Sequence preservation of osteocalcin protein and mitochondrial DNA in bison bones older than 55 ka

Christina M. Nielsen-Marsh* Fossil Fuels and Environmental Geochemistry, Newcastle Research Group, Drummond Building, University of Newcastle, Newcastle-upon-Tyne NE1 7RU, UK

Peggy H. Ostrom*
Hasand Gandhi*] Department of Geological Sciences, Michigan State University, East Lansing, Michigan 48824, USA

Beth Shapiro*
Alan Cooper*] Department of Zoology, University of Oxford, Oxford OX1 3PS, UK

Peter V. Hauschka* Harvard Schools of Medicine and Dental Medicine, Children's Hospital, Boston, Massachusetts 02115, USA

Matthew J. Collins* Fossil Fuels and Environmental Geochemistry, Newcastle Research Group, Drummond Building, University of Newcastle, Newcastle-upon-Tyne NE1 7RU, UK

ABSTRACT

We report the first complete sequences of the protein osteocalcin from small amounts (20 mg) of two bison bone (*Bison priscus*) dated to older than 55.6 ka and older than 58.9 ka. Osteocalcin was purified using new gravity columns (never exposed to protein) followed by microbore reversed-phase high-performance liquid chromatography. Sequencing of osteocalcin employed two methods of matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS): peptide mass mapping (PMM) and post-source decay (PSD). The PMM shows that ancient and modern bison osteocalcin have the same mass to charge (*m*/*z*) distribution, indicating an identical protein sequence and absence of diagenetic products. This was confirmed by PSD of the *m*/*z* 2066 tryptic peptide (residues 1–19); the mass spectra from ancient and modern peptides were identical. The 129 mass unit difference in the molecular ion between cow (*Bos taurus*) and bison is caused by a single amino-acid substitution between the taxa (Trp in cow is replaced by Gly in bison at residue 5). Bison mitochondrial control region DNA sequences were obtained from the older than 55.6 ka fossil. These results suggest that DNA and protein sequences can be used to directly investigate molecular phylogenies over a considerable time period, the absolute limit of which is yet to be determined.

Keywords: proteins, mass spectrometry, DNA, preservation, fossils, bison.

INTRODUCTION

Although most ancient DNA studies are confined to the past 50 k.y. (Wayne et al., 1999), significant interest exists in extending molecular records farther back in time (Pääbo, 2000). Certain mineral-associated proteins demonstrate marked stability and may offer a substrate that could extend the genetic record beyond that of DNA (Hauschka, 1980; Collins et al., 2000; Ostrom et al., 2000). Partial protein sequences from fossil shells (Robbins and Donachy, 1991) and bone (Huq et al., 1990) are encouraging, but direct sequencing by Edman degradation is hindered by amino-terminal blockage and the difficulty in preparing large quantities of purified extract (Robbins et al., 1993). Consequently, this reservoir of ancient genetic information remains largely untapped.

*E-mail addresses: Nielsen-Marsh—nielse42@msu.edu; Ostrom—Ostrom@msu.edu; Gandhi—Gandhiha@msu.edu; Shapiro—Beth.shapiro@zoology.oxford.ac.uk; Cooper—Alan.cooper@zoo.ox.ac.uk; Hauschka—peter.hauschka@tch.harvard.edu; Collins—M.collins@ncl.ac.uk. Corresponding author: present address: Christina M. Nielsen-Marsh, Department of Geological Sciences, Michigan State University, East Lansing, Michigan 48824, USA.

Here we report the first complete sequences of the protein osteocalcin from two fossil bison. The presence of DNA was also confirmed in one of the fossil samples by the sequencing of amplified mitochondrial DNA (mtDNA). Osteocalcin is a small extracellular bone matrix protein containing three γ-carboxyglutamic acid (Gla) residues at positions 17, 21, and 24, resulting in a high affinity for bone mineral (Hauschka and Carr, 1982; Hauschka et al., 1989). Osteocalcin is also exclusive to vertebrates, eliminating possible contamination by microbes, invertebrates, or plants (Hauschka, 1980; Muyzer et al., 1992). Our procedure involves purification using new gravity columns (never exposed to protein), followed by microbore reversed-phase high-performance liquid chromatography (rp-HPLC) (Schall et al., 1999). Sequencing is accomplished with matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS), which is unaffected by amino terminal blockage (Ostrom et al., 2000) and requires only 20 mg of bone. Screening by MALDI-MS for putative osteocalcin after partial purification with new gravity columns reduces the possibility of contamination.

METHODS

Two *Bison priscus* permafrost fossils were analyzed: one from Stanchikovsky Yar, Siberia (Kolyma River region) (CRS-SY-2), ^{14}C dated to older than 55.6 ka (Oxford Radiocarbon Laboratory OxA-10575), and the other from Eva Creek, Ester, Alaska, ^{14}C dated to older than 58.9 ka (Oxford Radiocarbon Laboratory OxP-10647).

Protein extraction and purification methods are modified from Schall et al. (1999). Bone fragments were cleaned (Dremel drill), powdered (SPEX CertiPrep 6750 Freezer/Mill), and 20 mg of powder demineralized (0.5 *N* sodium EDTA, pH 8.0, 4 h, 25 °C). The extract was centrifuged (13,000 rpm, 5 min) twice and the supernatant introduced to a new C18 gravity column (60 Å pore size, Fisher) equilibrated with solvent A (A) [950/49.9/0.1, water/acetonitrile (ACN)/trifluoroacetic acid (TFA) (v/v/v)] and eluted with solvent B (B) [900/9.9/0.1 ACN/water/TFA (v/v/v)] in 8, 1000 μL aliquots [20%, 24%, 28%, 31%, 32%, 33%, 34%, 36% B in A (v/v)]. These fractions were dried (SpeedVac) and reconstituted in 10 μL of 1% OGP (*n*-octyl-β-D-glucopyranoside) in 50 m*M* Tris buffer. Aliquots (0.5 μL) were placed on the MALDI target with 0.1% TFA (0.5 μL) and 0.5 μL saturated 4-hydroxy-α-cyanocinnamic acid (4-HCCA) solution (1:1 ACN/0.1% TFA). MALDI-MS spectra were acquired to identify fractions containing osteocalcin (31% and 32% B fractions).

Fractions containing osteocalcin were combined, diluted in A, and purified by rp-HPLC. An aliquot (~150 μL) was injected onto a peptide trap (1 × 10 mm, Michrom BioResources) equilibrated with 20% B. The trap was washed with 600 μL of A and its flow introduced to a 1 × 150 mm C18 column (300 Å, 5 μm, Reliasil, Michrom BioResources). The column and trap were maintained at 25 °C during the gradient: 20% B, changing to 25% B over 15 min, held for 5 min, then increased to 30% B over 15 min, held for 10 min and then increased to 95% over 5 min, then held for 5 min. The peak collected was

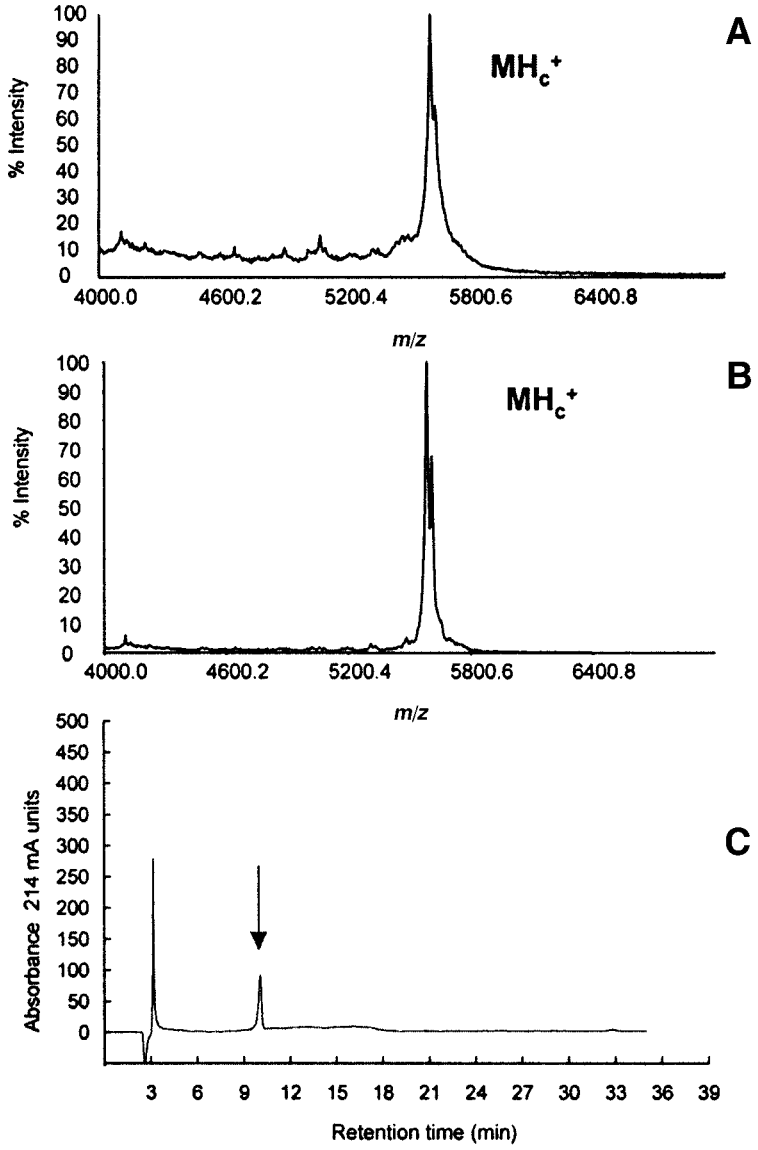

Figure 1. Matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) and reversed-phase high-performance liquid chromatography (rp-HPLC) data for bone extracts. High-mass regions of MALDI-MS (4-HCCA) showing predominant peak, *m/z* 5590 (MH^+), from gravity column eluent for (A) modern bison and (B) older than 55.6 ka bison. C: rp-HPLC of older than 58.9 ka bison peak extract eluted from gravity column; arrow indicates ~11 min osteocalcin peak.

further purified by reinjection into the rp-HPLC. MALDI-MS of blank runs showed no evidence of osteocalcin. The rp-HPLC columns used for fossil extracts were never used for modern osteocalcin.

For peptide mass mapping (PMM), rp-HPLC fractions were dried (SpeedVac) and digested with trypsin (4 μL, 1% OGP in 50 m*M* Tris/chloride-buffer, pH 8.0; 1 μL trypsin, 10 ng/μL) (Sigma) for 24 h at 37 °C. The digest was injected into the peptide trap and washed as before. The rp-HPLC gradient was: 5% B to 40% B over 30 min, held for 10 min, increased to 95% B over 1 min, held for 5 min, then returned to 5% B over 1 min. The putative osteocalcin precursor peak was identified by post-source decay (PSD) analysis.

MALDI-MS was performed on a PerSeptive Biosystems DE-STR MALDI-TOF-MS. A nitrogen laser (337 nm, 3 ns pulse width, acceleration voltage of 22 kV) was applied in delayed extraction mode (50 ns delay). Mass spectra were externally calibrated. Samples (0.5 μL) were prepared with 0.1% TFA (0.5 μL) and matrix (0.5 μL) [either 4-HCCA, or 0.2% 6-Aza-2-thiothymine (ATT) (1:1 ACN/water)].

DNA from the Siberian sample was extracted, amplified, and sequenced (Barnes et al., 2002). A 377–base pair (bp) sequence of mitochondrial control region DNA was amplified in two overlapping fragments by using primers Bis-CR-16633f (forward) (GCCCCATG CATATAAGCAAG), BisCR-16810r (reverse) (GCCTAGCGGGTTGCTGGTTTCACGC), BisCR-16765f (CACGAGCTTAACTACCAT GC), and BisCR-94r (GGCCATAGCTGA GTCCAAGC). Numbering is relative to the 3′ base in the cow mtDNA sequence NC001567. One fragment was cloned, and the combined consensus sequence was compared with modern bison (*Bison bison*) and cow (*Bos Taurus*) sequences from GenBank. The sequence was deposited in GenBank with the accession number AF538947.

Radioimmunoassay (RIA) was performed on heated bone powder from modern cow (*B. taurus*) and two fossil bison (*B. priscus*). The fossils were biostratigraphically dated to 120 ka (BMNH #49690; London, UK) and 300 ka (LA#1616; Roaix, France). RIA utilized a nonequilibrium assay (Ostrom et al., 2000). Bone powder (50–200 mg) was extracted (16 h, 4 °C; 0.5 *M* ammonium EDTA, pH 6.5) with protease inhibitors (Boehringer). Centrifuged (10,000 × g, 1 min) extracts were serially diluted and assayed. The primary antibody was rabbit anti-bovine osteocalcin, the tracer was 125I-bovine osteocalcin, and goat anti-rabbit IgG was used to pellet the tracer bound to the primary antibody. A standard curve was determined with bovine osteocalcin. Antibody recognition depends on osteocalcin epitopes involving both the NH2- and COOH-terminal sequences, essentially requiring intact osteocalcin (at least 45 of the 49 residues); antibody recognition is relatively insensitive to the carboxylation status of Gla.

DISCUSSION

MALDI-MS of ancient fractions from the gravity columns contained a peak mass to charge, *m/z* 5590, in the range of osteocalcin of modern vertebrates (M_r = 5210–5889) (Hauschka et al., 1989) (Figs. 1A and 1B). No reported sequences exist for bison. As shown for the older than 58.9 ka sample, injection of the gravity column fractions into the rp-HPLC resulted in a peak at 11 min (Fig. 1C), which was identified with PMM.

For the PMM, the rp-HPLC fractions containing putative osteocalcin were digested with trypsin, and prepared for MALDI-MS with 4-HCCA matrix. As illustrated for the older than 58.9 ka bison, the PMM of putative osteocalcin for each fossil is consistent with that of osteocalcin from modern bison (Figs. 2A and 2B). Tentative assignments were made for some of the masses by comparison to predicted tryptic fragments of modern cow osteocalcin (see later discussion). The lack of major differences between the ancient and modern PMM suggests the absence of diagenetic products. Because the matrix 4-HCCA promotes decarboxylation of Gla via charge-remote fragmentation, samples were reanalyzed with ATT. By reducing energy transfer to the analyte, ATT decreases the probability for de-

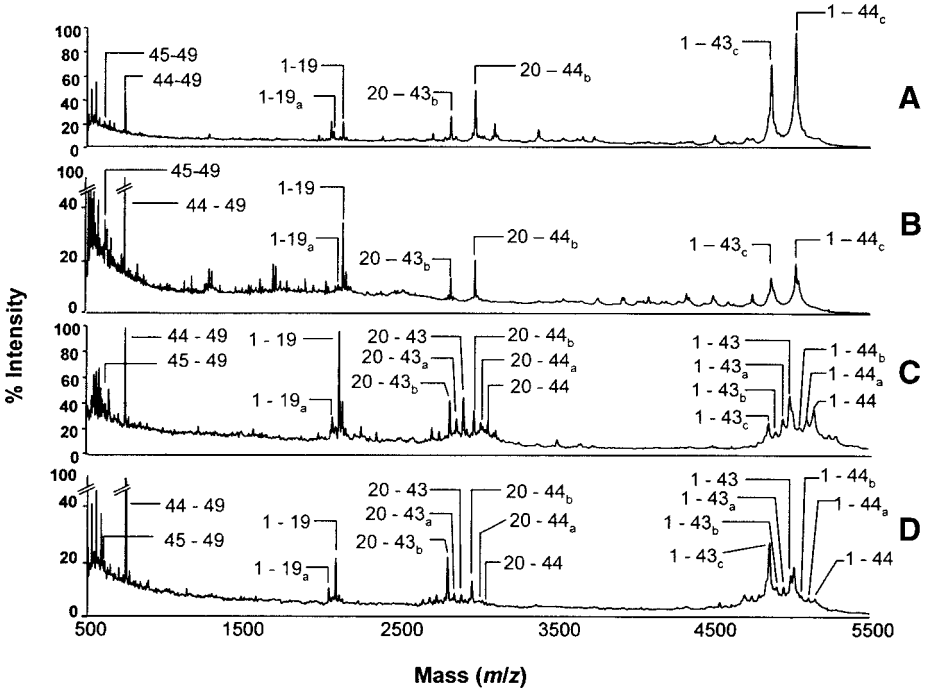

Figure 2. Matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) of osteocalcin tryptic digests. Mass spectra for (A) modern bison and (B) older than 58.9 ka bison obtained using 4-HCCA represent MH^+ ions of proteolytic peptides with fully decarboxylated Gla. Data for (C) modern bison and (D) older than 55.6 ka bison attained using ATT. Subscripts (a, b, and c) on MH^+ correspond to ions with one, two, or three decarboxylated Gla residues, respectively. Assignments were derived by comparison to tryptic peptides from cow osteocalcin. Assigned *m/z* values are (corresponding *m/z* value for modern cow analogues in parentheses) 1–44, 5158 (5287.6); 1–43, 5002 (5131.3); 20–44, 3066 (3066.3); 20–43, 2910 (2910.1); 1–19, 2110 (2240.3); 44–49, 738 (738.9); and 45–49, 582 (582.7).

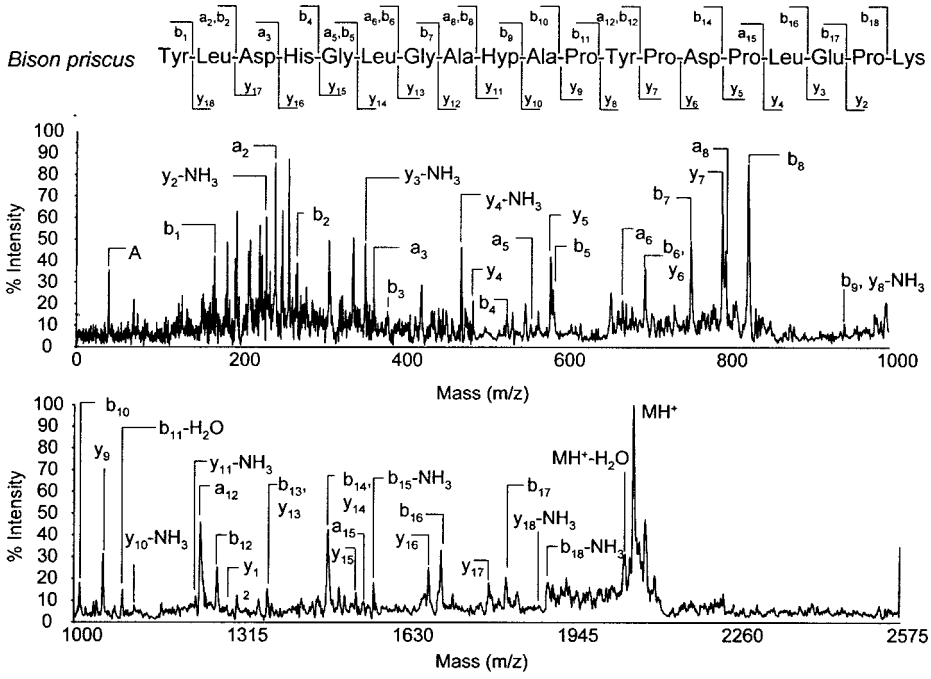

Figure 3. Post-source decay (PSD) of tryptic peptide 1–19 from older than 55.6 ka fossil bison osteocalcin. Precursor ion, MH^+, is represented by peak at the highest *m/z* (2066). Metastable decay products from cleavage of peptide backbone were labeled using Roepstorff nomenclature (Roepstorff and Fohlman, 1984). *B. priscus* sequence appears above mass spectra. For comparison, modern cow is Tyr-Leu-Asp-His-Trp-Leu-Gly-Ala-Hyp-Ala-Pro-Tyr-Pro-Asp-Pro-Leu-Glu-Pro-Lys (GenBank GEBO gi:538590). Two sequences differ by 129 Da (Gly vs. Trp at position 5).

carboxylation. With ATT, the data for peptide 1–43 show a peak at *m/z* 5002 (MH^+) and three lower mass signals (MH_a^+, MH_b^+, and MH_c^+) for the modern and older than 55.6 ka bison (Figs. 2C and 2D). The peak at *m/z* 5002 is consistent with the presence of three fully carboxylated Gla residues, resulting in an increase of 132 *m/z* (3 × 44). The peaks MH_a^+, MH_b^+, and MH_c^+ at *m/z* 4958, 4914, and 4870, respectively, indicate that some ionization-induced decarboxylation occurs. These data suggest that the *m/z* for carboxylated bison osteocalcin is 5722 (5590 + 132). This result was confirmed by reanalyzing putative osteocalcin from the rp-HPLC with ATT (data not shown). The PMM with ATT show a complement of fully carboxylated and ionization-induced decarboxylation peaks for each of the digest peptides in all samples. As with 4-HCCA, no difference existed between the PMM of osteocalcin from modern and fossil bison for ATT, indicating the lack of diagenetic alteration.

The *m/z* values of four peptides in the bison PMM were consistent with those of the C-terminal tryptic peptides of cow osteocalcin (20–43, 21–44, 44–49, and 45–49). Other tryptic peptides (1–19, 1–43, and 1–44) and the mass of putative bison osteocalcin differed from counterparts in cow by 129 Daltons (Da) (*m/z* of cow is 5851 [Hauschka et al., 1989; GenBank GEBO gi:538590]). An equivalent to a traditional tandem mass spectrometry (MS/MS) experiment, PSD MALDI-MS, resolved the 129 Da discrepancy. Only one tryptic peptide, 1–19, *m/z* 2066, was appropriate for PSD. As illustrated for the older than 55.6 ka sample, PSD of modern and ancient samples confirmed that one amino-acid substitution distinguishes cow from bison and accounts for the 129 Da discrepancy (residue 5 of cow is Trp, but for bison it is Gly) (Fig. 3).

Comparison of 377-bp of mtDNA from the older than 55.6 ka sample (CRS-SY-2) with modern bison and cow sequences supports these results (Fig. 4). Although the ancient bison sequence differs by only 4% from those of modern bison (15 of 377-bp), it differs from the cow sequence by 10% (37 out of 377-bp). This is the first case where both protein and DNA sequences have been isolated from the same ancient sample, confirming exceptional biomolecular preservation in permafrost bone (Barnes et al., 2002).

Although we have not yet sequenced osteocalcin from nonpermafrost bone, we have addressed its survivability with RIA data. RIA shows that, when heated to 165 °C for 5.25 h, modern cow bone powder retains 9.7% of osteocalcin of unheated bone (42 pmol osteocalcin/mg bone). These data are consistent with previous immunological results emphasizing the thermal stability of osteocalcin

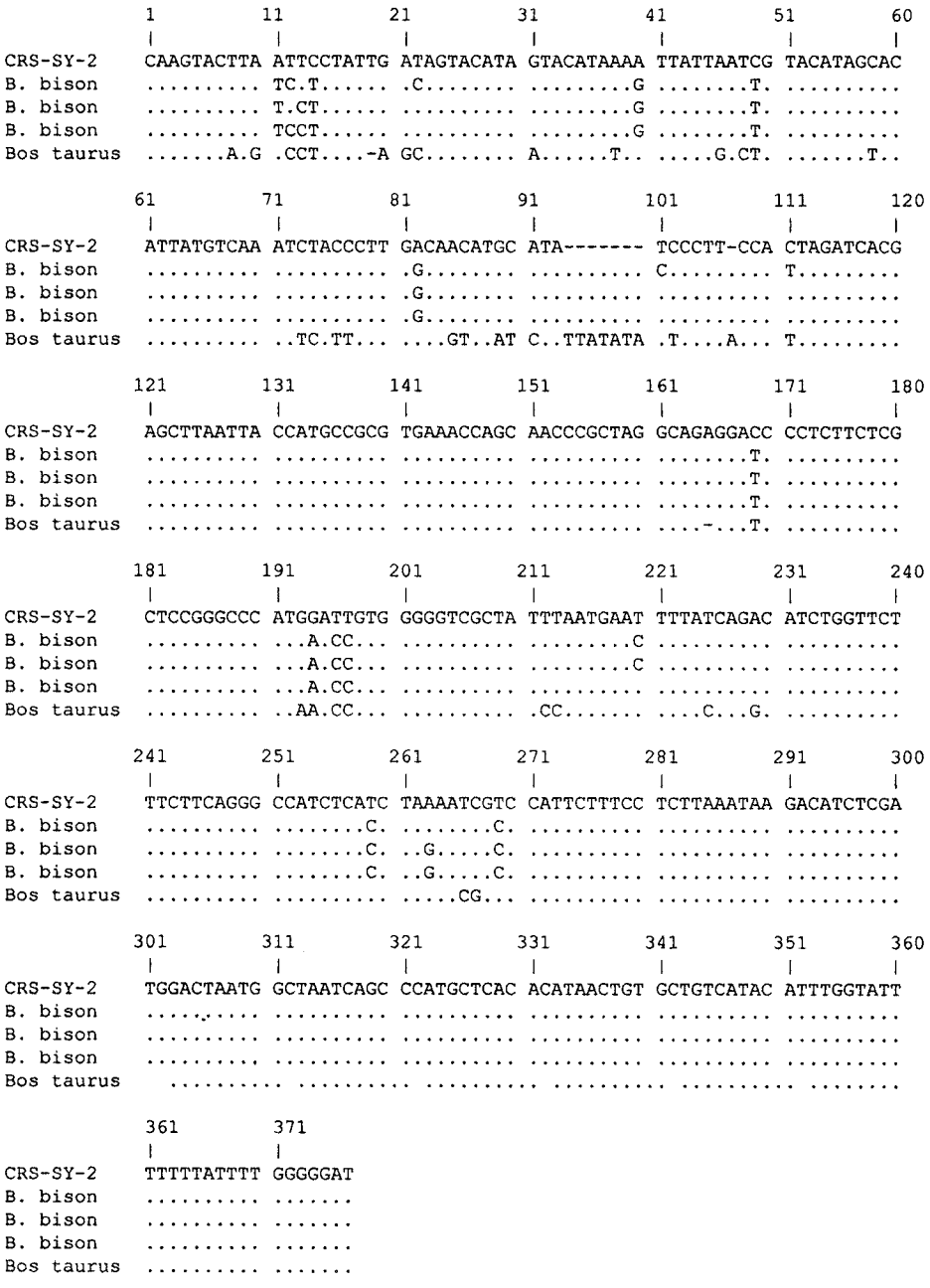

Figure 4. Mitochondrial DNA sequences for older than 55.6 ka bison (Oxford number CRS-SY-2), three modern bison (GenBank accession numbers AF083357, AF083360, and AF083361, second, third, and fourth rows, respectively) and modern cow (GenBank NC001567). Dots represent identity with top (fossil bison) sequence.

(Collins et al., 2000). RIA-confirmed osteocalcin also exists in ancient bison bones: a 120 ka bone with 1.5 pmol/mg and a 300 ka bone with 0.003 pmol/mg. Because MALDI-MS permits sequencing at low picomole levels (Zaluzec et al., 1995), it may be possible to sequence both fossils with small amounts (20–750 mg) of bone powder. While phylogenetic resolution with osteocalcin may be limited, its stability and potential for long-term survival may offer advantages over DNA. Although temporal limits for DNA and protein survival are still unknown, their combined analysis may allow future molecular paleontological investigations to extend farther back in time.

ACKNOWLEDGMENTS

We thank D.A. Gage, coinvestigator, for reviewing this manuscript and I. Barnes, M. Wigger, A. Sher, A. Currant, and B. Saint-Martin for samples. This study was supported by the Wellcome Trust (Nielsen-Marsh and Cooper), the National Science Foundation (NSF) (grant EAR-9805417 to Ostrom and Gage), and the Rhodes Trust (Shapiro). We acknowledge M. Schall for original sequencing of fossil osteocalcin (see Schall et al. [1999]; and other protein sequences); many analytical approaches employed here are modified from earlier work associated with NSF grants EAR-9805417 and EAR-9614235 (Ostrom and Gage).

REFERENCES CITED

Barnes, I., Matheus, P., Shapiro, B., Jensen, D., and Cooper, A., 2002, Dynamics of Pleistocene population extinctions in Beringian brown bears: Science, v. 295, p. 2267–2270.

Collins, M.J., Gernaey, A.M., Nielsen-Marsh, C.M., Vermeer, C., and Westbroek, P., 2000, Slow rates of degradation of osteocalcin: Green light for fossil bone protein?: Geology, v. 28, p. 1139–1142.

Hauschka, P.V., 1980, Osteocalcin: A specific protein of bone with potential for fossil dating, *in* Hare, P.E., ed., Biogeochemistry of amino acids: New York, Wiley, p. 75–82.

Hauschka, P.V., and Carr, S.A., 1982, Calcium-dependant α-helical structure in osteocalcin: Biochemistry, v. 21, p. 2538–2547.

Hauschka, P.V., Lian, J.B., Cole, D.E., and Gundberg, C.M., 1989, Osteocalcin and matrix Gla protein: Vitamin K-dependent proteins in bone: Physiology Review, v. 69, p. 990–1047.

Huq, N.L., Tseng, A., and Chapman, G.E., 1990, Partial amino acid sequence of osteocalcin from an extinct species of ratite bird: Biochemistry International, v. 21, p. 491–496.

Muyzer, G., Sandberg, P., Knapen, M.H.J., Vermeer, C., Collins, M., and Westbroek, P., 1992, Preservation of the bone protein osteocalcin in dinosaurs: Geology, v. 20, p. 871–874.

Ostrom, P.H., Schall, M., Gandhi, H., Shen, T.-L., Hauschka, P.V., Strahler, J.R., and Gage, D.A., 2000, New strategies for characterizing ancient proteins by using matrix-assisted laser desorption ionization mass spectroscopy: Geochimica et Cosmochimica Acta, v. 64, p. 1043–1050.

Pääbo, S., 2000, Of bears, conservation genetics, and the value of time travel: National Academy of Sciences Proceedings, v. 97, p. 1320–1321.

Robbins, L.L., and Donachy, J., 1991, Mineralization regulating proteins in fossil planktonic foraminifera, *in* Sikes, C.S., and Wheeler, A.P., eds., Commodity polypeptides: Washington, D.C., ACS Books, p. 139–148.

Robbins, L.L., Muyzer, G., and Brew, K., 1993, Macromolecules from living and fossil biominerals: Implications for the establishment of molecular phylogenies, *in* Engel, M.H., and Macko, S.A., eds., Organic geochemistry: Principles and applications: New York, Plenum Press, p. 799–816.

Roepstorff, P., and Fohlman, J., 1984, Proposal for a common nomenclature for sequence ions in mass-spectra of peptides: Biomedical Mass Spectrometry, v. 11, p. 601.

Schall, M., Strahler, J., Ostrom, P., Leykam, J., and Gage, D.A., 1999, Characterization of the bone matrix protein osteocalcin in fossils by MALDI-MS: Proceedings of the 47th American Society for Mass Spectrometry Conference on Mass Spectrometry and Allied Topics, June 13–17, 1999, Dallas, Texas, Abstract 207.

Wayne, R.K., Leonard, J.A., and Cooper, A., 1999, Full of sound and fury: The recent history of ancient DNA: Annual Review of Ecology and Systematics, v. 30, p. 457–477.

Zaluzec, E.J., Gage, D.A., and Throck Watson, J., 1995, Matrix assisted laser desorption ionization mass spectrometry applications in peptide and protein characterization: Protein Expression and Purification, v. 6, p. 109–123.

Manuscript received April 25, 2002
Revised manuscript received August 15, 2002
Manuscript accepted August 16, 2002

Printed in USA