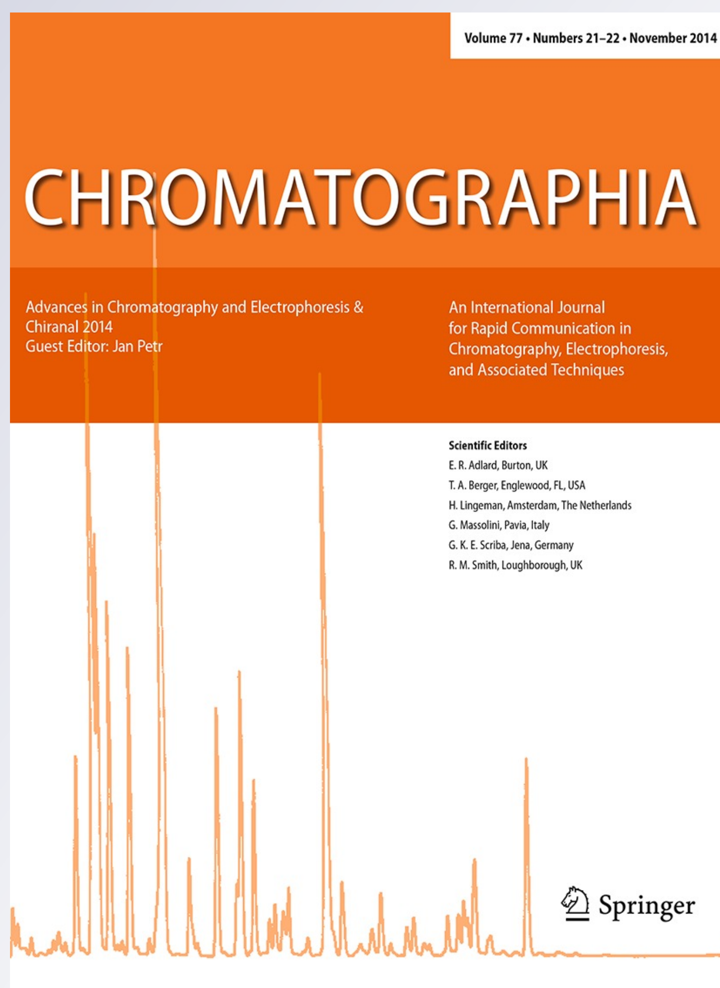
# *Prince Cangrande's Collagen: Study of Protein Modification on the Mummy of the Lord of Verona, Italy (1291–1329 AD)*

**Ivan Mikšík, Pavla Sedláková, Statis Pataridis, Federica Bortolotti, Rossella Gottardo & Franco Tagliaro**

**Chromatographia**
An International Journal for Rapid Communication in Chromatography, Electrophoresis and Associated Techniques

ISSN 0009-5893
Volume 77
Combined 21-22

Chromatographia (2014) 77:1503-1510
DOI 10.1007/s10337-014-2710-0




Springer

Your article is protected by copyright and all rights are held exclusively by Springer-Verlag Berlin Heidelberg. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com”.

Springer

Author's personal copy

Chromatographia (2014) 77:1503–1510
DOI 10.1007/s10337-014-2710-0

SHORT COMMUNICATION

# Prince Cangrande's Collagen: Study of Protein Modification on the Mummy of the Lord of Verona, Italy (1291–1329 AD)

**Ivan Mikšík · Pavla Sedláková · Statis Pataridis · Federica Bortolotti · Rossella Gottardo · Franco Tagliaro**

Received: 29 March 2014 / Revised: 16 May 2014 / Accepted: 3 June 2014 / Published online: 28 June 2014
© Springer-Verlag Berlin Heidelberg 2014

**Abstract** The natural mummy of prince Cangrande, Lord of Verona, Italy (1291–1329 AD) was studied. Two samples were taken: rib bone and muscle. These samples were cleaved with trypsin and analysed by liquid chromatographic methods coupled to mass spectrometry (Q-TOF, ion-trap). Special attention was devoted to nonenzymatic protein modification—the deamidation of asparagine and glutamine. A huge amount of collagen was determined in the tissues of the mummy (covering over 80 % of the sequence)—collagen type I was identified in the rib bone and collagen types I and III in the muscle. A high overall percentage of asparaginyl and glutaminyl residues were deamidated (up to 92 %). In agreement with the literature we can suppose that the deamidation of really old samples (at least 100-years-old) is mainly dependent on the burial conditions and/or thermal age and cannot serve as a precise "molecular clock".

**Keywords** Column liquid chromatography, Mass spectrometry · Collagen · Deamidation · Mummy · Protein modification

Published in the topical collection *Advances in Chromatography and Electrophoresis & Chiranal 2014* with guest editor Jan Petr.

I. Mikšík (✉) · P. Sedláková · S. Pataridis
Institute of Physiology, Academy of Sciences of the Czech Republic, v.v.i., Videnska 1083, 142 20 Prague 4, Czech Republic
e-mail: miksik@biomed.cas.cz

F. Bortolotti · R. Gottardo · F. Tagliaro
Unit of Forensic Medicine, Department of Public Health and Community Medicine, University of Verona, Verona, Italy

## Introduction

Collagen is the most abundant protein in animal tissues. The dominant and best described collagen types are type I and III. These collagens have a triple-helix structure that is highly stable and durable. Collagen type I was described in prehistoric samples, such as in the fossilized bones of a *Tyrannosaurus rex* (68 million-year-old) and Mastodon (*Mammut americanum*) (0.16–0.6 million-year-old) [1] as well as in a mammoth skull (0.1–0.3 million-year-old). However, the possible contamination of million-year-old samples has been widely discussed. An excellent preservation of collagen was also described in mummified human tissue. In this case, collagen was found in soft tissues such as in the skin of the 5,300-year-old Tyrolean Iceman [2], whose body was reported to have undergone a natural process of mummification through freeze-drying.

The asparaginyl (Asn) and glutaminyl (Gln) residues in protein(s) are uniquely unstable under physiological conditions. The deamidation of these amino acid residues is a common post-translational modification. It was shown that this deamidation procedure is pervasive. Deamidation of Asn and Gln is described as time dependent change in charge and conformation of peptides and proteins. It is proposed that deamidation is phenomenon in a remarkably large percentage of human proteins. It has been hypothesized that Asn and Gln may serve, through deamidation, as so called "molecular clock" [3, 4]. With collagen, deamidation is markedly suppressed by its higher order structure. It is questionable whether collagen plays a significant role in the aging process of fossil materials [5].

It seems that collagen can be an ideal protein for dating materials. For this reason discussion intensified about the authenticity of the collagen from a *T. rex* [1] when none of the three reported glutamine residues corresponded to

Author's personal copy



*T. rex*. peptides being deamidated [6]. It was explained that these modification events were difficult to detect and the new mass spectrometer clearly showed deamidation events. However, it is difficult to distinguish between post-translational modifications and diagenetic deamidation, since the *T. rex* and Mastodon sequences were not analyzed before fossilization [7]. The proteome-wide incidence of deamidation was estimated relative to protein recovery to further assess the molecular state of preservation of ancient proteins when Orlando et al. [8] analysed the early Middle Pleistocene horse (*Equus*) metapodial. The deamidation was comparable to the mammoth bone species (*Mammuthus primigenius*) [9].

Cangrande della Scala was the Lord of Verona (Northern Italy) in the fourteenth century (1291–1329). He is well known as the leading patron of the poet Dante Alighieri, but he was also a successful warrior and autocrat. After death, his body was buried in a stone sarcophagus placed on the façade of the Church S. Maria della Scala, in the very center of Verona. In 2004, the body underwent an archaeo-pathological investigation which found it in an excellent condition of natural mummification, probably resulting from a rapid drying due to the microenvironment in the stone sarcophagus. The aim of this work was to analyze the mummy's tissues to study the protein modification(s) over the centuries.

## Experimental

### Chemicals

Trypsin (TPCK treated from bovine pancreas, 13,500 units per mg) and ammonium bicarbonate were obtained from Sigma (Sigma-Aldrich, St. Louis, MO, USA); acetonitrile LC–MS Ultra Chromasolv and formic acid for mass spectrometry were from Fluka (Sigma-Aldrich). 2-mercaptoethanol, cyanogen bromide (CNBr) were from Merck (Darmstadt, Germany). Collagen standards (collagen from human placenta—Bornstein and Traub type I and III) were purchased from Sigma. All solutions were prepared in MilliQ water (Millipore, Bedford, MA, USA).

### Sample Preparation

Small pieces of rib bone and muscle samples, collected from the mummy of Cangrande della Scala, were stored at room temperature until analysis.

The data obtained from the mummy were compared with recent tissues collected from forensic autopsies performed in 2013, as well as with collagen standards (human collagen type I and III from Sigma).

Samples from the mummy (cca 2 mg at concentration 1 mg $mL^{-1}$) were treated with trypsin solution (1:20 enzyme:substrate ratio) in a 20 mmol $L^{-1}$ ammonium bicarbonate buffer (pH 7.8) at 37 °C for 3 h. After treatment, the reaction was stopped by adding 2 % formic acid and samples were stored at −18 °C.

Samples from recent tissues underwent CNBr/trypsin procedure. Samples (cca 2 mg) were incubated in 0.2 mol $L^{-1}$ ammonium bicarbonate, pH 7.0, containing 25 % (v/v) 2-mercaptoethanol to reduce the oxidized methionyl residues. After this reduction 2-mercaptoethanol was evaporated, and the samples were cleaved with CNBr in 70 % (v/v) formic acid under nitrogen for 24 h. The samples were lyophilised three times. Next procedure was trypsin digestion, the same as described above for samples from mummy.

All samples were done in triplicate.

### Analysis of Peptide Digests with LC–MS/MS

Analyses were performed on nanoLC (Proxeon) coupled with high-resolution Q-TOF MS (MaXis, Bruker) using a Biosphere C18 column (150 mm × 75 μm i.d., NanoSeparations) and on HPLC coupled with ion-trap MS (HCT Ultra ETD II, Bruker) with CID and ETD fragmentation using a Jupiter Proteo 90 A column (250 × 2 mm, Phenomenex).

Using nanoLC/MS, separation of the peptides was achieved via a linear gradient between mobile phase A (water) and B (acetonitrile), both containing 0.1 % (v/v) formic acid. Separation was started by running the system with 5 % mobile phase B, followed by a gradient elution to 30 % B at 70 min. The next step was gradient elution to 50 % B in 10 min, and then a gradient to 100 % B in 8 min. Finally, the column was eluted with 100 % B for 2 min. Equilibration before the next run was achieved by washing the column with 5 % mobile phase B for 10 min. The flow rate was 0.25 μL $min^{-1}$, and the column was held at ambient temperature (25 °C).

On-line nano-electrospray ionization (nESI) was used in positive mode. The ESI voltage was set to +4.5 kV, scan time 1.3 Hz. Operating conditions: drying gas (N2), 1 L $min^{-1}$; drying gas temperature, 160 °C; nebulizer pressure, 0.4 bar. Experiments were performed by scanning from 100 to 2,200 *m/z*. The reference ion used (internal mass lock) was a monocharged ion of C24H19F36N3O6P3 (*m/z* 1,221.9906). Mass spectra corresponding to each signal from the total ion current chromatogram were averaged, enabling an accurate molecular mass determination. All nLC-MS/MS analyses were done in duplicate.

For the LC/MS, the chromatographic separations were carried out in RP-HPLC mode using a Jupiter 4 μm Proteo 90A column (250 × 2 mm I.D., Phenomenex, Torrance, USA). A 20 μL sample was injected. Elution was achieved using a linear gradient (A = water with 0.1 % formic acid,

Author's personal copy



and B = acetonitrile with 0.085 % formic acid). Separation was initiated by running the system isocratically for 2 min with 2 % mobile phase B, followed by a gradient elution to 35 % B at 40 min. Finally, the column was eluted with 100 % B for 10 min. Equilibration before the next run was achieved by washing with buffer A for 10 min. The flow rate was 0.25 mL $min^{-1}$, the column temperature was held at 25 °C and UV absorbance detection was done at 214 nm.

Atmospheric pressure ionization-electrospray ionization (API-ESI) ion-trap mass spectrometry was used. The conditions used with the CID and ETD fragmentation were: drying gas (N2), 10 L $min^{-1}$; drying gas temperature, 350 °C; nebulizator pressure, 25 psi (172.4 kPa); ions were observed over the mass range *m/z* 300–1,400 (MS—standard mode, MS/MS—enhanced mode 100–2,300 *m/z*). Analysis was done in auto MS/MS mode (10 precursor ions, excluded after two spectra for 0.5 min).

### Database Searching

Data were processed using ProteinScape software (version 3.0.0.446). Proteins were identified by correlating tandem mass spectra to the IPI (v. 3.87) and SwissProt (v. 8/2013) databases, using the MASCOT search engine (v. 2.3.0) (http://www.matrixscience.com). The taxonomy was restricted to Homo sapiens to remove protein identification redundancy. Trypsin (or CNBr/trypsin) was chosen as the enzyme parameter. One missed cleavage was allowed, and an initial peptide mass tolerance of ±10.0 ppm was used for MS and ±0.05 Da for MS/MS analysis in the case of Q-TOF MS (settings for ion-trap were: an initial peptide mass tolerance of ±1.2 Da was used for MS and ±0.6 Da for MS/MS analysis). Lysines and prolines were assumed to be hydroxylated, methionine was allowed to be oxidated whereas asparagine and glutamine deamidated. All these possible modifications were set to be variable. The monoisotopic peptide charge was set to 1+, 2+ and 3+. The Peptide Decoy option was selected during the data search process to remove false-positive results. Only significant hits were accepted (MASCOT score ≥60 for proteins and MASCOT score ≥20 for peptides), http://www.matrixscience.com).

## Results and Discussion

The mummy was well preserved. It must be emphasized that the sarcophagus was placed on the façade of the church, and so was influenced by a higher temperature compared to the mummified Tyrolean Iceman [2] or Siberian mammoth [9]. The tissue could be simply treated with trypsin to obtain peptide fragments. This is mostly important for collagen. This protein in tissues normally forms a stable triple-helix structure that is relatively stable against proteolytic enzymes. Treatment of the control (recent) tissues with trypsin did not lead to a significant yield of collagenous peptides. For this reason these recent tissues were subsequently treated with CNBr followed by trypsin cleavage. This is a normal procedure in connective tissue research.

Three collagenous chains of two types were observed in the mummy: collagen type 1 (2 chains) and type III (1 chain) with good protein sequence coverage (see Tables 1, 2; Figs. 1, 2) when at least one-half of protein sequence was described.

This coverage was similar for both the analysed tissues (rib bone and muscle)—above 80 % for collagen type I (chain 1), above 50 % for collagen type I (chain 2) and almost 50 for collagen type III.

The deamidation was very high. Despite obtaining high sequence coverage, not all sequences containing Asn and Gln were determined. In terms of the deamidated Asn and Gln in the determined sequences, the deamidation rate was very high—above 90 % for collagen type I (chain 1) and above 80 % for collagen type III. In terms of the Asn and Gln overall, the deamidation was lower, but still high (around 80 % and almost 40 %, respectively). Compared to the control samples (tissues as well as standards) deamidation is significantly very higher, i.e. the deamidation of control samples was really low. These experiments verified that deamidation was not an artifact from the sample preparation procedure. Higher sequence coverage for archeological vs. modern materials can be caused by long aging of tissues, their natural "treatment" and so good acceptability for enzyme cleavage. However, due to its rigid structure collagen survive in fossils when proteome degradation is high [11]. On the opposite collagen from modern material is a relatively rigid protein resistant to enzymatic cleavage (before trypsin cleavage it was necessary to disrupt structure by CNBr). Discrepancy between sequence coverage of chains α1 and α2 can be caused by slightly different degradation procedure.

These results can be compared to published results on an early Middle Pleistocene horse (*Equus*) bone (560–780 1,000 years before present) [8] as well as a pleistocene Mammoth femur (*M. primigenius*) [9]. The mammoth bone specimens preserved in temperate environments had a high amount of deamidated sites (approx. 80 %) while the mammoth bone specimens preserved in the Siberian permafrost had only 53 %. Older early Middle Pleistocene horse bone from the Yukon had 53 % deamidated sites. This is in agreement with the findings that the extent of deamidation is more influenced by the burial conditions and thermal age than by chronological age [10]. The mummy of Cangrande was purely influenced by temperature, so a high content of deamidated sites was observed. There were also different

Author's personal copy



**Table 1** Coverage od deamidation process of collagens

| | Mummy | | | | | Control | | | | | Standard | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Rib bone | | Muscle | | | Rib bone | | Muscle | | | | | |
| | Alpha 1 (I) | Alpha 2 (I) | Alpha 1 (I) | Alpha 2 (I) | Alpha 1 (III) | Alpha 1 (I) | Alpha 2 (I) | Alpha 1 (I) | Alpha 2 (I) | Alpha 1 (III) | Alpha 1 (I) | Alpha 2 (I) | Alpha 1 (III) |
| **Sequence coverage [%]** | **82.4** | **64.4** | **80.2** | **55.2** | **49.7** | **74.6** | **75.5** | **51.6** | **35.8** | **48.3** | **77.6** | **64.3** | **81.4** |
| Asn coverage [%] | 100.0 | 79.2 | 100.0 | 50.0 | 45.8 | 100.0 | 91.7 | 54.5 | 33.3 | 37.5 | 100.0 | 66.7 | 83.3 |
| Gln coverage [%] | 83.3 | 45.5 | 80.0 | 36.4 | 46.2 | 73.3 | 50.0 | 46.7 | 9.1 | 38.5 | 70.0 | 45.5 | 80.8 |
| **Asn and Gln coverage [%]** | **87.8** | **63.0** | **85.4** | **43.5** | **46.0** | **80.5** | **71.7** | **48.8** | **21.7** | **38.0** | **78.0** | **56.5** | **82.0** |
| Deamidated Asn [% of whole content] | 100.0 | 50.0 | 100.0 | 45.8 | 41.7 | 18.2 | 25.0 | 0.0 | 0.0 | 8.3 | 0.0 | 4.2 | 12.5 |
| Deamidated Gln [% of whole content] | 73.3 | 22.7 | 70.0 | 22.7 | 34.6 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| **Deamidated Asn and Gln [% of whole content]** | **80.5** | **37.0** | **78.0** | **34.8** | **38.0** | **4.9** | **13.0** | **0.0** | **0.0** | **4.0** | **0.0** | **2.2** | **6.0** |
| Deamidated Asn [% of determined] | 100.0 | 63.2 | 100.0 | 91.7 | 90.9 | 18.2 | 27.3 | 0.0 | 0.0 | 22.2 | 0.0 | 6.3 | 15.0 |
| Deamidated Gln [% of determined] | 88.0 | 50.0 | 87.5 | 62.5 | 75.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

Author's personal copy



**Table 2** Overview of deamidation of collagen from mummy, control (biopsy) and standard

| Collagen alpha 1 (I) | | | | | | Collagen alpha 2 (I) | | | | | | Collagen alpha 1 (III) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Position AA | Mummy | | Control | | Standard | Position AA | Mummy | | Control | | Standard | Position AA | Mummy | Control | Standard |
| | Rib bone | Muscle | Rib bone | Muscle | | | Rib bone | Muscle | Rib bone | Muscle | | | Muscle | Muscle | |
| 162 Q | *nd* | *nd* | *nd* | *nd* | *nd* | 80 Q | *nd* | *nd* | *nd* | *nd* | *nd* | 149 Q | *nd* | *nd* | *nd* |
| 199 Q | *nd* | *nd* | | *nd* | *nd* | 111 Q | *nd* | **X** | *nd* | *nd* | *nd* | 150 N | *nd* | *nd* | *nd* |
| 202 Q | *nd* | *nd* | | *nd* | *nd* | 114 Q | *nd* | **X** | *nd* | *nd* | *nd* | 154 Q | *nd* | *nd* | *nd* |
| 229 N | **X** | **X** | | | | 125 Q | *nd* | | *nd* | *nd* | *nd* | 200 Q | *nd* | *nd* | *nd* |
| 250 Q | **X** | **X** | | | | 162 Q | **X** | **X** | | | | 208 Q | *nd* | *nd* | *nd* |
| 295 N | **X** | **X** | **X** | *nd* | | 183 N | *nd* | *nd* | | *nd* | *nd* | 272 N | *nd* | *nd* | |
| 300 Q | **X** | **X** | | *nd* | | 191 Q | | | *nd* | *nd* | *nd* | 287 N | **X** | *nd* | |
| 324 N | **X** | **X** | | *nd* | | 207 N | **X** | *nd* | **X** | *nd* | | 293 N | **X** | *nd* | |
| 358 Q | **X** | **X** | | | | 212 Q | **X** | *nd* | | *nd* | | 322 N | *nd* | *nd* | *nd* |
| 367 Q | **X** | **X** | | | | 272 N | **X** | **X** | **X** | *nd* | | 331 Q | *nd* | *nd* | |
| 387 N | **X** | **X** | | *nd* | | 299 N | | *nd* | | *nd* | *nd* | 362 N | **X** | **X** | **X** |
| 393 Q | **X** | **X** | | *nd* | | 303 N | | *nd* | **X** | *nd* | *nd* | 367 Q | **X** | | |
| 400 N | **X** | **X** | | | | 359 N | | | | *nd* | *nd* | 374 Q | *nd* | *nd* | |
| 421 Q | **X** | **X** | | | | 369 Q | | | | *nd* | *nd* | 380 Q | *nd* | *nd* | |
| 432 N | **X** | **X** | | | | 384 N | **X** | **X** | **X** | *nd* | | 389 N | *nd* | *nd* | |
| 457 Q | **X** | **X** | | | | 445 N | **X** | **X** | | | *nd* | 422 N | **X** | **X** | **X** |
| 561 Q | **X** | **X** | *nd* | *nd* | *nd* | 458 N | **X** | **X** | | | | 437 N | *nd* | *nd* | |
| 576 Q | **X** | **X** | | *nd* | | 491 N | **X** | **X** | | | | 476 N | **X** | | |
| 619 Q | **X** | **X** | | | | 507 N | *nd* | *nd* | *nd* | *nd* | *nd* | 497 N | **X** | | **X** |
| 634 Q | **X** | | *nd* | *nd* | *nd* | 527 N | **X** | *nd* | | *nd* | | 557 Q | *nd* | *nd* | |
| 643 Q | | | *nd* | *nd* | *nd* | 528 N | | *nd* | | *nd* | | 574 Q | | *nd* | |
| 661 Q | **X** | | *nd* | *nd* | *nd* | 531 Q | | *nd* | | *nd* | | 586 N | *nd* | *nd* | |
| 688 Q | **X** | **X** | | | | 537 Q | | *nd* | | *nd* | | 593 N | *nd* | *nd* | *nd* |
| 700 N | **X** | **X** | | *nd* | | 540 Q | | *nd* | | *nd* | | 605 Q | **X** | | |
| 705 N | **X** | **X** | | *nd* | | 546 Q | *nd* | *nd* | *nd* | *nd* | | 611 N | *nd* | | |
| 721 Q | | **X** | | *nd* | | 555 Q | *nd* | *nd* | *nd* | *nd* | | 617 Q | *nd* | | |
| 727 Q | | **X** | | *nd* | | 620 N | **X** | **X** | | *nd* | | 638 Q | *nd* | *nd* | |
| 822 Q | **X** | **X** | | *nd* | | 668 N | **X** | **X** | | | | 641 Q | *nd* | *nd* | |
| 855 N | **X** | **X** | | | | 723 N | | **X** | **X** | *nd* | | 653 N | *nd* | *nd* | |
| 891 N | **X** | **X** | | | | 734 Q | **X** | **X** | | *nd* | | 725 Q | *nd* | | |
| 952 Q | **X** | **X** | | | | 750 N | | *nd* | | *nd* | | 775 Q | **X** | | |
| 957 Q | **X** | **X** | | | | 768 N | **X** | *nd* | | *nd* | | 817 Q | | *nd* | |

Author's personal copy



**Table 2** continued

| Collagen alpha 1 (I) | | | | | | Collagen alpha 2 (I) | | | | | | Collagen alpha 1 (III) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Position AA | Mummy | | Control | | Standard | Position AA | Mummy | | Control | | Standard | Position AA | Mummy | Control | Standard |
| | Rib bone | Muscle | Rib bone | Muscle | | | Rib bone | Muscle | Rib bone | Muscle | | | Muscle | Muscle | |
| 966 Q | **X** | **X** | | | | 825 Q | **X** | **X** | *nd* | *nd* | | 818 N | **X** | *nd* | |
| 985 Q | **X** | **X** | | | | 864 Q | *nd* | *nd* | | *nd* | *nd* | 857 Q | *nd* | | |
| 1090 Q | **X** | **X** | | | | 918 N | **X** | **X** | **X** | | **X** | 887 N | *nd* | *nd* | |
| 1102 Q | **X** | *nd* | *nd* | *nd* | | 929 N | | **X** | | | | 889 N | *nd* | *nd* | |
| 1117 Q | **X** | **X** | | | | 932 N | **X** | **X** | | | | 910 N | **X** | | |
| 1129 Q | **X** | **X** | | | | 941 Q | *nd* | *nd* | *nd* | *nd* | *nd* | 928 Q | | *nd* | |
| 1156 N | **X** | **X** | **X** | | | 953 N | *nd* | *nd* | | | | 938 Q | **X** | | |
| 1203 Q | *nd* | *nd* | *nd* | *nd* | *nd* | 977 N | *nd* | *nd* | *nd* | *nd* | *nd* | 974 Q | **X** | | |
| 1206 Q | *nd* | *nd* | *nd* | *nd* | *nd* | 1002 Q | **X** | *nd* | | | | 986 N | *nd* | | |
| | | | | | | 1026 N | *nd* | *nd* | | *nd* | *nd* | 998 Q | **X** | | |
| | | | | | | 1029 Q | *nd* | *nd* | | *nd* | *nd* | 1015 N | **X** | | |
| | | | | | | 1041 Q | *nd* | *nd* | | *nd* | *nd* | 1037 N | | | |
| | | | | | | 1083 Q | *nd* | *nd* | *nd* | *nd* | *nd* | 1088 Q | **X** | | |
| | | | | | | 1086 Q | *nd* | *nd* | *nd* | *nd* | *nd* | 1114 N | **X** | *nd* | |
| | | | | | | | | | | | | 1126 Q | **X** | *nd* | |
| | | | | | | | | | | | | 1127 Q | **X** | *nd* | |
| | | | | | | | | | | | | 1168 N | *nd* | *nd* | *nd* |
| | | | | | | | | | | | | 1183 Q | *nd* | *nd* | *nd* |

*nd* not detected, X deamidated, *empty cell* unmodified

Author's personal copy



**Fig. 1** Sequence of human collagen type I, chain alpha 1—rib bone from mummy Cangrande. Matched peptides are in *yellow*; propeptides are highlighted as *lower-case letters*; asparaginyl and glutaminyl residues are in **bold** while deamidated ones are in *red*

```
   1 mfsfvdlrll lllaatallt hgqeegqveg qdedippitc vqnglryhdr   50
  51 dvwkpepcri cvcdngkvlc ddvicdetkn cpgaevpege ccpvcpdgse  100
 101 sptdqettgv egpkgdtgpr gprgpagppg rdgipgqpgl pgppgppgpp  150
 151 gppglggnfa pQLSYGYDEK STGGISVPGP MGPSGPRGLP GPPGAPGPQG  200
 201 FQGPPGEPGE PGASGPMGPR GPPGPPGKNG DDGEAGKPGR PGERGPPGPQ  250
 251 GARGLPGTAG LPGMKGHRGF SGLDGAKGDA GPAGPKGEPG SPGENGAPGQ  300
 301 MGPRGLPGER GRPGAPGPAG ARGNDGATGA AGPPGPTGPA GPPGFPGAVG  350
 351 AKGEAGPQGP RGSEGPQGVR GEPGPPGPAG AAGPAGNPGA DGQPGAKGAN  400
 401 GAPGIAGAPG FPGARGPSGP QGPGGPPGPK GNSGEPGAPG SKGDTGAKGE  450
 451 PGPVGVQGPP GPAGEEGKRG ARGEPGPTGL PGPPGERGGP GSRGFPGADG  500
 501 VAGPKGPAGE RGSPGPAGPK GSPGEAGRPG EAGLPGAKGL TGSPGSPGPD  550
 551 GKTGPPGPAG QDGRPGPPGP PGARGQAGVM GFPGPKGAAG EPGKAGERGV  600
 601 PGPPGAVGPA GKDGEAGAQG PPGPAGPAGE RGEQGPAGSP GFQGLPGPAG  650
 651 PPGEAGKPGE QGVPGDLGAP GPSGARGERG FPGERGVQGP PGPAGPRGAN  700
 701 GAPGNDGAKG DAGAPGAPGS QGAPGLQGMP GERGAAGLPG PKGDRGDAGP  750
 751 KGADGSPGKD GVRGLTGPIG PPGPAGAPGD KGESGPSGPA GPTGARGAPG  800
 801 DRGEPGPPGP AGFAGPPGAD GQPGAKGEPG DAGAKGDAGP PGPAGPAGPP  850
 851 GPIGNVGAPG AKGARGSAGP PGATGFPGAA GRVGPPGPSG NAGPPGPPGP  900
 901 AGKEGGKGPR GETGPAGRPG EVGPPGPPGP AGEKGSPGAD GPAGAPGTPG  950
 951 PQGIAGQRGV VGLPGQRGER GFPGLPGPSG EPGKQGPSGA SGERGPPGPM 1000
1001 GPPGLAGPPG ESGREGAPGA EGSPGRDGSP GAKGDRGETG PAGPPGAPGA 1050
1051 PGAPGPVGPA GKSGDRGETG PAGPTGPVGP VGARGPAGPQ GPRGDKGETG 1100
1101 EQGDRGIKGH RGFSGLQGPP GPPGSPGEQG PSGASGPAGP RGPPGSAGAP 1150
1151 GKDGLNGLPG PIGPPGPRGR TGDAGPVGPP GPPGPPGPPG PPSAGFDFSF 1200
1201 LPQPPQEKAH DGGRYYRAdd anvvrdrdle vdttlkslsq qienirspeg 1250
1251 srknpartcr dlkmchsdwk sgeywidpnq gcnldaikvf cnmetgetcv 1300
1301 yptqpsvaqk nwyisknpkd krhvwfgesm tdgfqfeygg qgsdpadvai 1350
1351  ltflrlmst easqnityhc knsvaymdqq tgnlkkalll qgsneieira 1400
1401 egnsrftysv tvdgctshtg awgktvieyk ttktsrlpii dvapldvgap 1450
1451 dqefgfdvgp vcfl                                        1500
```

**Fig. 2** Sequence of human collagen type I, chain alpha 2—rib bone from mummy Cangrande. Matched peptides are in *yellow*; propeptides are highlighted as *lower-case letters*; asparaginyl and glutaminyl residues are in **bold** while deamidated ones are in *red*

```
   1 mlsfvdtrtl lllavtlcla tcqslqeetv rkgpagdrgp rgergppgpp   50
  51 grdgedgptg ppgppgppgp pglggnfaaQ YDGKGVGLGP GPMGLMGPRG  100
 101 PPGAAGAPGP QGFQGPAGEP GEPGQTGPAG ARGPAGPPGK AGEDGHPGKP  150
 151 GRPGERGVVG PQGARGFPGT PGLPGFKGIR GHNGLDGLKG QPGAPGVKGE  200
 201 PGAPGENGTP GQTGARGLPG ERGRVGAPGP AGARGSDGSV GPVGPAGPIG  250
 251 SAGPPGFPGA PGPKGEIGAV GNAGPAGPAG PRGEVGLPGL SGPVGPPGNP  300
 301 GANGLTGAKG AAGLPGVAGA PGLPGPRGIP GPVGAAGATG ARGLVGEPGP  350
 351 AGSKGESGNK GEPGSAGPQG PPGPSGEEGK RGPNGEAGSA GPPGPPGLRG  400
 401 SPGSRGLPGA DGRAGVMGPP GSRGASGPAG VRGPNGDAGR PGEPGLMGPR  450
 451 GLPGSPGNIG PAGKEGPVGL PGIDGRPGPI GPAGARGEPG NIGFPGPKGP  500
 501 TGDPGKNGDK GHAGLAGARG APGPDGNNGA QGPPGPQGVQ GGKGEQGPPG  550
 551 PPGFQGLPGP SGPAGEVGKP GERGLHGEFG LPGPAGPRGE RGPPGESGAA  600
 601 GPTGPIGSRG PSGPPGPDGN KGEPGVVGAV GTAGPSGPSG LPGERGAAGI  650
 651 PGGKGEKGEP GLRGEIGNPG RDGARGAPGA VGAPGPAGAT GDRGEAGAAG  700
 701 PAGPAGPRGS PGERGEVGPA GPNGFAGPAG AAGQPGAKGE RGAKGPKGEN  750
 751 GVVGPTGPVG AAGPAGPNGP PGPAGSRGDG GPPGMTGFPG AAGRTGPPGP  800
 801 SGISGPPGPP GPAGKEGLRG PRGDQGPVGR TGEVGAVGPP GFAGEKGPSG  850
 851 EAGTAGPPGT PGPQGLLGAP GILGLPGSRG ERGLPGVAGA VGEPGPLGIA  900
 901 GPPGARGPPG AVGSPGVNGA PGEAGRDGNP GNDGPPGRDG QPGHKGERGY  950
 951 PGNIGPVGAA GAPGPHGPVG PAGKHGNRGE TGPSGPVGPA GAVGPRGPSG 1000
1001 PQGIRGDKGE PGEKGPRGLP GLKGHNGLQG LPGIAGHHGD QGAPGSVGPA 1050
1051 GPRGPAGPSG PAGKDGRTGH PGTVGPAGIR GPQGHQGPAG PPGPPGPPGP 1100
1101 PGVSGGGYDF GYDGDFYRAd qprsapslrp kdyevdatlk slnnqietll 1150
1151 tpegsrknpa rtcrdlrlsh pewssgyywi dpnqgctmda ikvycdfstg 1200
1201 etciraqpen ipaknwyrss kdkkhvwlge tinagsqfey nvegvtskem 1250
1251 atqlafmrll anyasqnity hcknsiaymd eetgnlkkav ilqgsndvel 1300
1301 vaegnsrfty tvlvdgcskk tnewgktiie yktnkpsrlp fldiapldig 1350
1351 gadqeffvdi gpvcfk                                       1400
```

Author's personal copy



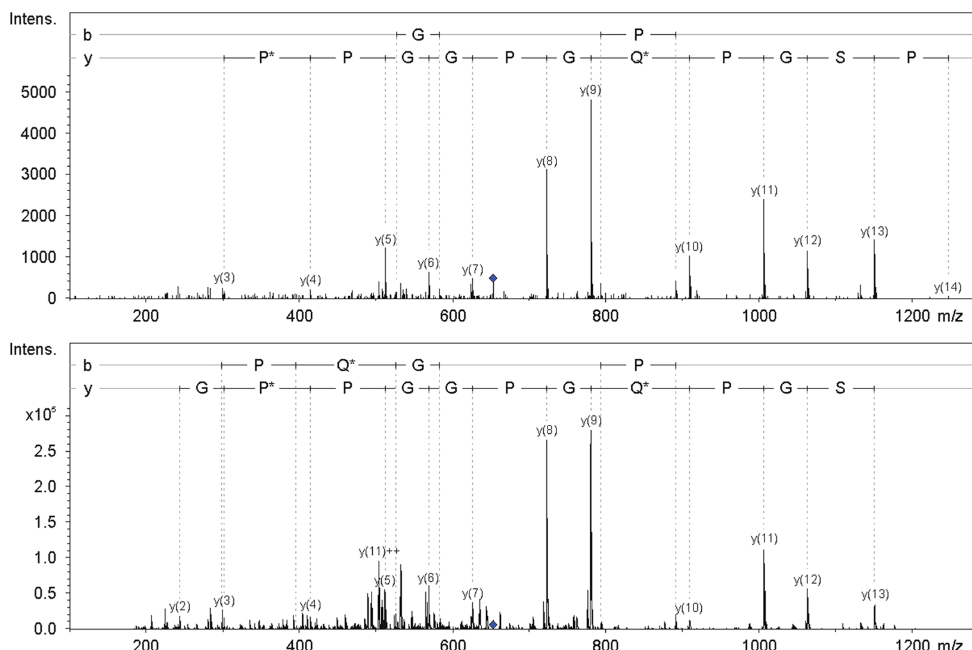



**Fig. 3** MS/MS spectra of peptide GPSGPQGPGGPPGPK from collagen type I, chain alpha 1 from muscle of mummy Cangrande. Above analysis was made by high-resolution Q-TOF (Mascot score 142.4), below analysis by ion-trap (Mascot score 96.8)

deamidation rates of asparaginyl and glutaminyl observed (see Table 1) when deamidation of asparaginyl was significantly higher. This is in agreement to previously published results on early Middle Pleistocene horse bone and Mammouth femur [8-Supplementary information].

The deamidation of collagen from a mummy 675 years old was comparable to the deamidation of the Siberian mammoth and was higher than in the mammoth preserved in temperate environments as well as a pleistocene horse from the Yukon. For these reasons and in agreement with other literature ([10]) we can propose that the deamidation of really old samples (at least a 100-years-old) is mainly dependent on the burial conditions and/or thermal age, and cannot serve as a precise "molecular clock".

Collagen was analysed by two mass spectrometry instruments: high-resolution Q-TOF and ion-trap. Results obtained by these methods were the same as it is demonstrated at Fig. 3 on the spectra of peptide GPSGPQGPGGPPGPK from collagen type I, chain alpha 1 (peptide with deamidated glutamine—G*).

**Acknowledgments** This work was supported by the Czech Science Foundation (Grants Nos. P206/12/0453 and 13-17224S). It also received support as part of the long-term conceptual development of the research organization RVO:67985823.

## References

1. Asara JA, Schweitzer MH, Freimark LM, Philips M, Cantley LC (2007) Science 316:280–285
2. Janko M, Zink A, Gigler AM, Heckl WM, Stark RW (2010) Proc R Soc B 277:2301–2309
3. Robinson NE, Robinson AB (2004) Molecular clocks: deamidation of asparaginyl and glutaminyl residues in peptides and proteins. Althouse Press, Cave Junction
4. Robinson NE, Robinson AB (2001) PNAS 98:12409–12413
5. Hurtado PP, O'Connor PB (2012) Anal Chem 84:3017–3025
6. Buckley M, Walker A, Ho SYW, Yang Y, Smith C, Ashton P, Oates JT, Koon H, Penkman K, Elsworth B, Ashford D, Solazzo C, Andrews P, Strahler J, Shapiro B, Ostrom P, Gandhi H, Miller W, Raney B, Zylber MI, Gilbert MTP, Prigodich RV, Ryan M, Rijsdijk KF, Janoo A, Collins MJ (2008) Comment on "Protein sequences from Mastodon and *Tyrannosaurus rex* revealed by mass spectrometry". Science 319:33
7. Asara JA, Schweitzer MH (2008) Response to comment on protein sequences from mastodon and *Tyrannosaurus rex* revealed by mass spectrometry. Science 319:33
8. Orlando L, Ginolhac A, Zhang G, Froese D, Albrechtsen A, Stiller M, Schubert M, Cappellini E, Petersen B, Moltke I, Johnson PLF, Fumagalli M, Vilstrup JT, Raghavan M, Korneliussen T, Malaspinas A-S, Vogt J, Szklarczyk D, Kelstrup CD, Vinther J, Dolocan A, Stenderup J, Velazquez AMV, Cahill J, Rasmussen M, Wang X, Min J, Zazula GD, Seguin-Orlando A, Mortensen C, Magnussen K, Thompson JF, Weinstock J, Gregersen K, Røed KH, Eisenmann V, Rubin CJ, Miller DC, Antczak DF, Bertelsen MF, Brunak S, Al-Rasheid KAS, Ryder O, Andersson L, Mundy J, Krogh A, Gilbert MTP, Kjær K, Sicheritz-Ponten T, Jensen LJ, Olsen JV, Hofreiter M, Nielsen R, Shapiro B, Wang J, Willerslev E (2013) Nature 499:74–78
9. Cappellini E, Jensen LJ, Szklarczyk D, Ginolhac A, da Fonseca RAR, Stafford TW Jr, Holen SR, Collins MJ, Orlando L, Willerslev E, Gilbert MTP, Olsen JV (2012) J Prot Res 11:917–926
10. van Doorn NL, Wilson J, Hollund H, Soressi M (2012) Collins MJ. Rapid Commun Mass Spectrom 26:2319–2327
11. Wadsworth C, Buckley M (2014) Rapid Commun Mass Spectrom 28:605–615