# Comment on "Protein Sequences from Mastodon and Tyrannosaurus rex Revealed by Mass Spectrometry"

Mike Buckley, *et al.*

*Science* **319**, 33c (2008);
DOI: 10.1126/science.1147046

# Comment on "Protein Sequences from Mastodon and *Tyrannosaurus rex* Revealed by Mass Spectrometry"

Mike Buckley,[1] Angela Walker,[2] Simon Y. W. Ho,[3] Yue Yang,[1] Colin Smith,[4] Peter Ashton,[1] Jane Thomas Oates,[1] Enrico Cappellini,[1] Hannah Koon,[1] Kirsty Penkman,[1] Ben Elsworth,[1] Dave Ashford,[1] Caroline Solazzo,[1] Phillip Andrews,[2] John Strahler,[2] Beth Shapiro,[6] Peggy Ostrom,[5] Hasand Gandhi,[5] Webb Miller,[6] Brian Raney,[7] Maria Ines Zylber,[8] M. Thomas P. Gilbert,[9] Richard V. Prigodich,[10] Michael Ryan,[11] Kenneth F. Rijsdijk,[12] Anwar Janoo,[13] Matthew J. Collins[1]*

We used authentication tests developed for ancient DNA to evaluate claims by Asara *et al*. (Reports, 13 April 2007, p. 280) of collagen peptide sequences recovered from mastodon and *Tyrannosaurus rex* fossils. Although the mastodon samples pass these tests, absence of amino acid composition data, lack of evidence for peptide deamidation, and association of α1(I) collagen sequences with amphibians rather than birds suggest that *T. rex* does not.

Early reports of DNA preservation in multimillion-year-old bones (i.e., from dinosaurs) have been largely dismissed (*1*, *2*) (table S1), but reports of protein recovery are persistent [see (*3*) for review]. Most of these studies used secondary methods of detection, but Asara *et al*. (*2*) recently reported the direct identification of protein sequences, arguably the gold standard for molecular palaeontology, from fossil bones of an extinct mastodon and *Tyrannosaurus rex*. After initial optimism generated by reports of dinosaur DNA, there has been increasing awareness of the problems and pitfalls that bedevil analysis of ancient samples (*1*), leading to a series of recommendations for future analysis (*1*, *4*). As yet, there are no equivalent standards for fossil protein, so here we apply the recommended tests for DNA (*4*) to the authentication of the reported mastodon and *T. rex* protein sequences (*2*) (Table 1).

First, the likelihood of collagen survival needs to be considered. The extremely hierarchical structure of collagen results in unusual, catastrophic degradation (*5*) as a consequence of fibril collapse. The rate of collagen degradation in bone is slow because the mineral "locks" the components of the matrix together, preventing helical expansion, which is a prerequisite of fibril collapse (*6*). The packing that stabilizes collagen fibrils (*6*) also increases the temperature sensitivity of degradation ($E_a$ 173 kJ mol$^{-1}$) (Fig. 1). Collagen decomposition would be much faster in the *T. rex* buried in the then-megathermal (>20°C) (*7*) environment of the Hell Creek formation [collagen half-life ($T_{½}$) = ~2 thousand years (ky)] than it would have been in the mastodon lying within the Doeden Gravel Beds (present-day mean temperature, 7.5°C; collagen $T_{½}$ = 130 ky) (Fig. 1).

This risk of contamination also needs to be evaluated. Collagen is an ideal molecular target for this assessment because the protein has a highly characteristic motif that is also sufficiently variable to enable meaningful comparison between distant taxa if enough sequence is obtained (Fig. 2). Compared with ancient DNA amplification, contamination by collagen is inherently less likely. Furthermore, because the bones sampled in (*2*) were excavated by the

**Table 1.** Key questions to ask about ancient biomolecular investigations [adapted from (*4*)].

| Test | Sample | Pass | Observation |
|---|---|---|---|
| Do the age, environmental history, and preservation of the sample suggest collagen survival? | Mastodon, 300 to 600 ky old | Yes | Collagen $T_{½}$ at 7.5°C = 130 ky |
| | *T. rex*, 65 million years old | No | Collagen $T_{½}$ at 20°C = 2 ky |
| Do the biomolecular and/or macromolecular preservation of the sample, the molecular target, the innate nature of the sample, and its handling history suggest that contamination is a risk? | Biomolecular preservation | ? | Range of evidence presented (*8*) but no amino acid compositional data |
| | Macromolecular preservation | Yes | Macromolecular preservation is not the equivalent of biomolecular preservation (*9*) |
| | Molecular target | Yes | |
| | Handling history | Yes? | Large (2.5 g) samples increase risk of contamination? |
| Do the data suggest that the sequence is authentic, rather than the result of damage and contamination? | Mastodon and *T. rex* | No | Errors in interpretation of spectra [see table S1 and (*13*)]? |
| | | | Damage-induced errors in sequence |
| Do the results make sense, and are there enough data to make the study useful and/or to support the conclusions? | Mastodon | Yes | Weak affinity to mammals |
| | *T. rex* | No | Affinity of α1(I) peptides to amphibians, not birds or reptiles |

[1]BioArch, Departments of Biology, Archaeology, Chemistry and Technology Facility, University of York, Post Office Box 373, York YO10 5YW, UK. [2]Department of Biological Chemistry, University of Michigan Medical School, Ann Arbor, MI 48109–0404, USA. [3]Evolutionary Biology Group, Department of Zoology, University of Oxford, OX1 3PS, UK. [4]Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, D-04103, Leipzig, Germany. [5]Department of Zoology, Michigan State University, East Lansing, MI 48824, USA. [6]Department of Biology, Pennsylvania State University, University Park, PA 16802, USA. [7]Center for Biomolecular Science and Engineering, University of California–Santa Cruz, CA 95064, USA. [8]Department of Parasitology, Kuvin Center, Hebrew University of Jerusalem, Israel. [9]Biological Institute, University of Copenhagen, Universitetsparken 15, 2100 Copenhagen, Denmark. [10]Chemistry Department, Trinity College, 300 Summit Street, Hartford, CT 06106, USA. [11]Cleveland Museum of Natural History, 1 Wade Oval Drive, University Circle, Cleveland, OH 44106, USA. [12]National Museum of Natural History "Naturalis," P.O. Box 9517, 2300 RA Leiden, Netherlands. [13]National Heritage Trust Fund Mauritius, Mauritius Institute, La Chaussée Street Port Louis, Mauritius.

*To whom correspondence should be addressed. E-mail: mc80@york.ac.uk

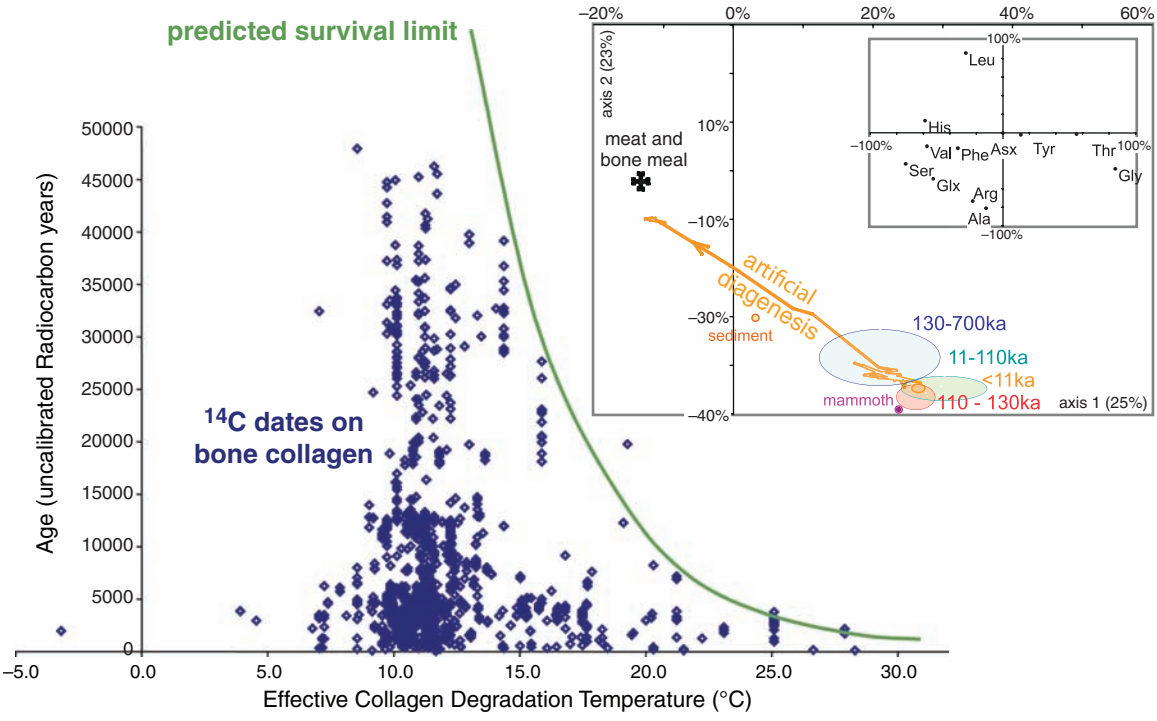


**Fig. 1.** Plot of radiocarbon age versus estimated effective collagen degradation temperature for radiocarbon-dated bones from laboratory databases (principally Oxford and Groningen). The line represents the expected calendar age at which 1% of the original collagen remains following a zero-order reaction; almost no bone collagen survives beyond this predicted limit. **(Inset)** The 99% confidence intervals of amino acid compositions by first two principal component analyses (48% of total variance) for bones from NW Europe aged <11 ky ($n$ = 324), 11 to 110 ky ($n$ = 210), 110 to 130 ky ($n$ = 26), and 130 to 700 ky ($n$ = 31). Pliocene samples are not plotted, as their composition ($n$ = 8) is highly variable and yields of amino acids are low. The orange line indicates a compositional trend observed when compact bone is heated for 32 days at 95°C, which reduces collagen to 1% of the initial concentration [each inflection represents a separate analysis; $n$ = 32)]. The composition becomes more similar to mixed tissue samples (meat and bone meal; $n$ = 32), principally due to the depletion of Gly. An amino acid profile for mammoth is consistent with collagen, unlike the associated sediment sample [data from (*11*)].

authors, obvious contamination sources such as animal glue (used in conservation) can be excluded. However, concentrating protein from the large amounts of bone used (2.5 g) may have heightened the risk of extraneous proteins entering the sample during extraction, although there have been no systematic studies of this phenomenon. Independent extraction and analyses would have strengthened claims for the authenticity of the origin of the peptides (and potentially ameliorated the original problems of data interpretation) (*4*).

The remarkable soft-tissue preservation of the investigated *T. rex* specimen (MOR 1125) has been documented (*8*). However, microscopic preservation does not equate with molecular preservation (*9*). Immunohistochemistry provides support for collagen preservation, but Asara *et al.* (*2*) presented no data regarding inhibition assays with collagen from different species or cross-reactivity with likely contaminants [e.g., fungi (*10*)]. Curiously, no amino acid compositional analysis was conducted [see (*11*)], although immonium ions were identified by time-of-flight secondary ion mass spectrometry. In our experience, collagen-like amino acid profiles have been obtained in all bones from which we could obtain collagen sequence (Fig. 1, inset).

Regarding the proof of sequence authenticity, the spectra reported by Asara *et al.* (*12*) are inconsistent with some of the sequence assignments (*13*) (table S1). A common diagenetic modification, deamidation, not considered in (*2*), may shed light on authenticity. The facile succinimide-mediated deamidation (*14*) of asparagine occurred at $N_{229}G$ and $N_{1156}G$ in ostrich peptides (Ost 4 and Ost5) (see table S1 for nomenclature), presumably during sample preparation. Direct hydrolytic deamidation is slower (*14*), and an expectation of elevated levels of such products is reasonable for old samples. We agree with the most recent interpretation (*13*) of the spectrum illustrated in Fig. 2B as α1(I) $G_{362}$SEGPEGVR$_{370}$, the deamidated (Q→$E_{367}$) form of the sequence found in most mammals (*12*). By way of contrast, none of the three glutamine residues in the reported *T. rex* peptides are deamidated (table S1). Only time will tell if Q→E is a useful marker for authentically old collagen, but from the evidence presented, the mastodon sequence looks more diagenetically altered than *T. rex*.

The unusual, fragmented nature of the reported *T. rex* sequence does not make it amenable to standard, model-based phylogenetic analysis. Instead, we examined the phylogenetic signal of the α1(I) fragments of mastodon and *T. rex* using Neighbor-Net analysis and uncorrected genetic distances. Using the sequences reported in (*13*), both the *T. rex* and mastodon signal display an affinity with amphibians (Fig. 2A). Our reinterpretation of the spectra (*12*) changes the affinity of mastodon but not of *T. rex* (Fig. 2B). In addition to the α1(I) peptides used in the Neighbor-Net analysis, Asara *et al.* reported two other peptides from *T. rex* (*13*); we question the interpretation of the α1(II) spectra (identical to frog) but not the α2(I) spectra (identical to chicken).

We require more data to be convinced of the authenticity of the *T. rex* collagen sequences reported by Asara *et al.* Nevertheless, the handful of spectra reported for the temperate Pleistocene mastodon fail neither phylogenetic nor diagenetic tests, thus highlighting the potential of protein mass spectrometry to bridge the present gulf in our understanding between the fate of archaeological and fossil proteins. To avoid past mistakes of ancient DNA research (*1*), we recommend that future fossil protein claims be considered in light of tests for authenticity such as those presented here.

### Reference and Notes


1. E. Willerslev, A. Cooper, *Proc. R. Soc. London. B. Biol. Sci.* **272**, 3 (2005).
2. J. M. Asara, M. H. Schweitzer, L. M. Freimark, M. Phillips, L. C. Cantley, *Science* **316**, 280 (2007).
3. M. H. Schweitzer, *Palaeontologia Electronica* **5**, editorial 2 (2003); http://palaeo-electronica.org/2002_2/r_and_p.pdf
4. M. T. P. Gilbert, H.-J. Bandelt, M. Hofreiter, I. Barnes, *Trends Ecol. Evol.* **20**, 541 (2005).
5. M. J. Collins, M. S. Riley, A. M. Child, G. Turner-Walker, *J. Archaeol. Sci.* **22**, 175 (1995).
6. C. A. Miles, M. Ghelashvili, *Biophys. J.* **76**, 3243 (1999).
7. K. R. Johnson, D. J. Nichols, J. H. Hartman, *The Hell Creek Formation and the Cretaceous-Tertiary Boundary in the Northern Great Plains: Geological Society of America Special Paper* **361**, 503–510 (2002).
8. M. H. Schweitzer *et al.*, *Science* **316**, 277 (2007).
9. N. S. Gupta, D. E. G. Briggs, R. D. Pancost, *J. Geol. Soc. London* **163**, 897 (2006).
10. P. Sepulveda *et al.*, *Infect. Immun.* **63**, 2173 (1995).

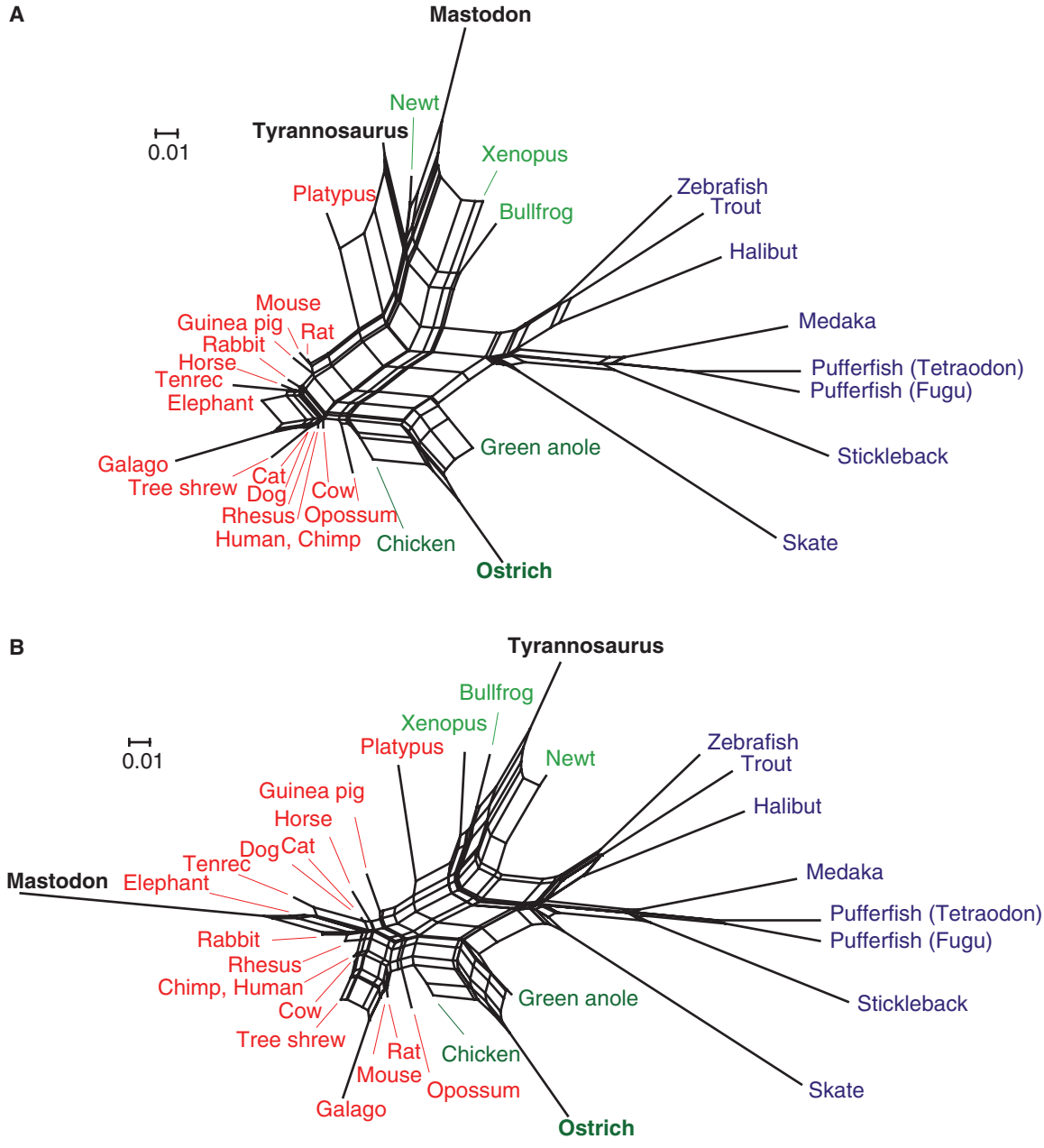


**Fig. 2.** Phylogenetic networks of α1(I) sequences using Neighbor-Net analysis (**A**) with the most recent Asara *et al*. assignments (*13*) and (**B**) after our reinterpretation of the mass spectrometric data (*12*). *T. rex* does not group with bird/reptile using either set of sequence alignments. More sequence is required for a full, model-based phylogenetic analysis.


11. M. Schweitzer, C. L. Hill, J. M. Asara, W. S. Lane, S. H. Pincus, *J. Mol. Evol.* **55**, 696 (2002).
12. See supporting material on *Science* Online.
13. J. M. Asara *et al*., *Science* **317**, 1324 (2007).
14. N. E. Robinson *et al*., *J. Pept. Res.* **63**, 426 (2004).


15. This work was supported by NSF (EAR-0309467), National Environment Research Council (NE511148, GR9/01656), National Center for Research Resources (P41-18627), European Commission (MEST-CT-2004-007909, MEST-CT-2005-020601), Wellcome Trust Bioarchaeology Fellowships (K.P. and H.K.) Analytical Chemistry Trust Fund, the Royal Society of Chemistry Analytical Division, Engineering and Physical Sciences Research Council, and the Michigan Proteome Consortium.


**Supporting Online Material**
www.sciencemag.org/cgi/content/full/319/5859/33c/DC1
SOM Text
Table S1
References


26 June 2007; accepted 20 November 2007
10.1126/science.1147046