

Proteins and their modifications in a medieval mummy

Ivan Mikšík,^{1*} Pavla Sedláková,¹ Stasis Pataridis,¹ Federica Bortolotti,² and Rossella Gottardo²

¹Institute of Physiology, Academy of Sciences of the Czech Republic, Prague 14220, Czech Republic

²Department of Diagnostic and Public Health, Unit of Forensic Medicine, University of Verona, Verona, Italy

Received 11 May 2016; Accepted 17 August 2016

DOI: 10.1002/pro.3024

Published online 20 August 2016 proteinscience.org

Abstract: Proteins and their modifications of the natural mummy of Cangrande della Scala (Prince of Verona, Northern Italy, 1291–1329) were studied. The nano-LC-Q-TOF analysis of samples of rib bone and muscle from the mummy showed the presence of different proteins including Types I, III, IV, V, and XI collagen, hemoglobin (subunits alpha and beta), ferritin, biglycan, vitronectin, prothrombin, and osteocalcin. The structure of Type I and Type III collagen was deeply studied to evaluate the occurrence of modifications in comparison with Type I and Type III collagen coming from tissues of recently died people. This analysis showed high percentage of asparaginyl and glutaminyl deamidation, carbamylation and carboxymethylation of lysine, as well as oxidation and dioxidation of methionine. The most common reaction during the natural mummification process was oxidation—the majority of lysine and proline of collagen Type I was hydroxylated whereas methionine was oxidated (oxidated or dioxidated). To the best of our knowledge, this is the first study which reports the protein profile of a natural mummified human tissue and the first one which describes the carbamylation and carboxymethylation of lysine in mummified tissues.

Keywords: mummy; collagen; protein modification; deamidation; carbamylation; carboxymethylation

Introduction

Collagens are a family of extracellular matrix proteins that play a dominant role in maintaining the structure of various tissues. Collagens are the most abundant proteins in the human body, constituting approximately 30% of its protein mass. The origin of name collagen comes from the Greek κόλλα (kólla), meaning “glue” and suffix -γέν, -gen, denoting “producing”—it refers to the obsolete process of boiling the skin and sinews of animals to obtain glue. So far at least 28 collagen types have been described. The most abundant collagen forms extracellular fibrils or network-like structures. Fibril-

forming collagens represent a set of at least nine different polypeptide chains which constitute the molecular species of Type I–III, V, XI, XXIV, and XXVII collagens. Type I collagen is usually a heterotrimeric molecule composed of two $\alpha 1(I)$ chains and one $\alpha 2(I)$ chain [$\alpha 1(I)$]₂ $\alpha 2(I)$. Type I collagen is the dominant type of collagen and it is frequently associated to other collagen types such as Type III and Type V collagen.^{1–4} The Type I and III collagen show a triple helix structure, which is stable and durable. Type I collagen shows a so rigid and stable structure that it has been described in prehistoric samples, such as in the fossilized bones of *Tyrannosaurus rex* (68-million-year old) and *Mastodon* (0.16–0.6-million-year old) as well as in mammoth skull (0.1–0.3-million-year old).⁵ However, the possibility of contamination of million-years old samples is widely discussed. Excellent preservation of collagen has been also described in mummified human tissue. Indeed Type I collagen has been found in the skin of

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: Czech Science Foundation; Grant number: 15-01948S.

*Correspondence to: Ivan Mikšík, Institute of Physiology, Academy of Sciences of the Czech Republic, Videnska 1083, 14220 Prague, Czech Republic. E-mail: ivan.miksik@fgu.cas.cz

the 5300-year-old Tyrolean Iceman, whose body was reported to have been undergone a natural process of mummification through freeze-drying.⁶

Deamidation of asparaginyl (Asn) and glutaminyl (Gln) residues in protein(s) is a widely occurring posttranslational modification and the progress of this process was proposed as a “molecular clock.”⁷ Taking into consideration its endurance, collagen could be used for the study of this phenomenon, even if the high order structure of this protein has been reported to suppress deamidation.⁸ The formation of hydroxylysine glycosides is another posttranslational modification of ancient collagen which has been reported.⁹

Cangrande della Scala was Prince of Verona (Northern Italy) in the 14th century (1291–1329). He was a successful warrior and autocrat, being also well known as the leading patron of the poet Dante Alighieri. After death, his body was buried in a stone sarcophagus placed on the façade of the Church S. Maria della Scala, in the center of Verona. In 2004, the body underwent an archaeopathological investigation which found it very well preserved because of the occurrence of a natural mummification process.

In a previous study on rib bone and muscle samples of Cangrande della Scala, based on the use of HPLC-Ion Trap MS and nano-LC-Q-TOF, we demonstrated a high deamidation percentage of asparaginyl and glutaminyl residues in both Type I and Type III collagen.¹⁰

The aim of this work was to enlarge the study of the protein profile of the mummified tissues of the Prince using a new nanoLC-Q-TOF method to identify new modifications of collagen, in addition to that reported in the previous study, and to verify which other proteins, besides collagen, constituted mummified tissues.

Results and Discussion

Type I collagen was the most abundant protein in the rib bone sample; whereas, Type I and III collagen were the most abundant proteins in muscle sample. In comparison to tissues from the recently dead subject, a very high percentage (up to 100%) of asparaginyl and glutaminyl residues were deamidated in mummified tissues, as already described by our group.¹⁰ Moreover, new modifications were identified, such as carboxymethylation and carbamylation of lysine and oxidation and dioxidation of methionine, as reported in Table I.

Figure 1 shows the coverage of Type I collagen, chain $\alpha 1$. Mass spectra of individual modifications are reported in Figure 2. The high sequence coverage of Type I collagen is an agreement with the well described rigid structure of its molecule and then with its high “survival” in fossils where proteome

degradation is high.¹¹ If we compare the results of coverage for modern material (i.e., control) the discrepancy (lower coverage) with mummified material can be caused by the rigid structure of natural collagen resulting in a high resistance to trypsin treatment. Collagen from mummy was probably partly damaged and so it is more susceptible to enzymatic cleavage. Among the fossil collagen analyses which are reported in the literature, we can mention that performed on a horse (*Equus*) bone from early Middle Pleistocene¹² as well as that performed on a Mammoth femur (*Mammuthus primigenius*) from Pleistocene.¹³ The data reported in these articles, in agreement with the study of van Doorn *et al.*,¹⁴ show that deamidation of collagen in fossil material is more influenced by the burial conditions and thermal age than by chronological age. These results are in agreement with that reported in our previous work where we concluded that the deamidation process is not suitable for a precise “molecular clock.”¹⁰

The most common reaction (beyond asparagine and glutamine deamidation) we observed in the mummified tissues was oxidation, being the majority of Type I collagen lysine and proline residues hydroxylated, whereas methionine was oxidated (oxidated or dioxidated).

Additional modifications were represented by carboxymethylation and carbamylation of lysine which looked uncommon in tissues of the recently dead subject.

Carboxymethylation is a well-known reaction occurring during glycation (nonenzymatic glycosylation) between amine group of lysine and oxo group of sugars or various aldehydes. As it has been previously described, collagen accumulates reactive metabolites through reactions that are not regulated by enzymes.^{15,16} A typical example of these nonenzymatic changes is glycation (e.g., Maillard reaction), resulting from the reaction of the oxo-group of sugars with the free amino group of lysine and arginine. The initial labile Schiff base and Amadori products undergo a series of rearrangement, dehydration, and fragmentation reactions to produce more complex and irreversibly covalently crosslinked structures.^{17,18} One of the best known monotopic modification is N^ε-(carboxymethyl)-lysine (CML). Some years ago, our group performed a *in vitro* study on the modification of Type I collagen (obtained from bovine Achilles tendon) using glucose and ribose, demonstrating that reactive lysines are at locations 504 and 1032, respectively 504, 519, 750, 861, and 1032.¹⁹ The presence of CML was also evaluated in this study on the mummified samples. Table II reports CML location numbers, which are different from that identified in the previous study since the structure

Table I. Percentage of Modified Aminoacids in Collagen Chains Type I (Chains $\alpha 1$ and 2) and Type III in Muscle and Rib Tissues and Their Coverage

Modification	CO1 α 1		CO1 α 2		CO3 α 1	
	Mummy (%)	Control (%)	Mummy (%)	Control (%)	Mummy (%)	Control (%)
Muscle						
Lysine						
Hydroxylation	94	52	81	50	72	43
Carboxymethyl	35	5	38	0	3	0
Carbamyl	24	0	19	0	14	0
Proline						
Hydroxylation	97	62	88	57	83	64
Methionine						
Dioxidation	86	60	80	100	88	0
Oxidation	100	40	100	100	88	50
Asparagine						
Deamidation	100	0	100	0	57	0
Glutamine						
Deamidation	100	2	94	10	87	0
Rib						
Lysine						
Hydroxylation	92	79	88	67	17	14
Carboxymethyl	32	10	27	0	33	29
Carbamyl	27	14	19	0	17	0
Proline						
Hydroxylation	96	86	89	75	55	49
Methionine						
Dioxidation	86	43	60	33	50	—
Oxidation	100	71	100	33	100	—
Asparagine						
Deamidation	100	40	92	0	100	0%
Glutamine						
Deamidation	100	30	95	6	50	20
Coverage of collagen sequence						
Tissue	CO1 α 1		CO1 α 2		CO3 α 1	
	Mummy	Control	Mummy	Control	Mummy	Control
Muscle	92	57	86	46	73	21
Rib	97	80	91	75	14	19

of human collagen is different from that of bovine collagen.

Even if the formation of CML is mainly described during glycooxidation, it has been also described during lipoxidation (i.e., oxidation of lipids).^{20–22} We can suppose that in the case of mummy the CML origins from lipids. The reason of this assumption is a relative high occurrence of lipids in human body (in comparison to a relative low presence of free sugars like glucose or ribose) and the widespread oxidation reactions during mummification process.

It is important to note that our previous study was based on *in vitro* experiments (7-days incubation of collagen Type I ($\alpha 1$) with reactive sugars, glucose, and ribose) and we found only two, respectively, five reactive lysines at the structure. In the case of the analyzed mummy, where a “675 years natural incubation” occurred we found 18 reactive lysines in Type I ($\alpha 1$) collagen, 14 reactive lysines in

the Type I ($\alpha 2$) collagen and 6 reactive lysines in Type III ($\alpha 1$) collagen (see Table II).

Concerning carbamylation, it constitutes a post-translational modification of proteins/amino acids which has been described also in tissues at physiological conditions, resulting from different pathways *in vivo*. Among them, the most important is the nonenzymatic reaction between isocyanic acid, a decomposition product of urea, and either the N-terminus or the ϵ -amino group of lysine residues. Isocyanic acid levels, while low *in vivo*, are in equilibrium with urea and are thus increased in chronic and end-stage renal diseases. An alternative pathway involves the leukocyte heme protein myeloperoxidase, which catalyzes the oxidation of thiocyanate in the presence of hydrogen peroxide, producing isocyanate at inflammation sites. Notably, plasma thiocyanate levels are increased in smokers, and leukocyte-driven protein carbamylation occurs both within human and animal atherosclerotic plaques, as well as on plasma

1	mfsfvdlrll	lllaatallt	hgqeegqveg	qdedippitc	vqnglryhdr	50
51	dvwkpepcrri	cvcdngkvlc	ddvicdetkn	cpgaevpege	ccpvcpdgse	100
101	sptdqettgv	egpkgdtgpr	gprgpagppg	rdgipgqpgl	pgppgppgpp	150
151	gppglggnfa	pQLSYGYDEK	STGGISVPGP	MGPSGPRGLP	GPPGAPGPQG	200
201	FQGPPEPGE	PGASGPMGPR	GPPGPPGKNG	DDGEAGKPGR	PGERGPFPQO	250
251	GARGLPGTAG	LPGMKGHRGE	SGLDGAKGDA	GPAGPKGEPG	SPGENGAPGO	300
301	MGPRGLPGER	GRPGAPGPAG	ARGNDGATGA	AGPPGPTGPA	GPPGFPGAVG	350
351	AKGEAGPQGP	RGSEGPQGVV	GEPGPPGPAG	AAGPAGNPGA	DGQPGAKGAN	400
401	GAPGIAGAPG	FPGARGPSGP	QGGGPPGPK	GNSGEPGAPG	SKGDTGAKGE	450
451	PGPVGVQGP	GPAGEEGKRG	ARGEPPPTGL	PGPPGERGGP	GSRGFPDADG	500
501	VAGPKGPAGE	RGSPPGAPPK	GSPGEAGRPG	EAGLPGAKGL	TGSPGSPGPD	550
551	GKTGPPGPAG	QDGRFPPPP	PGARGQAGVM	GFPGPKGAAG	EPGKAGERGV	600
601	PGPPGAVGPA	GKDGEAGAOG	PPGPAGPAGE	RGEQGPAGSP	GFQGLPGPAG	650
651	PPGEAGKPGE	QGVPGDLGAP	GPSGARGERG	FPGERGVQGP	PGPAGPRGAN	700
701	GAPGNDGAKG	DAGAPGAPGS	QGAPGLOQMP	GERGAAGLPG	PKGDRGDAGP	750
751	KGADGSPGKD	GVRGLTGPIG	PPGPAGAPGD	KGESGSPGPA	GPTGARGAPG	800
801	DRGEPGPPGP	AGFAGPPGAD	GQPGAKGEPG	DAGAKGDAGP	PGPAGPAGPP	850
851	GPIGNVGAAPG	AKGARGSAGP	PGATGFPGAA	GRVGPSPGSG	NAGPPGPPGP	900
901	AGKEGGKGP	GETGPAGRPG	EVGPPGPPGP	AGEKGSFGAD	GPAGAPGTPG	950
951	PQGIAGQRGV	VGLPGQRGER	GFPLPGPSG	EPGKQPSGA	SGERGPSPGM	1000
1001	GPPGLAGPPG	ESGREGAPGA	EGSPGRDGSP	GAKGDRGETG	PAGPPGAPGA	1050
1051	PGAPGPVGP	GKSGDRGETG	PAGPTGPVGP	VGARGPAGPQ	GPRGDKGETG	1100
1101	EQGDRGIKGH	RFGSGLQGGP	GPPGSPGEGG	PSGASGPAGP	RGPPGSAGAP	1150
1151	GKDGLNGLPG	PIGPPGPRGR	TGDAGPVGPP	GPPGPPGPPG	PPSAGFDFSF	1200
1201	LPQPPQEK KAH	DGGRYYRAdd	anvvrdrdle	vdttlkslsq	qienirspeg	1250
1251	srknpartcr	dlkmchsdwk	sgeywidpnq	gcnldaikvf	cnmetgetcv	1300
1301	yptqpsvaqk	nwyisknpkd	krhvwfgesm	tdgfgfeygg	qgsdpadvai	1350
1351	qltflrlmst	easqnityhc	knsvaymdqq	tgnlkkalll	qgsneieira	1400
1401	egnsrftysv	tvdgctsthtg	awgktvieyk	ttktsrlpii	dvapldvgap	1450
1451	dqefgfdvvgp	vcfl				1500

Figure 1. Sequence of the human collagen, Alpha 1, Type I. The matched peptides are in yellow; propeptides are highlighted by lower letters; asparaginyl and glutaminyl residues are in bold when deamidated ones are underlined, as well as proline and lysine are in bold when hydroxylated ones are underlined.

proteins. Protein carbamylation is considered a hallmark of molecular aging and is implicated in many pathological conditions. Protein carbamylation levels have emerged as a particularly strong predictor of both prevalent and incident cardiovascular disease risk. Recent studies also suggest that protein carbamylation may serve as a potential therapeutic target for the prevention of atherosclerotic heart disease.²³

In the case of the mummified tissues, we suppose that mechanism of nonenzymatic reaction between isocyanic acid (as decomposition product of urea) ϵ -amino group of lysine residues plays a role in the carbamylation of lysines.

Hydroxylation of lysine and proline is a common process for collagen and it is important for the formation of the triple helix structure of this molecule.

These hydroxylation reactions are enzymatically controlled (by two enzymes: prolyl-4-hydroxylase and lysyl-hydroxylase). There are many sites for this modification and one of the most used methods for determination/quantification of collagen in tissues was based on the determination of hydroxyproline. In the tissues of mummy, we found higher level of hydroxylation of proline and lysine (in comparison to control) which is probably caused by nonenzymatic oxidation more likely than enzymatic reaction (Table I). The nonenzymatic hydroxylation of proline and lysine was described by Trelstad *et al.* at 1981²⁴ who demonstrated that reduced oxygen derivatives can hydroxylate both free and polypeptide-bound proline and lysine. Now it is well described idea that reactive radical HO^{*} in presence of O₂ produces

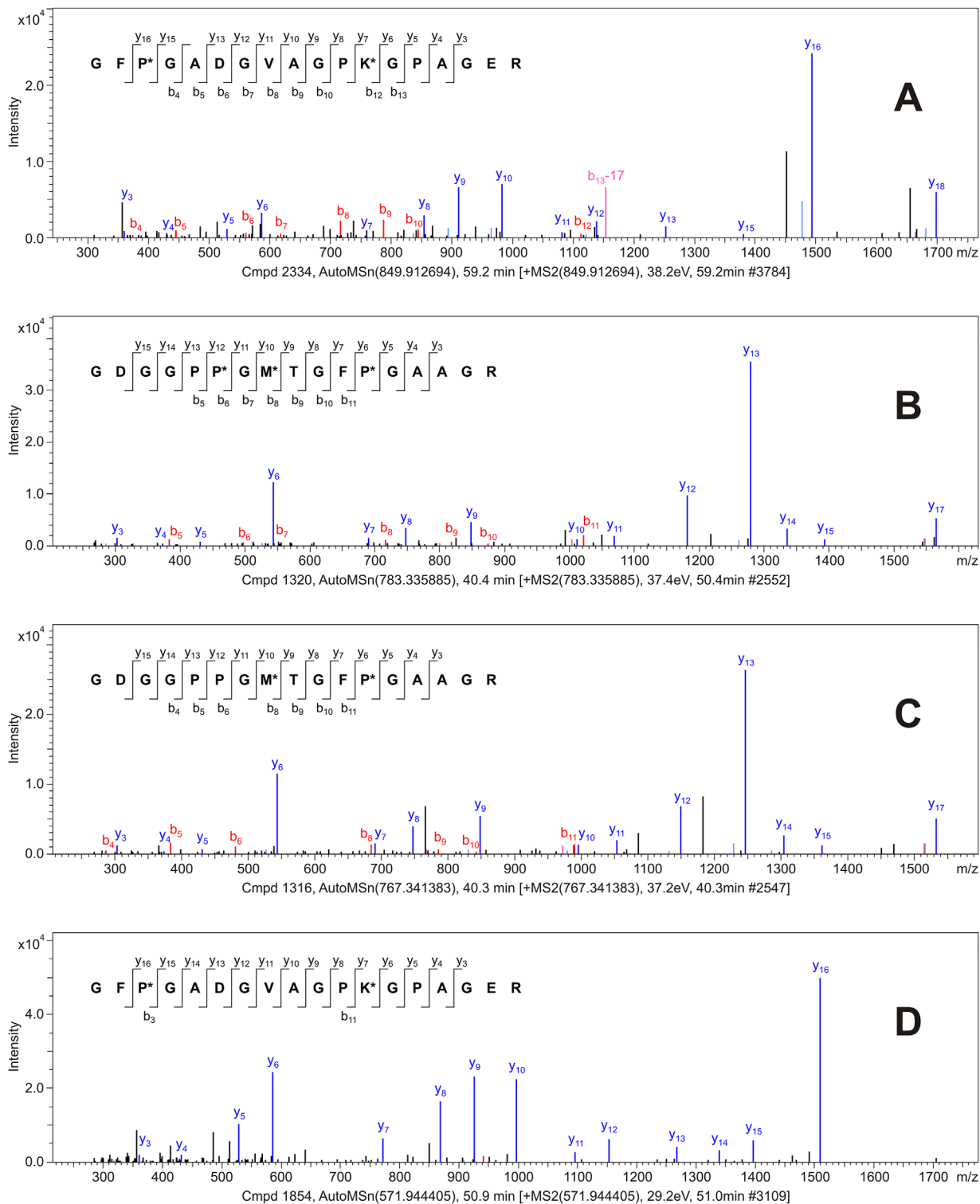


Figure 2. Mass spectra of identified modifications. (A) CO1A1 (rib)—seq. 494–511, carbamyl (K): 12, hydroxyl (P): 3; (B) CO1A2 (rib)—seq. 778–794, dioxidation (M): 8, hydroxyl (P): 6, 12; (C) CO1A2 (rib)—seq. 778–794, oxidation (M): 8, hydroxyl (P): 12; and (D) CO1A1 (muscle)—seq. 494–511, carboxymethyl (K): 12, hydroxyl (P): 3.

hydroxyprolines (in the case of proline) as well as hydroxylysines (in the case of lysine).^{25–27} However, it is possible that quantification of collagen in mummy based on the old method for its determination by hydroxyproline content would lead to false results.

Besides these modifications mainly observed in Type I and Type III collagen, we verified the presence in the mummy tissues of other types of collagen (IV, V, and XI) and of some other proteins, such as hemoglobin (subunits alpha and beta), ferritin, biglycan, vitronectin, prothrombin, and osteocalcin (see

Table II. Sites of Carboxymethylation (CML) and Carbamylation (CL) of Collagens Lysines in Mummy

Collagen type I ($\alpha 1$)			Collagen type I ($\alpha 2$)			Collagen type III ($\alpha 1$)		
#AA	Muscle	Rib	#AA	Muscle	Rib	#AA	Muscle	Rib
228	CML	CML, CL	140	CML, CL	CML, CL	436	CL	
237	CML	CL	149	CML		440	CL	
265	CML		177		CL	662		CML
277	CML	CML, CL	309		CML, CL	674	CML, CL	CL
286	CML	CML, CL	464	CML, CL	CML	833	CL	
397	CML, CL	CML, CL	498	CML, CL		977		CML
430		CML	506		CML			
505	CML, CL	CML, CL	510	CML				
520	CML		543	CL	CML, CL			
538	CML, CL	CML, CL	621	CML	CML			
586	CL		744	CML	CML			
594		CML	747	CML				
709		CML	815	CML, CL	CL			
781	CML, CL	CML, CL	846	CML				
934	CL							
1062	CL	CL						
1096	CML	CML						
1152	CML, CL	CML, CL						

Table III). The comparison between mummified tissues and recent tissues showed some differences, as it could be supposed, being Type I, II, III, and XI present in both type of samples. Proteins different from collagen showed lower sequence coverage (lower amount of observed peptides) in comparison to Type I and III collagens. It is probably caused by lower “structural rigidity” of these proteins.

Conclusion

The protein analysis of the tissues of the Cangrande’s mummy showed that the majority of lysine and proline of type I collagen were hydroxylated, methionine was oxidated (oxidated or dioxidated) and many lysine residues were carbamylated and carboxymethylated.

To the best of our knowledge this is the first study which reports the protein profile of a natural mummified human tissue and the first one which shows that during natural mummification process the most common modification (besides deamidation of asparagine and glutamine) is oxidation. In addition, this is the first time that carbamylation and carboxymethylation of lysine are described in mummified tissues

Materials and Methods

Chemicals

Trypsin (Type IX-S from porcine pancreas, E.C. 3.4.21.4, 15,450 units per mg), ammonium bicarbonate and acetonitrile (HPLC-MS grade) were obtained

Table III. Proteins Found in Muscle and Rib Bone Samples of Mummy

Accession number	Protein	Scores		#Peptides		SC (%)	
		Muscle	Rib	Muscle	Rib	Muscle	Rib
P02452	Collagen alpha-1(I) chain	51,797	44,198	1084	901	92 ^a	97 ^a
P08123	Collagen alpha-2(I) chain	33,054	32,433	604	531	86 ^a	91 ^a
P02461	Collagen alpha-1(III) chain	12,063	2447	191	12	73 ^a	14 ^a
P13942	Collagen alpha-2(XI) chain	2980	2139	16	15	40	39
P20908	Collagen alpha-1(V) chain	2771	2237	8	10	40	44
P05997	Collagen alpha-2(V) chain	1622	—	6	—	36	—
P08572	Collagen alpha-2(IV) chain	1544	—	6	—	30	—
P69905	Hemoglobin subunit alpha	385	165	7	3	57	29
P02792	Ferritin light chain	312	—	3	—	29	—
F8W6P5	Hemoglobin subunit beta (fragment)	282	—	4	—	61	—
P21810	Biglycan	—	556	—	9	—	30
P04004	Vitronectin	—	384	—	5	—	27
P00734	Prothrombin	—	367	—	8	—	17
P02818	Osteocalcin	—	315	—	3	—	19
P68871	Hemoglobin subunit beta	—	230	—	3	—	59
P36955	Pigment epithelium-derived factor	—	212	—	3	—	18

Scores, Mascot score; SC, sequence coverage.

^a Coverage recalculated using collagen sequence without telopeptides.

from Sigma (St. Louis, MO, USA). Two-mercaptoethanol and cyanogen bromide (CNBr) were from Merck (Darmstadt, Germany). All solutions were prepared in MilliQ water (Millipore, Bedford, MA, USA). ProteaseMAX were purchased from Promega (Madison, WI, USA). Empore Octadecyl C18 Extraction disks were purchased from Supelco (Bellefonte, PA, USA).

Sample preparation

Rib bone (500 mg) and muscle samples (500 mg), which were collected during the autopsy of the mummy, were sealed and stored at room temperature until analysis.

To perform a comparison study, samples of rib bone (500 mg) and muscle (500 g) were collected from a human body which underwent forensic autopsy in 2013.

All the samples (including that from the mummy and that from the subject recently dead) were brought up in 0.1% ProteaseMax in 50 mM ammonium bicarbonate (400 μ L/mg tissue) and vortexed at room temperature for 1 h.

Trypsin procedure: the samples were incubated at 37°C in pH 7.8 ammonium bicarbonate buffer (20 mmol/L) added with trypsin (1:50 enzyme:substrate ratio). After 3 h, the cleavage was stopped by acidification by acetic acid.

After trypsin cleavage the samples were purified by StageTips using Empore C18 Extraction disks according to the published protocol.²⁸

Analysis of tryptic digests with LC-MS/MS. The nano-HPLC apparatus used for protein digest analysis was a Proxeon Easy-nLC (Proxeon, Odense, Denmark). It was coupled to a ultrahigh resolution MaXis Q-TOF (quadrupole—time of flight) mass spectrometer (Bruker Daltonics, Bremen, Germany) by nanoelectrosprayer. The nLC-MS/MS instruments were controlled by the software packages HyStar 3.2 and micrOTOF-control 3.0. The data were collected and manipulated with the software packages ProteinScape 3.0 and DataAnalysis 4.0 (Bruker Daltonics).

Five microliters of the peptide mixture were injected into an NS-AC-12dp3-C18 Biosphere C18 column (particle size: 3 μ m, pore size: 12 nm, length: 200 mm, and inner diameter: 75 μ m) with a NS-MP-10 Biosphere C18 precolumn (particle size: 5 μ m, pore size: 12 nm, length: 20 mm, and inner diameter: 100 μ m), both manufactured by NanoSeparations (Nieuwkoop, Holland).

The separation of 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 7% B at 5 min 30% B at 180 min.

The next step was a gradient elution to 50% B in 10 min and then a gradient to 100% B in 10 min. Finally, the column was eluted with 100% B for 20 min. Equilibration between the runs was achieved by washing the column with 5% mobile phase B for 10 min. The flow rate was 0.20 μ L/min and the column was held at ambient temperature (25°C).

On-line nano-electrospray ionization (easy nano-ESI) in positive mode was used. The ESI voltage was set at +4.5 kV, scan time: 1.3 Hz. Operating conditions: drying gas (N₂): 1 L/min; drying gas temperature: 160°C; nebulizer pressure: 40 kPa. Experiments were performed by scanning from 100 to 2200 *m/z*. The reference ion used (internal mass lock) was a monocharged ion of C₂₄H₁₉F₃₆N₃O₆P₃ (*m/z* 1221.9906). Mass spectra corresponding to each signal from the total ion current chromatogram were averaged, enabling an accurate molecular mass determination. All LC-MS and LC-MS/MS analyses were done in duplicate.

Database searching. Data were processed using ProteinScape software. Proteins were identified by correlating tandem mass spectra to the Uniprot database, using the MASCOT searching engine (<http://www.matrixscience.com>). The taxonomy was restricted to *Homo sapiens* to remove protein identification redundancy. Trypsin (or semitrypsin) 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. Lysines were assumed to be hydroxylated, carboxymethylated, carbamylated, galactosylated, and glucosylgalactosylated, prolines were assumed to be hydroxylated, methionine was allowed to be oxidated and dioxidated whereas asparagine and glutamine deamidated. All these possible modifications were set to be variable. In the first experiments, these modifications were searched by nonspecific searching of modifications (errors). 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>). The amino acid(s) were considered to be modified when the modification was found to occur at least once within the structure of proteins.

Conflict of Interest

We declare that we have no conflict of interest in relation to this article.

References

1. Pataridis S, Eckhardt A, Mikulikova K, Sedlakova P, Miksik I (2008) Identification of collagen types in tissues using HPLC-MS/MS. *J Sep Sci* 31:3483–3488.
2. Pataridis S, Eckhardt A, Mikuliková K, Sedláková P, Miksik I (2009) Determination and quantification of collagen types in tissues using HPLC-MS/MS. *Curr Anal Chem* 5:316–323.
3. Myllyharju J, Kivirikko KI (2004) Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends Genet* 20:33–43.
4. Prockop DJ, Kivirikko KI (1995) Collagens: molecular biology, diseases, and potentials for therapy. *Annu Rev Biochem* 64:403–434.
5. Asara JM, Schweitzer MH, Freimark LM, Phillips M, Cantley LC (2007) Protein sequences from mastodon and *Tyrannosaurus rex* revealed by mass spectrometry. *Science* 316:280–285.
6. Janko M, Zink A, Gigler AM, Heckl WM, Stark RW (2010) Nanostructure and mechanics of mummified type I collagen from the 5300-year-old Tyrolean Ice-man. *Proc R Soc B* 277:2301–2309.
7. Robinson NE, Robinson AB (2004) Molecular clocks: deamidation of asparaginyl and glutaminyl residues in peptides and proteins. Cave Junction, OR: Althouse Press.
8. Hurtado PP, O'Connor PB (2012) Deamidation of collagen. *Anal Chem* 84:3017–3025.
9. Hill RC, Wither MJ, Nemkov T, Barrett A, D'Alessandro A, Dzieciatkowska M, Hansen KC (2015) Preserved proteins from extinct *Bison latifrons* identified by tandem mass spectrometry; hydroxylysine glycosides are a common feature of ancient collagen. *Mol Cell Proteomics* 14:1946–1958.
10. Miksik I, Sedlakova P, Pataridis S, Bortolotti F, Gottardo R, Tagliaro F (2014) Prince Cangrande's collagen: study of protein modification on the Mummy of the Lord of Verona, Italy (1291-1329 AD). *Chromatographia* 77:1503–1510.
11. Wadsworth C, Buckley M (2014) Proteome degradation in fossils: investigating the longevity of protein survival in ancient bone. *Rapid Commun Mass Spectrom* 28:605–615.
12. 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, Korneliusson T, Malaspina 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, Roed 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, Kjaer K, Sicheritz-Ponten T, Jensen LJ, Olsen JV, Hofreiter M, Nielsen R, Shapiro B, Wang J, Willerslev E. (2013) Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* 499:74–78.
13. Cappellini E, Jensen LJ, Szklarczyk D, Ginolhac A, da Fonseca RAR, Stafford TW, Jr, Holen SR, Collins MJ, Orlando L, Willerslev E, et al. (2012) Proteomic analysis of a Pleistocene mammoth femur reveals more than one hundred ancient bone proteins. *J Proteome Res* 11:917–926.
14. van Doorn NL, Wilson J, Hollund H, Soressi M, Collins MJ (2012) Site-specific deamidation of glutamine: a new marker of bone collagen deterioration. *Rapid Commun Mass Spectrom* 26:2319–2327.
15. Deyl Z, Miksik I (2000) Advanced separation methods for collagen parent [alpha]-chains, their polymers and fragments. *J Chromatogr B Biomed Sci Appl* 739:3–31.
16. Miksik I, Deyl Z (1997) Post-translational non-enzymatic modification of proteins II. Separation of selected protein species after glycation and other carbonyl-mediated modifications. *J Chromatogr B Biomed Sci Appl* 699:311–345.
17. Niwa T (1997) Mass spectrometry in the search for uremic toxins. *Mass Spectrom Rev* 16:307–332.
18. Yim MB, Yim HS, Lee C, Kang SO, Chock PB, Protein glycation—creation of catalytic sites for free radical generation. In: Park SC, Hwang ES, Kim HS, Park WY, Eds. (2001) Healthy aging for functional longevity: molecular and cellular interactions in senescence, *Annals of the New York Academy of Sciences*, Vol. 928, pp 48–53.
19. Mikulikova K, Eckhardt A, Pataridis S, Miksik I (2007) Study of posttranslational non-enzymatic modifications of collagen using capillary electrophoresis/mass spectrometry and high performance liquid chromatography/mass spectrometry. *J Chromatogr A* 1155:125–133.
20. Fu MX, Requena JR, Jenkins AJ, Lyons TJ, Baynes JW, Thorpe SR (1996) The advanced glycation end product, Nepsilon-(carboxymethyl)lysine, is a product of both lipid peroxidation and glycoxidation reactions. *J Biol Chem* 271:9982–9986.
21. Saxena AK, Saxena P, Wu X, Obrenovich M, Weiss MF, Monnier VM (1999) Protein aging by carboxymethylation of lysines generates sites for divalent metal and redox active copper binding: relevance to diseases of glycoxidative stress. *Biochem Biophys Res Commun* 260:332–338.
22. Miyata T, Inagi R, Asahi K, Yamada Y, Horie K, Sakai H, Uchida K, Kurokawa K (1998) Generation of protein carbonyls by glycoxidation and lipoxidation reactions with autoxidation products of ascorbic acid and polyunsaturated fatty acids. *FEBS Lett* 437:24–28.
23. Verbrugge FH, Tang WHW, Hazen SL (2015) Protein carbamylation and cardiovascular disease. *Kidney Int* 88:474–478.
24. Trelstad RL, Lawley KR, Holmes LB (1981) Nonenzymatic hydroxylations of proline and lysine by reduced oxygen derivatives. *Nature* 289:310–312.
25. Davies MJ, Fu SL, Wang HJ, Dean RT (1999) Stable markers of oxidant damage to proteins and their application in the study of human disease. *Free Rad Biol Med* 27:1151–1163.
26. Xu G, Chance MR (2007) Hydroxyl radical-mediated modification of proteins as probes for structural proteomics. *Chem Rev* 107:3514–3543.
27. Dean RT, Fu S, Stocker R, Davies MJ (1997) Biochemistry and pathology of radical-mediated protein oxidation. *Biochem J* 324:1–18.
28. Rappsilber J, Mann M, Ishihama Y (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2:1896–1906.