Available online at www.sciencedirect.com
SCIENCE DIRECT®
Geochimica et Cosmochimica Acta 70 (2006) 2034–2044
Geochimica
www.elsevier.com/locate/gca

Unraveling the sequence and structure of the protein osteocalcin from a 42 ka fossil horse

Peggy H. Ostrom [a,*], Hasand Gandhi [a], John R. Strahler [b], Angela K. Walker [b], Philip C. Andrews [b], Joseph Leykam [c], Thomas W. Stafford [d], Robert L. Kelly [e], Danny N. Walker [f], Mike Buckley [g], James Humpula [h]

[a] *Department of Zoology, Michigan State University, East Lansing, MI 48824, USA*
[b] *Department of Biological Chemistry, University of Michigan, Ann Arbor, MI 48109, USA*
[c] *Biochemistry and Molecular biology, Michigan State University, East Lansing, MI 48824, USA*
[d] *Stafford Research Labs, 200 Acadia Avenue, Lafayette, CO 80026, USA*
[e] *Anthropology Department, University of Wyoming, Laramie, WY 82071, USA*
[f] *Wyoming State Archaeologist's Office, Laramie, WY 82071, USA*
[g] *Department of Biology, University of York, YO10 5YW, UK*
[h] *Department of Geological Sciences, Michigan State University, East Lansing, MI 48824, USA*

Received 19 October 2005; accepted in revised form 12 January 2006

Abstract

We report the first complete amino acid sequence and evidence of secondary structure for osteocalcin from a temperate fossil. The osteocalcin derives from a 42 ka equid bone excavated from Juniper Cave, Wyoming. Results were determined by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-MS) and Edman sequencing with independent confirmation of the sequence in two laboratories. The ancient sequence was compared to that of three modern taxa: horse (*Equus caballus*), zebra (*Equus grevyi*), and donkey (*Equus asinus*). Although there was no difference in sequence among modern taxa, MALDI-MS and Edman sequencing show that residues 48 and 49 of our modern horse are Thr, Ala rather than Pro, Val as previously reported (Carstanjen B., Wattiez, R., Armory, H., Lepage, O.M., Remy, B., 2002. Isolation and characterization of equine osteocalcin. *Ann. Med. Vet.* **146**(1), 31–38). MALDI-MS and Edman sequencing data indicate that the osteocalcin sequence of the 42 ka fossil is similar to that of modern horse. Previously inaccessible structural attributes for ancient osteocalcin were observed. Glu_{39} rather than Gln_{39} is consistent with deamidation, a process known to occur during fossilization and aging. Two post-translational modifications were documented: Hyp_9 and a disulfide bridge. The latter suggests at least partial retention of secondary structure. As has been done for ancient DNA research, we recommend standards for preparation and criteria for authenticating results of ancient protein sequencing.

1. Introduction

Studies of ancient DNA have advanced our understanding of phylogenetic relationships among extinct taxa and genetic diversity through space and time (Leonard et al., 2000; Paxinos et al., 2002; Hadley et al., 2004; Weinstock et al., 2005). However, the majority of the samples analyzed are young (<50 ka) and/or from permafrost. Because some proteins have a greater likelihood of survival than DNA (Collins et al., 2000) they may provide genetic information that is not accessible via DNA analysis. We putatively identified the bone protein osteocalcin in fossils via matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-MS) and later demonstrated that it could be sequenced from >53,000 ka permafrost bones (Ostrom et al., 2000; Nielsen-Marsh et al., 2002). Partial amino acid sequences of osteocalcin from ca.

* Corresponding author.
E-mail address: Ostrom@msu.edu (P.H. Ostrom).

0016-7037/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.gca.2006.01.004

75 ka Neandertal specimens from Iraq (Nielsen-Marsh et al., 2005) and radioimmunoassay data in 300 ka bones from warm environments (Ostrom et al., 2000) document this protein's survivability.

The recent emergence of proteomics as a molecular tool for evolutionary biology is an exciting prospect. However, efforts to extend ancient biomolecular records in time and space must be increasingly aware of technological limitations and pitfalls. Our challenges are to (1) expand mass spectral databases for proteins from previously uncharacterized taxa, (2) define temporal and spatial limits of survival of osteocalcin, and (3) develop strict criteria for sample handling and analysis and interpretation of mass spectra. Relative to ancient DNA, the science and technology associated with ancient protein sequencing is in its infancy. We should avoid mishaps that befell early ancient DNA research (Wayne et al., 1999). Our efforts to provide the first complete osteocalcin sequence from a non-permafrost fossil speak to the challenges noted above. The sequence was obtained from a ca. 42 ka horse bone from Juniper Cave, Wyoming and compared to that of three phylogenetically related taxa (horse, zebra, and donkey). The ancient osteocalcin sequence was derived by multiple MALDI-MS techniques, confirmed in two laboratories and substantiated by Edman sequencing. The results identify post-translational modifications in the fossil and diage-

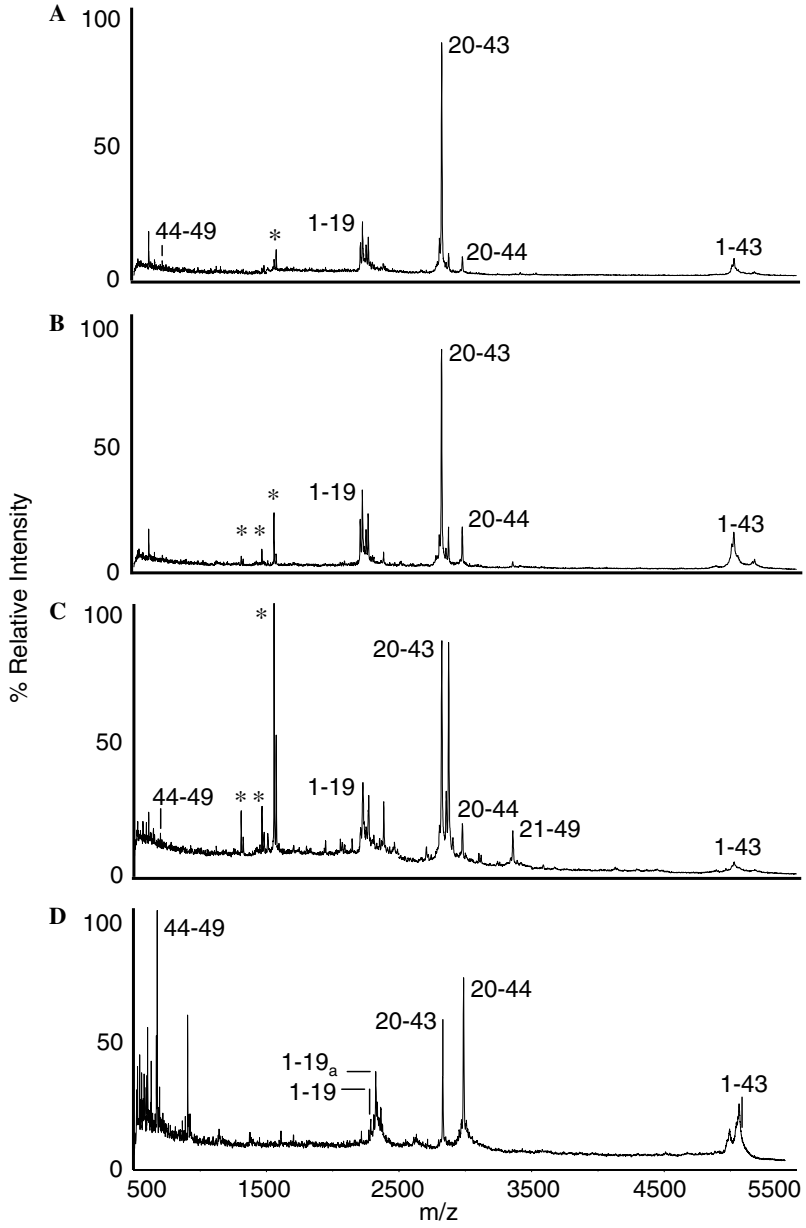

Fig. 1. Peptide mass fingerprint for osteocalcin tryptic digest from: (A) extant horse; (B) zebra; (C) donkey; and (D) partially purified osteocalcin from the 42 ka horse produced from an ABI-4700 (confirmed on an ABI-STR). Peptide 1–19a, 43 Da greater than 1–19, is from a modified form of osteocalcin. *Self-digest products of trypsin.