. The fractions were then aliquoted and lyophilized. Proteins were analyzed for protein content and purity by amino acid analysis and mass spectrometry.

In Vitro Nucleation Assay—Hydroxyapatite-nucleating activity was assayed with a modification of the steady-state agarose gel system described previously (15, 16). Steady-state buffers contained either 7.1 mM of $Ca(NO_3)_2$ or 4.3 mM Na_2HPO_4. Proteins were studied in triplicate over a range of concentrations, with the lowest concentration of protein capable of nucleating HA used as a means to compare nucleating activity between the different proteins. Total mineral formation was determined by measuring calcium and phosphate contents within the gels after ashing and were expressed as weight of Ca + PO_4 per gel. Total phosphate content was quantified using the ammonium molybdate method (28), and calcium was quantified by atomic absorption spectrophotometry using a Varian SpectAA 30/40. The experimental data were compared with the negative control (no protein added) data using one-way analysis of variance.

Circular Dichroism Spectroscopy—The far-UV spectra of rBSP, rBSP-pE1,2D, rBSP-pE1,2A, rBSP-(43–101), and rBSP-(134–206) were recorded in quartz cells of 1-mm optical path length using a Jasco-J810 spectropolarimeter between 190 and 260 nm, in 0.5-nm steps. Proteins were studied at 0.2 mg/ml concentrations in 5 mM Tris-HCl, 150 mM NaCl, pH 7.4. A base line with buffer only was recorded separately and subtracted from each spectrum. The effect of calcium on the conformation of each protein was determined by the addition of 5 and 10 mM $CaCl_2$. $CaCl_2$ in buffer alone did not give measurable spectra within the 190–260 nm range. All spectra were recorded at room temperature. The molar ellipticity (θ) expressed in degrees cm^2 $dmol^{-1}$ was calculated on the basis of mean residue molecular mass.

Estimates of protein secondary structure from the CD data were made using the Dicroprot package (29), which incorporates the Contin (30), K2D (31, 32), VARSLC (33) and Selcon (34, 35) methods, as well as a calculation from $[\Phi]_{220}$ nm and a simple least squares method based on the Gauss-Jordan elimination.

Small Angle X-ray Scattering—All measurements were made at the European Molecular Biology Laboratory Outstation at the Deutsches Elektronen-Synchrotron (Hamburg, Germany), beamline X33, at 15 °C using radiation with a wavelength of 0.15 nm and a path length of 1 mm. A sample detector distance of 3 m (low angle) was used to cover the

TABLE I
Effects of mutations in the polyglutamic acid regions of BSP on HA-nucleation in vitro

The minimum concentration of protein required to nucleate HA is shown.

Protein	C	C
	µg/ml	*nmol*
Native BSP	0.087	0.0025
rBSP	1.7–3.4	0.05–0.1
rBSP-pE1A	3	0.087
rBSP-pE2A	3	0.087
rBSP-pE1,2A	17.5	0.5
rBSP-pE1D	3	0.087
rBSP-pE2D	3	0.087
rBSP-pE1,2D	3.4	0.1

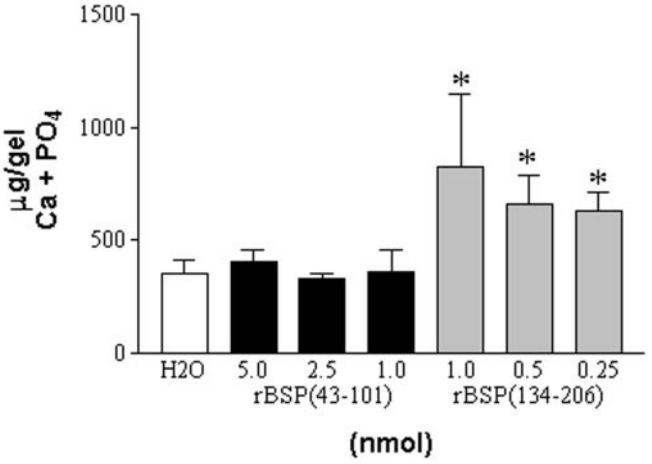

FIG. 4. **Effect of rBSP peptides on HA nucleation *in vitro*.** rBSP-(43–101), a peptide encompassing the first Glu-rich domain of BSP, and rBSP-(134–206), a peptide incorporating the second Glu-rich domain, were tested for their ability to nucleate HA at subthreshold concentrations of calcium (7.1 mM) and phosphate (4.3 mM). A negative control containing no protein was included. The rBSP-(43–101) peptide did not have HA nucleating activity at 1.0–5.0 nmol. The rBSP-(134–206) peptide induced nucleation at a minimum concentration of 0.25 nmol. *, statistical significance compared with negative control value determined by one-way analysis of variance, $p < 0.05$.

range of momentum transfer (S)2 sin θ/λ, where 2 θ is the scattering angle) from 0.003 to 0.085 Å-1. Fifteen successive 1-min exposures were recorded for each sample. Each protein sample was preceded and followed by recording of the buffer alone to ensure the cleanliness of the cell between readings of protein solutions. Averaging of frames, corrections for detector response and beam intensity, and buffer subtraction were done using the programs SAPOKO[2] and OTOKO (36). Protein samples were dialyzed against 5 mM Tris-HCl, 150 mM NaCl, pH 7.0. Samples were run in the presence and absence of 5 mM $CaCl_2$. Protein concentrations were rBSP, 4.7 mg/ml, and rBSP-pE1,2A, 7.4 mg/ml.

RESULTS

Expression and Purification of Recombinant BSP—Rat BSP was expressed in *E. coli* BL21(DE3) cells and purified to an apparent 99%+ purity. The recombinant rat BSP with the thrombin cleavage site and 5×His tag was detected by electrospray mass spectrometry as a single peak at 34802 Da, corresponding to the theoretical mass of 34796 Da based on amino acid composition. SDS-PAGE showed a single band at 67 kDa. The mutant proteins and recombinant peptides were similarly assessed for purity by SDS-PAGE, amino acid analysis, and mass spectrometry (data not shown).

Effects of Full-length rBSP on HA Nucleation—To determine whether the bacterially expressed BSP was capable of nucleating HA and to examine the involvement of post-translational

[2] D. I. Svergun and M. H. J. Koch, unpublished observations.

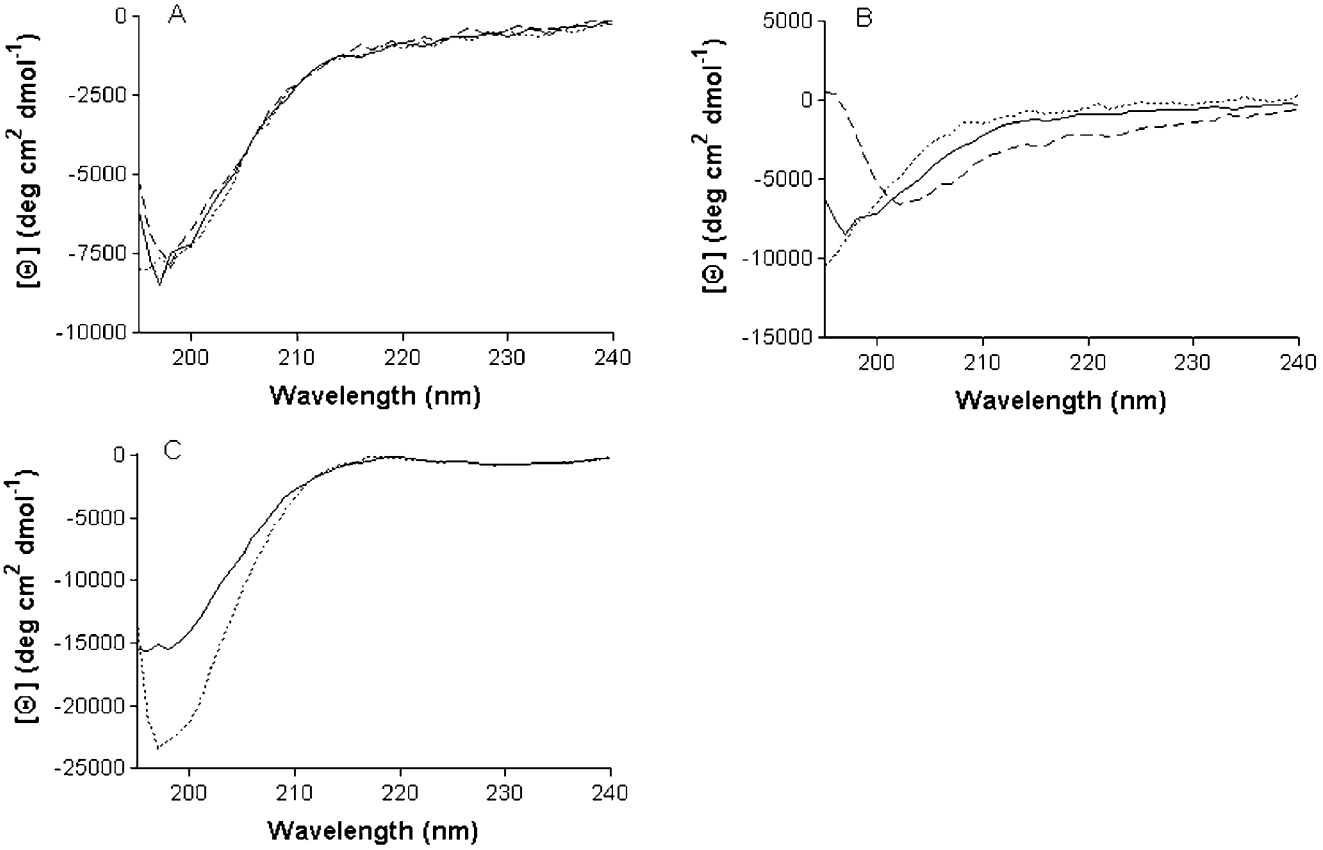

FIG. 5. **Circular dichroism spectra of rBSP peptides and mutants.** All proteins were studied at 0.2 mg/ml in 5 mM Tris-HCl, 150 mM NaCl. *A*, rBSP with 0 mM $CaCl_2$ (——), 5 mM $CaCl_2$ (- - - –), or 10 mM $CaCl_2$ (-·-·-). *B*, rBSP (——), rBSP-pE1,2D (– – –), and rBSP-pE1,2A with no calcium added (-·-·-). *C*, rBSP-(43–101) (——) and rBSP-(134–206) with no calcium added (– – –).

modifications in HA nucleation, various concentrations of native and recombinant BSP were incorporated into steady-state agarose gels, and the lowest level of protein required to induce nucleation was determined (Fig. 3). Although both native and recombinant forms of BSP were capable of nucleating HA, native BSP was found to nucleate HA at concentrations as low as 0.0025 nmol (0.087 μg/ml), whereas rBSP required 0.05–0.1 nmol (1.7–3.4 μg/ml).

To examine the contribution that each domain makes to the nucleating activity of full-length rBSP, the single-domain mutants (rBSP-pE1D, rBSP-pE2D, rBSP-pE1A, and rBSP-pE2A) were tested at various concentrations to determine the minimum concentration required for HA nucleation (Table I). Each of these single domain mutants, as well as the double domain mutant rBSP-pE1,2D, was found to have nucleating activity at ~0.1 nmol, comparable with unmutated rBSP (Table I). A decrease in activity was seen, however, with rBSP-pE1,2A, where both domains were mutated to poly[A]. rBSP-pE1,2A required 0.5 nmol of protein (Table I).

To more clearly establish the nucleating activity of each poly[E] domain, two peptides were expressed, each incorporating one of the Glu-rich domains. The rBSP-(134–206) peptide was shown to have nucleating activity at a concentration of 0.25 nmol; however, the rBSP-(43–101) peptide did not promote nucleation at concentrations as high as 5 nmol (Fig. 4). This finding is consistent with our previous study on porcine BSP peptides, which showed different activities for each domain.

Structural Determination—The CD spectrum of rBSP at pH 7.4 is shown in Fig. 5*A*. It has a pattern typical of an unfolded protein, as generally observed by a denatured protein in urea or guanidine hydrochloride. Because of the strong electrostatic repulsion of the high degree of negatively charged residues, 5 and 10 mM $CaCl_2$ were added to determine whether rBSP could be transformed into a partially folded conformation (Fig. 5*A*). The presence of calcium did not significantly alter the spectrum of rBSP.

The two full-length, double domain mutants (rBSP-pE1,2A and rBSP-pE1,2D) were also studied by CD (Fig. 5*B*). Although the rBSP and rBSP-pE1,2D spectra were different in intensity, they exhibited similar shape. The rBSP-pE1,2A, however, appears to have a different conformation; most noticeably, the minimum is shifted from ~195 nm for rBSP and rBSP-pE1,2D to ~204 nm for rBSP-pE1,2A. The addition of 5 and 10 mM $CaCl_2$ did not effect the spectra of either rBSP-pE1,2A or rBSP-pE1,2D (data not shown).

The CD spectra (Fig. 5*C*) of the two peptides, rBSP-(43–101) and rBSP-(134–206), although differing in intensity, were similar in shape and were unchanged with the addition of 5 and 10 mM $CaCl_2$ (data not shown).

All CD spectra were analyzed for estimates of secondary structure using the Dicroprot program. Only the rBSP-pE1,2A mutant exhibited a spectrum similar enough to those found in the data bases for secondary structure estimates to be given. It is estimated that this protein exhibits 13.8% α-helix, 43.5% β-sheet, 10.8% β-turn, 8.1% poly-pro (II) helix, and 18.9% unordered structure.

Small angle x-ray scattering (SAXS) is a useful tool for investigating the conformation, shape, and dimensions of proteins in solution. The degree of unfolding of a protein is best viewed with a Kratky plot, $S^2I(S)$ *versus* (S is scattering wave vector; $I(S)$ is scattering profile intensity), which emphasizes the high angle part of the scattering profile (37). Globular proteins scatter as S^{-4} at high S values, yielding a Kratky plot that is proportional to S^{-2} (38), whereas a random coil scatters

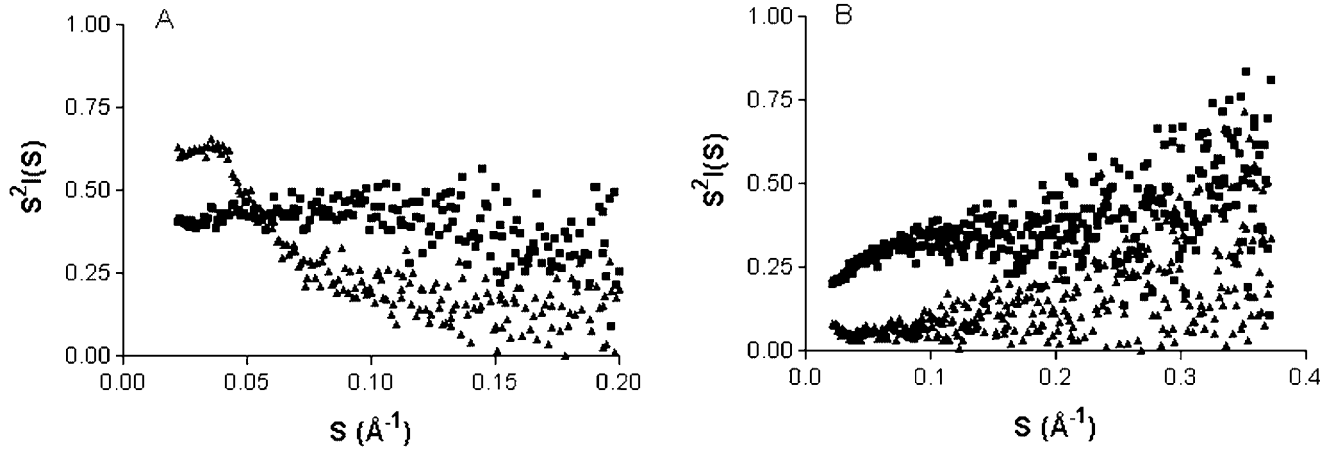

FIG. 6. **Kratky plots of scattering curves from rBSP and rBSP-pE1,2A.** *A*, rBSP at pH 7 (■) and pH 4 (▲). *B*, rBSP-pE1,2A at pH 7 (■) and pH 4 (▲).

as S^{-1} at high angles, yielding a plot with S^1 dependence (39). Therefore, a Kratky plot of a native globular protein will have a characteristic maximum that is dependent on its radius of gyration (R_g), whereas unfolded and partially folded proteins give a plateau and then rise (37). Analysis of the x-ray scattering of rBSP and rBSP-pE1,2A in the form of a Kratky plot shows that neither has a well developed globular structure under any condition studied (Fig. 6). The profile of the Kratky plots for both proteins at neutral pH, pH 4, and in the presence of 5 mM $CaCl_2$ (data not shown) is typical of a random coil.

DISCUSSION

Studies on the structure and function of BSP have been hindered by difficulties in protein expression. Although full-length BSP has been successfully expressed eukaryotically, expression of full-length prokaryotic BSP has been problematic (40). A successful expression system for full-length prokaryotic rat BSP was established recently (41, 42). This expression system was used to investigate the domains responsible for the nucleating activity of BSP and additionally to examine the contribution of charge and conformation to this nucleating activity.

Using the bacterially expressed full-length rat rBSP, the nucleating activities of native BSP and rBSP were compared. Although the rBSP was capable of nucleating HA, it required approximately a 40-fold increase in protein concentration over native BSP. This decrease in activity is in agreement with previous studies on porcine peptides derived from both native bone extracts and prokaryotic expression, which indicated that although post-translational modifications are not necessary for nucleation, lack of these modifications does decrease the potency of the protein (19). In this study it was shown that a native porcine BSP peptide encompassing residues 42–125 was capable of nucleating HA at concentrations as low as 0.3 nmol, whereas the recombinant form of the peptide required 1.6 nmol or higher to nucleate HA. The exact role of the post-translational modifications on BSP in HA nucleation remains unclear. Although it has been proposed that the post-translational modifications on BSP may stabilize a particular conformation for the protein, studies have shown that at least for the amino-terminal half of the eukaryotically expressed BSP, there is no defined secondary structure (22, 23). However, it is possible that the phosphate groups on the post-translationally modified BSP may provide a higher charge density, which allows for calcium ion accumulation at lower protein concentrations. Future studies utilizing eukaryotically expressed BSP will examine the role(s) the post-translational modifications may have in HA nucleation.

Because chemical modification of the carboxylate groups has been shown to abolish the nucleation activity of BSP, and synthetic homopolymers of glutamic acid have demonstrated HA nucleating activity (17), the nucleation activity of BSP is believed to reside primarily in two poly[E] domains found in the amino-terminal half of the molecule. This belief has also been reinforced through the investigation of poly[E]-containing tryptic peptides and recombinant peptides of BSP (18, 19). To further define the contribution of the poly[E] sequences to nucleation, either one or both of the poly[E] domains were replaced with either poly[D] or poly[A]. Poly[D] was chosen because aspartic acid residues possess the same charge as glutamic acid, and poly[A] was chosen because it shares the same propensity as poly[E] to form a helical structure (43). To alleviate any anomalies in which one domain may be compensating for the loss of activity in the other domain, two peptides, rBSP-(43–101) and rBSP-(134–206), were expressed so that each domain could be studied independently.

All four single poly[E] domain full-length mutants were found to possess similar nucleation activities to rBSP, because nucleation activity was observed at a concentration of ~0.1 nmol (Table I). No variation was observed between the poly[A] and poly[D] mutants. These analyses indicate that both poly[E] domains in rBSP appear to have distinct and similar nucleating activity. These results are in contrast to what is seen with the individual peptides. In this study, it was found that the first domain peptide, rBSP-(43–101) is inactive at concentrations as high as 5.0 nmol, whereas the second domain peptide, rBSP-(134–206), is almost as effective a nucleator as full-length rBSP, showing activity at 0.25 nmol (Fig. 4). It is interesting that in previous studies using recombinant porcine BSP, the reverse was found to be true. The peptide encompassing the first domain was found to be active, and the second domain peptide was inactive (19). Analysis of the amino acid sequences for these two species reveals that the first poly[E] domains are quite similar (rat, EEEGEEEE; porcine, EEEEEEEE), with the only difference being the presence of a glycine residue in the middle of the rat sequence. The second domain exhibits a similar trend (rat, EEEEEEEEEE; porcine, EDEEEEEENEE); in this case an aspartic acid and an asparagine residue disrupt the poly[E] sequence. From these sequences, it may be hypothesized that the nucleating activity of BSP requires a minimum number of contiguous glutamic acid residues.

When both poly[E] domains of rBSP were simultaneously mutated (rBSP-pE1,2D and rBSP-pE1,2A), it was found that nucleating activity was retained, although a higher concentration of protein was required for the poly[A] mutant compared

with unmutated rBSP or poly[D] mutant (Table I). Aspartic acid and alanine residues were used in the mutagenesis to deduce the relative contribution of charge and structure to the nucleating activity of BSP. Although synthetic homopolymers of poly[D] have been shown not to possess nucleating activity and have in fact been shown to inhibit HA formation (44, 45), other nucleating proteins such as dentin phosphophoryn (16) and an aspartic acid-rich protein from mollusk shell (46, 47) have been shown to nucleate HA and calcium carbonate crystals, respectively, via their aspartic acid-rich sequences. It is therefore not surprising that the poly[D] BSP mutant is capable of nucleating HA. Poly[D] is not known, however, to adopt a helical conformation but has a more random coil conformation (18) and can, under acidic conditions (or in presence of calcium), form a β-sheet. If a helical conformation were required for nucleation, the poly[A] mutant would likely be a more potent nucleator of HA (*i.e.* require lower protein concentration) than the poly[D] mutant. This, however, was not observed, and it appears that the charge of the contiguous domains, rather than ordered secondary structure, is critical in maintaining the nucleating activity of BSP.

The fact that some nucleating activity is still evident after both domains have been mutated to poly[A] indicates that perhaps another region or domain may also be involved in nucleation. This notion is supported by HA binding studies, which have shown that BSP binds to HA, at least in part, via the contiguous poly[E] sequences. Synthetic homopolymers of poly[E], however, did not completely inhibit this binding (48). It was therefore suggested that additional domain(s), or specific conformational motifs on BSP, are involved in HA binding and perhaps in HA nucleation. Examination of the rat BSP amino acid sequence reveals that there are a number of glutamic acid residues downstream of each of the poly[E] domains (first, EEENNEDSEGNEDQEAEAE; second, ENEEAEVDENE). Although these glutamic acid-rich regions do not appear to possess independent nucleating activity, as evidenced by the lack of activity shown by the rat first domain peptide and the porcine second domain peptide (19), these regions may be working cooperatively in the altered full-length rBSP sequence to compensate for the mutation of both of the poly[E] domains.

Previous structural studies on native and eukaryotic, fully modified BSP have been conflicting. Fisher *et al.* (22) found by one-dimensional proton NMR that BSP was flexible along its entire length with no significant structural regions. Structural studies on native and eukaryotic BSP by Wuttke *et al.* (23), however, showed that BSP is 5% α-helix, 32% β-sheet, 17% β-turn, and 46% random coil; they proposed that the carboxyl-terminal part of the protein, which is free of glycans, forms a globular structure, whereas the highly glycosylated regions are thread-like due to lack of secondary structure. Here, we have studied the conformation of full-length rBSP, as well as the double-domain mutants (rBSP-pE1,2D and rBSP-pE1,2A), by circular dichroism and small angle x-ray scattering.

The three proteins were studied in the presence of 5 and 10 mM calcium chloride in an attempt to mimic *in vivo* conditions whereby calcium ions may induce folding of the protein by neutralizing the negative charges. The calcium did not appear to have any effect on protein folding and no ordered structure was evident.

Although secondary structure predictions were attempted on the CD spectra of all of the proteins, predictions were possible only for the rBSP-pE1,2A mutant. The difficulty in estimating secondary structure elements for the BSP proteins is due to the differences in our spectra and those found in the program data bases. These programs, along with secondary structure prediction programs based on amino acid sequence, are biased toward globular proteins and tend to be inaccurate for more extended, unordered proteins and peptides, often over-estimating ordered conformations for these proteins. The CONTIN program used by Wuttke *et al.* (23) is one of the better programs for estimating α-helix, β-sheet, and β-turn conformations from CD data. However, the method suffers from the choice of proteins in the data base of standards (49). The inclusion of denatured proteins in the data base has been shown to significantly improve the estimates for unordered proteins (50); however, these have not been included in these programs at this time. Even the structural estimates obtained for the rBSP-pE1,2A mutant appear to be inflated when considering the small angle x-ray scattering data. A protein with such high secondary structure content would likely appear as more of a globular protein in the Kratky plot. The rBSP and rBSP-pE1,2A proteins, however, are both lacking the typical maximum, which is indicative of its R_g, and show the characteristic plateau of unordered, unfolded proteins.

If it is only the poly[E] and downstream glutamic acid-rich regions that adopt an α-helical conformation, it may be possible that these regions are being masked by the unstructured nature of the bulk of the protein. Because the poly[E] and glutamic acid-rich regions make up a large part of the peptides, it should be possible to observe any α-helical conformation by CD. Although the rBSP-(134–206) peptide is predicted by Consensus Secondary Structure Prediction (51) to be 37% helical, and the rBSP-(43–101) is predicted to have 22% helical character, virtually no α-helical content was detected even in the presence of calcium. Although a small difference in the shape of the two spectra was seen, it was difficult to determine how great the structural differences in these two proteins may be without any reliable means of ascertaining their structure. The fact that no ordered secondary structure was observed by CD between the nucleating rBSP-(134–206) peptide and the nonnucleating rBSP-(43–101) peptide indicates once again that ordered secondary structure does not appear to play an important role in the nucleating activity of BSP.

Normally, proteins need a specific three-dimensional structure to perform their particular function. In the case of BSP, however, it has been suggested that the unstructured, flexible nature of BSP may be advantageous to its function (22). The amino-terminal half of BSP is known to have a strong affinity for HA (48, 52, 53), whereas a carboxyl-terminal RGD sequence allows for binding to $\alpha_v\beta_3$ integrin (19, 21, 54). The flexibility seen with BSP would thus allow it to serve as a bridge for attachment of cells to HA. In the case of mineralization, the flexibility of the protein may allow binding to type I collagen in the hole zones, as proposed by Fujisawa *et al.* (55), followed by initiation of mineral formation. Ongoing studies on the localization of the collagen-binding domain(s) of BSP and nucleation studies within collagen gels will provide further insight into this idea.

In conclusion, we have used a prokaryotic expression system to produce full-length recombinant rat BSP and have created mutations of desired domains to investigate its function. We have also tested the hypotheses that the polyglutamic acid domains of BSP are responsible for the nucleation of HA and that these domains adopt an α-helical conformation. Our results have revealed that a sequence of at least eight contiguous glutamic acid residues followed by a glutamic acid-rich region may be requisite for the nucleation of HA by wild-type BSP and that this region is not α-helical in conformation. Our secondary structure results are in agreement with those in the current literature and support the idea that the highly flexible nature of BSP may be advantageous to its function.

Acknowledgments—We thank T. M. Underhill, B. S. Shilton, and D. M. Currie for insights and assistance.

REFERENCES

1. Gorski, J. P. (1992) *Calcif. Tissue Int.* **50,** 391–396
2. Glimcher, M. J. (1989) *Anat. Rec.* **224,** 139–153
3. Veis, A. (1985) in *The Chemistry and Biology of Mineralized Tissue* (Butler, W., ed) pp. 170–177, Ebsco Press, Birmingham, AL
4. Sodek, J., Chen, J., Kasugai, S., Nagata, T., Zhang, Q., Shapiro, H. S., Wrana, J. L., and Goldberg, H. A. (1992) in *Biological Mechanisms of Tooth Movement and Craniofacial Adaption* (Davidovitch, Z., ed) pp. 127–136, Ohio University Press, Columbus, OH
5. Ganss, B., Kim, R. H., and Sodek, J. (1999) *Crit. Rev. Oral Biol. Med.* **10,** 79–98
6. Bianco, P., Fisher, L. W., Young, M. F., Termine, J. D., and Robey, P. G. (1991) *Calcif. Tissue Int.* **49,** 421–426
7. Bianco, P., Riminucci, M., Silvestrini, G., Bonucci, E., Termine, J. D., Fisher, L. W., and Robey, P. G. (1993) *J. Histochem. Cytochem.* **41,** 193–203
8. Chen, J., McKee, M. D., Nanci, A., and Sodek, J. (1994) *Histochem. J.* **26,** 67–78
9. Chen, J., Singh, K., Mukherjee, B. B., and Sodek, J. (1993) *Matrix.* **13,** 113–123
10. Hultenby, K., Reinholt, F. P., Norgard, M., Oldberg, A., Wendel, M., and Heinegard, D. (1994) *Eur. J. Cell Biol.* **63,** 230–239
11. Riminucci, M., Silvestrini, G., Bonucci, E., Fisher, L. W., Gehron Robey, P., and Bianco, P. (1995) *Calcif. Tissue Int.* **57,** 277–284
12. Wang, D., Christensen, K., Chawla, K., Xiao, G., Krebshach, P. H., and Franceschi, R. T. (1999) *J. Bone Miner. Res.* **14,** 893–903
13. Gerstenfeld, L. C., Shih, S., George, C., Mizuno, S., and Glowacki, J. (2000) in *Chemistry and Biology of Mineralized Tissues: Proceedings of the Sixth International Conference* (Goldberg, M., Boskey, A., and Robinson, C., eds) pp. 143–48, American Academy of Orthopaedic Surgery, Rosemont, IL
14. Thomas, H. F., Jiang, H., and Chen, J. (1999) *J. Dent. Res.* **78 SI**, 455 (abstr.)
15. Hunter, G. K., and Goldberg, H. A. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 8562–8565
16. Hunter, G. K., Hauschka, P. V., Poole, A. R., Rosenberg, L. C., and Goldberg, H. A. (1996) *Biochem. J.* **317,** 59–64
17. Hunter, G. K., and Goldberg, H. A. (1994) *Biochem. J.* **302,** 175–179
18. Goldberg, H. A., Warner, K. J., Stillman, M. J., and Hunter, G. K. (1996) *Connect. Tissue Res.* **35,** 385–392
19. Harris, N. L., Rattray, K. R., Tye, C. E., Underhill, T. M., Somerman, M. J., D'Errico, J. A., Chambers, A. F., Hunter, G. K., and Goldberg, H. A. (2000) *Bone* **27,** 795–802
20. Shapiro, H. S., Chen, J., Wrana, J. L., Zhang, Q., Blum, M., and Sodek, J. (1993) *Matrix* **13,** 431–440
21. Stubbs, J. T., III, Mintz, K. P., Eanes, E. D., Torchia, D. A., and Fisher, L. W. (1997) *J. Bone Miner. Res.* **12,** 1210–1222
22. Fisher, L. W., Torchia, D. A., Fohr, B., Young, M. F., and Fedarko, N. S. (2001) *Biochem. Biophys. Res. Commun.* **280,** 460–465
23. Wuttke, M., Muller, S., Nitsche, D. P., Paulsson, M., Hanisch, F. G., and Maurer, P. (2001) *J. Biol. Chem.* **276,** 36839–36848
24. Pogulis, R. J., Vallejo, A. N., and Pease, L. R. (1996) *Methods Mol. Biol.* **57,** 167–176
25. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning*: *A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
26. Goldberg, H. A., and Sodek, J. (1994) *J. Tissue Cult. Methods* **16,** 211–215
27. Goldberg, H. A., and Warner, K. J. (1997) *Anal. Biochem.* **251,** 227–233
28. Chen, P. S., Toribara, T. Y., and Warner, H. (1956) *Anal. Chem.* **28,** 1756–1758
29. Deléage, G. G. C. (1993) *Comput. Appl. Biosci.* **9,** 197–199
30. Provencher, S. W., and Glockner, J. (1981) *Biochemistry* **20,** 33–37
31. Merelo, J. J., Andrade, M. A., Prieto, A., and Morán, F. (1994) *Neurocomputing* **6,** 443–454
32. Andrade, M. A., Chacon, P., Merelo, J. J., and Moran, F. (1993) *Protein Eng.* **6,** 383–390
33. Manavalan, P., and Johnson, W. C., Jr. (1987) *Anal. Biochem.* **167,** 76–85
34. Sreerama, N., and Woody, R. W. (1994) *J. Mol. Biol.* **242,** 497–507
35. Sreerama, N., and Woody, R. W. (1994) *Biochemistry* **33,** 10022–10025
36. Boulin, C., Kempf, R., Koch, M. H. J., and McLaughlin, S. (1986) *Nucl. Instrum. Methods Phys. Res. A* **249,** 399–407
37. Doniach, S. (2001) *Chem. Rev.* **101,** 1763–1778
38. Porod, G. (1951) *Kolloid-Z* **124,** 83
39. Kirste, R. G., and Oberthür, R. C. (1982) in *Small Angle X-ray Scattering* (Glatter, O., and Kratky, O., eds) pp. 387–431, Academic Press, London, England
40. Stubbs, J. T., III (1996) *Connect. Tissue Res.* **35,** 393–399
41. Rattray, K. R., Currie, D. M., Warner, K. J., Hunter, G. K., and Goldberg, H. A. (2000) *J. Dent. Res.* **79,** 520
42. Razzouk, S., Brunn, J. C., Qin, C., Tye, C. E., Goldberg, H. A., and Butler, W. T. (2002) *Bone* **30,** 40–47
43. Tiffany, M. L., and Krimm, S. (1968) *Biopolymers* **6,** 1379–1382
44. Hunter, G. K., Kyle, C. L., and Goldberg, H. A. (1994) *Biochem. J.* **300,** 723–728
45. Goldberg, H. A., and Hunter, G. K. (1995) *Ann. N. Y. Acad. Sci.* **760,** 305–308
46. Worms, D., and Weiner, S. (1986) *J. Exp. Zool.* **237,** 11–20
47. Addadi, L., and Weiner, S. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82,** 4110–4114
48. Goldberg, H. A., Warner, K. J., Li, M. C., and Hunter, G. K. (2001) *Connect. Tissue Res.* **42,** 25–37
49. Greenfield, N. J. (1996) *Anal. Biochem.* **235,** 1–10
50. Venyaminov, S., Baikalov, I. A., Shen, Z. M., Wu, C. S., and Yang, J. T. (1993) *Anal. Biochem.* **214,** 17–24
51. Deleage, G., Blanchet, C., and Geourjon, C. (1997) *Biochimie* (*Paris*) **79,** 681–686
52. Fujisawa, R., Mizuno, M., Nodasaka, Y., and Kuboki, Y. (1997) *Matrix Biol.* **16,** 21–28
53. Fujisawa, R., Wada, Y., Nodasaka, Y., and Kuboki, Y. (1996) *Biochim. Biophys. Acta* **1292,** 53–60
54. Flores, M. E., Norgard, M., Heinegard, D., Reinholt, F. P., and Andersson, G. (1992) *Exp. Cell Res.* **201,** 526–530
55. Fujisawa, R., Nodasaka, Y., and Kuboki, Y. (1995) *Calcif. Tissue Int.* **56,** 140–144