OMICS A Journal of Integrative Biology
Volume 21, Number 7, 2017
© Mary Ann Liebert, Inc.
DOI: 10.1089/omi.2017.0067

Research Article

Omics for Precious Rare Biosamples: Characterization of Ancient Human Hair by a Proteomic Approach

Margaux Fresnais,[1] Pascale Richardin,[2] Marcela Sepúlveda,[3] Emmanuelle Leize-Wagner,[1] and Armelle Charrié-Duhaut[1]

Abstract

Omics technologies have far-reaching applications beyond clinical medicine. A case in point is the analysis of ancient hair samples. Indeed, hair is an important biological indicator that has become a material of choice in archeometry to study the ancient civilizations and their environment. Current characterization of ancient hair is based on elemental and structural analyses, but only few studies have focused on the molecular aspects of ancient hair proteins–keratins–and their conservation state. In such cases, applied extraction protocols require large amounts of raw hair, from 30 to 100 mg. In the present study, we report an optimized new proteomic approach to accurately identify archeological hair proteins, and assess their preservation state, while using a minimum of raw material. Testing and adaptation of three protocols and of nano liquid chromatography–tandem mass spectrometry (nanoLC-MS/MS) parameters were performed on modern hair. On the basis of mass spectrometry data quality, and of the required initial sample amount, the most promising workflow was selected and applied to an ancient archeological sample, dated to about 3880 years before present. Finally, and importantly, we were able to identify 11 ancient hair proteins and to visualize the preservation state of mummy's hair from only 500 μg of raw material. The results presented here pave the way for new insights into the understanding of hair protein alteration processes such as those due to aging and ecological exposures. This work could enable omics scientists to apply a proteomic approach to precious and rare samples, not only in the context of archeometrical studies but also for future applications that would require the use of very small amounts of sample.

Keywords: hair characterization, keratins, mass spectrometry, mummy's hair, proteomics

Introduction

OMICS TECHNOLOGIES HAVE BEEN APPLIED in a variety of settings such as medicine and ecology. Moreover, during the past decade, interest in hair analysis has strongly grown, particularly of ancient samples, with a view to deciphering the impact of aging and environment on the human or the animal host. Hair, as a biomaterial, is, indeed, a powerful biological indicator, whose composition is known to represent the blood composition to a certain extent (Kempson and Lombi, 2011; Tobin, 2005). During fiber growth, compounds from blood interact with hair molecules, thus enabling time-resolved monitoring along the fiber after the exposition to a specific compound (Kamata et al., 2015; Kumtabtim et al., 2011).

Hair shaft can be of 15–150 μm in diameter, and it exhibits three tubular morphological layers: the medulla, the cortex, and the cuticle. The medulla is the innermost part of the fiber and is not always observable in human hair, according to the hair type. It consists of a narrow tube filled with air and loose cells that are mostly empty. The cortex enfolds the medulla and represents 80–90% of the hair mass. It is mainly composed of 18 hydrophobic fibrous proteins called keratins, which can be of two types: type I, the acidic forms, and type II, the basic forms. They are α-helical proteins that associate one with another into tetrameric cylinder-like filaments called Keratin Intermediate Filaments (KIFs).

Assemblies of these KIFs are called macrofibrils, and their cohesion is ensured by an amorphous sulfur-rich matrix made

[1]Laboratoire de spectrométrie de masse des interactions et des systèmes (LSMIS), Université de Strasbourg, CNRS, CMC UMR 7140, Strasbourg, France.
[2]Centre de recherche et de restauration des musées de France (C2RMF), Palais du Louvre, Paris, France.
[3]Laboratorio de Análisis e Investigación Arqueométricas y Laboratorio de Arqueologia y Paleoambiente, Instituto de Alta Investigación, Universidad de Tarapacá, Arica, Chile.

up of Keratin Associated Proteins (KAPs). The network of KIFs and KAPs is characterized by the presence of an important number of cysteine residues (Sulphur based amino acid), which contributes to the stability of hair fibers and their high resistance to degradations through the existence of inter-filament, inter-filament-matrix, or even inter-matrix disulfide bridges. On the surface, the cuticle constitutes the robust protective outer layer of the hair fiber. It is about 3–5 μm thick and is made up of several layers of overlaying scales that are joined one to another by an intercellular lipid-rich substance (Barthélémy, 2011; Swift and Bews, 1974; Wolfram, 2003).

In the past several years, hair has become a major medium for toxicology (Kales and Christiani, 2005; Kamata et al., 2015), environmental (Bencko, 2005; Kintz and Villain, 2005), and forensic studies (Haglund and Sorg, 1996; Kintz et al., 2007; Wilson et al., 2007), providing complementary information to urine (Hardy et al., 2015) and blood (Goodrich et al., 2016) analyses, and enabling monitoring over a longer time period than that offered by urine and blood, which can be of only a few hours or days (Hardy et al., 2015).

When it is well preserved thanks to extreme environmental conditions (aridity, extreme coldness, or anoxia), ancient hair also represents a material of choice in archeology (Wilson, 2005) for dating (Richardin et al., 2011, 2013), dietary (Aufderheide et al., 1994), health (Arriaza et al., 2010; Boston and Arriaza, 2009), environmental (Arriaza, 2005; Bartkus et al., 2011; Byrne et al., 2010; Fresnais et al., 2015; Hallégot et al., 2008), and cultural (Arriaza, 2005) studies. In archeometry, many research works focus on the study of ancient proteins from collagen-based materials, such as bones (Aufderheide et al., 1994; Buckley and Wadsworth, 2014; Collins et al., 2006) or skin (Bianucci et al., 2008), and from teeth (Buckley and Kansa, 2011; Farell et al., 2013), but hair also represents a good alternative or a complement to these other biological remnants that can be rare or whose sampling can be much more invasive. In that regard, obtaining a maximum of information from the minimum of raw archeomaterials is a crucial issue.

In archeometrical studies, characterization of ancient hair is mainly based on elemental (Laser Ablation-Inductively Coupled Plasma-Mass Spectrometry [LA-ICP-MS], X-Ray Fluorescence spectroscopy [XRF], and Particle-Induced X-ray Emission imaging [PIXE] imaging), structural (Fourier Transformed Infra-Red spectroscopy [FTIR] and X-Ray Diffraction [XRD]), and biological analyses (Bertrand et al., 2003, 2014; Lubec et al., 1987; Wilson et al., 2007) but only few research works focus on the molecular study of ancient hair proteins and their conservation state (Lubec et al., 1997) and in such cases, they often apply extraction protocols requiring important amounts of raw hair. Current research focuses not only on the optimization of protocols for sample preparation but also on developing new analytical techniques to reduce the initial amounts of raw archeomaterial without losing any important information.

Proteomic analysis of modern hair is well described in literature (Barthélémy, 2011; Barthélémy et al., 2012; Lee et al., 2006) and has led to the molecular characterization of cortical proteins (mainly keratins and KAPs). Our proteomic approach for the characterization of hair in the archeological context will, of course, be based on these studies but will also use specific protocols, adapted to be robust, reproducible, and applicable to small amounts of samples. The aim is to ascertain whether keratins and KAPs can still be identified in ancient hair and in that case, to verify whether it is possible to go further in their characterization and to detect the post-translational modifications (PTMs). These aspects are complex and challenging since, to the best of our knowledge, there are only few studies on the degradation of ancient hair proteins, and only few references for the search of PTMs specific to aging in the archeological context.

Materials and Methods

Materials and chemicals

Sinapic acid (SA), ammonium bicarbonate (NH_4HCO_3), and sodium dodecyl sulfate (SDS) were purchased from Fluka (Buchs, Switzerland). Iodoacetamide (IAM), DL-dithiothreitol (DTT), Trizma® hydrochloride (Tris-HCl, buffer solution for proteins), thiourea, sodium deoxycholate (DCO), proteomics grade trypsin, trichloroacetic acid (TCA), trifluoroacetic acid (TFA), formic acid (FA), dichloromethane (DCM), acetonitrile (ACN), and acetone were from Sigma-Aldrich (St. Louis, MO, USA). Urea was bought from Acros Organic (Geel, Belgium). Methanol (MeOH) and propan-2-ol (*i*PrOH) were supplied by Hipersolv CHROMANORM® (VWR Chemical Prolabo®, Radnor, PA, USA). Ethanol (EtOH) was from AnalaR Normapur® (VWR Chemical Prolabo). Ultrapure water was produced with a water purification Purelab® UHQ system (ELGA LabWater VEOLIA Water, Antony, France). Dialysis cartridges (Slide-A-Lyzer 3.5 kDa, 3–12 mL) were from Thermo Scientific Pierce (Thermo Fisher Scientific, Inc., Waltham, MA, USA), and SeP-Pak C18 1cc Solid Phase Extraction (SPE) cartridges (100 mg sorbent, particle size 55–105 μm, pore size 125 Å) were from Waters (Milford, MA, USA).

Modern and archeological samples

Modern hair samples were taken from a man of European descent (country of origin France) who had not been exposed to any cosmetic or chemical treatment, such as perming, bleaching, or dyeing. This sample was used as control for the first steps of the protocol optimization. We focused in the present study on one hair sample belonging to a larger corpus studied in the framework of a project in collaboration between the Laboratory of Archaeometrical Analyses and Research (LAIA) from the University of Tarapacá (UTA, Arica, Northern Chile), the Research and Restoration Centre for French Museums (C2RMF, Paris, France), and the Laboratory of Mass Spectrometry of Interactions and Systems (LSMIS, Strasbourg, France).

The mummies of the corpus were recovered from excavations conducted in the 1960s in the north of the Atacama Desert, on the coast of Arica and Parinacota region in Chile, and are conserved at the Museum San Miguel de Azapa of the UTA (Arica, Northern Chile). Moreover, the mummy bundle studied in this work, PLM7_T119, was found in the Playa Miller 7 archeological site (Arica valley) by Guillermo Focacci in 1974. Thanks to radiocarbon dating of its hair (Richardin et al., 2011, 2013), this mummy bundle was dated to 3880 ± 30 years before present (BP)—2123–1777 cal. years before Christ (BC)—at the transition between Archaic (6500–1700 years BC) and Formative (1700 years BC—500 years *Anno domini* [AD]) periods.

As stated in the R.1211-49 article of the French Public Health Code, hair is not subject to the CODECOH regulation, which frames the preparation, the conservation, the importation, or the exportation of human biological materials. No ethics committee approval was then required for this work in a French laboratory.

Hair cleaning workflow

Hair samples were cut into small sections, immersed in ultrapure water, and placed in an ultrasonic bath for a few seconds to remove the solids and mineral surface contaminations (Fig. 1). After water removal, hair samples were then washed with a mixture of DCM/MeOH 1:1 (v/v) to remove organic coating materials, decomposition fluids, and eventual varnishes or synthetic resins. The samples were then thoroughly rinsed once with ultrapure water, once with acetone, and three times with ultrapure water.

Protocol 1 for extraction, precipitation, and digestion of hair proteins (P1)

Cortical proteins from 15 mg of cleaned hair were extracted by following the protocol described by Richardin et al. (2011, 2013), and this was completed by a digestion protocol adapted from the work of Barthélémy (2011) (Fig. 1). Samples were immersed in 6 mL of a 0.13 M DTT, 0.03 M SDS, and 25 mM Tris-HCl solution and left for 2 days at 50°C without agitation. Under these conditions, the only residue was the cuticle emptied of the cortex and medulla. After extraction, the soluble and insoluble materials were separated, and the proteins present in the soluble fraction were purified by precipitation with 50 μL of a 2% DCO solution and 500 μL of a 100% TCA solution, forming a pellet. After centrifugation, the pellets were purified by using acetone in an ice bath, abundantly rinsed with ultrapure water, so that there was no trace of acetone left, and the pad was finally lyophilized overnight.

One milligram of protein extract was then resuspended and reduced at 60°C for 1 h in a solution at 25 mM NH_4HCO_3 and 20 mM DTT. The sample was alkylated with 40 mM IAM during 1 h at 25°C in the dark and precipitated with 2.5 volumes of EtOH, rinsed with 70% EtOH, and dried with N_2. Alkylated proteins were resuspended in 1 mL of 100 mM NH_4HCO_3 and 2 M urea solution, and digested at 37°C overnight with 1 μg of trypsin. Digestion was quenched with 0.2% FA, and samples were finally desalted by using Sep-Pak C18 1 cc SPE cartridges. Digests were rinsed with H_2O 0.1% FA, and a mixture of ACN/H_2O 0.1% FA (1:1, v/v) was used for the elution of the peptides. Desalted digests were concentrated to dryness by vacuum centrifugation and stored at −20°C or resuspended in 50 μL ACN/H_2O 0.1% FA (1:1, v/v) for further analyses.

Protocol 2 for extraction of hair proteins (P2)

The extraction protocol was adapted from the experimental procedures given by Lee et al. (2006) and Barthélémy (2011) dedicated to proteomic studies of modern hair (Fig. 1). After the cleaning step, 15 mg of cleaned hair sample were immersed in 6 mL of the extraction solution containing 7 M urea, 2 M thiourea, 50 mM DTT, 50 mM Tris-HCl (buffer solution, pH 7.5), and 0.1% Triton X-100. Hair was incubated without agitation for 18 h at 37°C. After decantation, the soluble (protein extract) and insoluble materials (mostly cuticle) were separated and the protein extract was alkylated with 1 M IAM and 3 M Tris-HCl for 10 min in the dark at room temperature. The alkylated extract was then dialyzed in 3500 MWCO Slide-A-Lyzer dialysis cartridges against water (two changes) over 2 days, and the extract was lyophilized overnight. Finally, the extract was reduced and digested by following the same workflow as for Protocol 1.

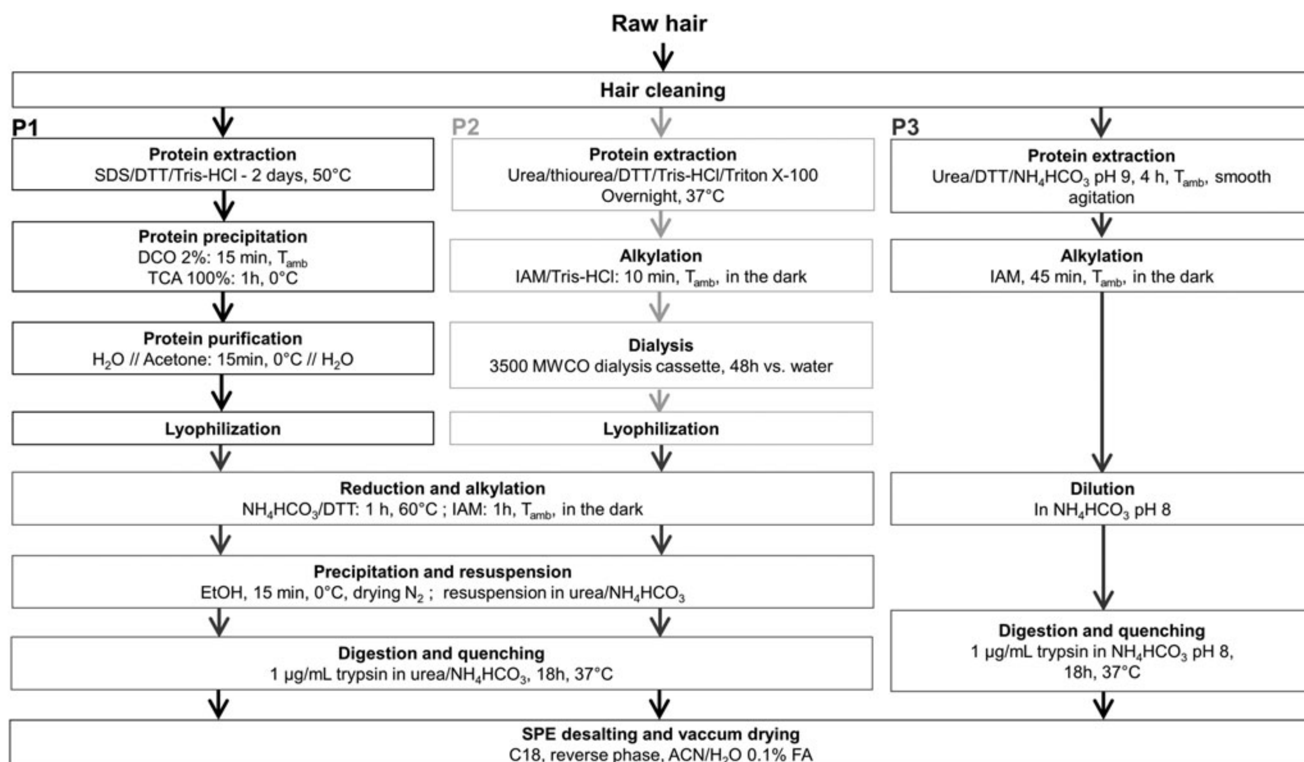

FIG. 1. Schemes of the three chosen proteomic protocols.

Protocol 3 for extraction and digestion of hair proteins (P3)

Initial experiment was led on 1 mg of cleaned hair by using a protocol adapted from the work of Solazzo et al. (2014) dedicated to the study of archeological keratinized fibers (Fig. 1). Samples were immersed in 1 mL of a solution of 8 M urea and 50 mM DTT in a 100 mM NH_4HCO_3 pH 9.1 buffer, and they were incubated with gentle agitation for 4 h at room temperature. The soluble (protein extract) and insoluble (mostly cuticle) materials were separated, and the protein extract was alkylated with 40 mM IAM for 45 min in the dark at room temperature. The alkylated extract was then digested at 37°C overnight by adding an equivalent volume of 50 mM NH_4HCO_3 pH 8.3 buffer and 10 μg of trypsin. Digestion was quenched with 0.2% FA, and samples were finally reduced to 1 mL by vacuum centrifugation and desalted by using Sep-Pak C18 1 cc SPE cartridges. Digests were rinsed with water 0.1% FA, and ACN/H_2O 0.1% FA (1:1, v/v) was used for the elution of the peptides. Desalted digests were concentrated to dryness by vacuum centrifugation and stored at –20°C or resuspended in 50 μL ACN/H_2O 0.1% FA (1:1, v/v) for further analyses.

Peptide mass fingerprinting by matrix-assisted laser desorption/ionization—time of flight–mass spectrometry

The matrix-assisted laser desorption-ionization mass spectrometry (MALDI-MS) analyses were conducted on an Autoflex II MALDI–Time of Flight (TOF)–mass spectrometer (Bruker Daltonics, Bremen, Germany). The device was featured with a pulsed nitrogen laser emitting at 337 nm and operated at an extraction voltage of 20 kV. The SPE desalted hair digests were resuspended in 50 μL ACN/H_2O 0.1% FA and deposited on a stainless steel target as dry droplets with a saturated solution of SA, in ACN/H_2O 0.1% TFA (1:1, v/v). Gated suppression was applied to prevent any saturation of the detector by matrix ions. One thousand laser shots were averaged for each spectrum and acquisitions were realized in a positive linear mode, adapted to compounds between 2 and 30 kDa. Laser power was optimized and kept constant for the different samples. Bruker Protein Calibration Standard I (4000–20,000 Da) was used for calibration, and spectra were processed with the Flex Analysis 3.4 software (Bruker Daltonics).

Peptide analysis and protein identification by nano liquid chromatography—electrospray ionization—tandem mass spectrometry

Desalted tryptic digests, resuspended in 50 μL ACN/H_2O 0.1% FA, were separated on an Ultimate 3000 nano-flow liquid chromatography device (Dionex; Thermo Scientific). The solvent system consisted of water with 0.2% FA (Solvent A) and ACN with 0.2% FA (Solvent B). Samples were first loaded onto a reversed-phase C18 pre-column (PepMap100, 5 μm particles, 100 Å pore size, 300 μm internal diameter (i.d.)) at a flow rate of 20 μL/min with 2% B. They were then transferred online to the analytical reversed-phase C18 column (75 μm i.d. × 15 cm, 3 μm particle, 100 Å pore size, Acclaim PepMap100 NanoViper from Thermo Fisher Scientific, Inc.), and peptides were separated at a flow rate of 300 nL/min, using a gradient from 2% B to 55% B. The column outlet was connected to a microTOF-Q II mass spectrometer (Bruker Daltonics) through a stainless steel nanospray needle.

Acquisition for the detection of precursor ions (MS) and fragmented product ions (MS/MS) was made in positive mode. Capillary voltage was set to –4.5 kV, nebulizer gas to 0.4 bar, dry gas to 4.0 L/min, source temperature to 180°C, and data acquisition for spectra collection was set to 2 Hz. For each MS scan, three precursor ions were selected for fragmentation and the total duty cycle was then 2 sec. Mass range for both MS and MS/MS scans was *m/z* 100–3000.

MS/MS data were then processed through SwissProt (Sp_20130109_decoy) and Human_Up11-6_decoy databases by using the Mascot 2.2.06 search algorithm developed by Matrix Science in Biotools 3.2 and ProteinScape 3.1 softwares (Bruker Daltonics). Trypsin cleavage rules were applied and the mass tolerance for precursor ions was set to ±30 ppm and to ±0.5 Da for fragments, allowing a maximum of three missed cleavages. N-terminal glutamine cyclization (–17.03 Da), N-terminal acetylation (+42.01 Da), carbamidomethylation of cysteine (+52.02 Da), N-terminal carbamylation (+43.01 Da), deamidation of asparagine and glutamine (+0.98 Da), and methionine and cysteine oxidation (+15.99 Da) were selected as variable modifications.

Research was processed by using decoy databases to obtain the peptide false discovery rates (FDR) for each sample. A manual validation was made on the identification results by fixing the protein FDR to a maximum of 1%. Only proteins respecting this fixed FDR value and yielding a Mascot score above 80, as well as at least two unique peptides, were accepted as identified. Moreover, peptides were accepted only when they yielded a Mascot score above 20 and *p*-value under 0.05.

To verify the reproducibility of the workflows and the analysis while preserving the samples, preparation of hair with each protocol was duplicated and each digested sample was analyzed twice by nano liquid chromatography–tandem mass spectrometry (nanoLC-MS/MS). To check the absence of contaminants from the laboratory environment, such as modern keratins, control samples were realized for each set of samples with ultrapure water, using the same protocol and at the same time as hair samples.

Results and Discussion

Three protocols adapted from the literature were applied to cleaned hair samples. Protocol 1 (P1) was adapted from the workflow for sample preparation for accelerator mass spectrometry (AMS) measurements (Richardin et al., 2011, 2013) on ancient hair. It was designed to selectively extract proteins from archeological keratinized materials, while removing all traces of exogenous compounds to prevent any misleading results for the dating of ancient hair. We also added several steps for protein digestion and peptide digest desalting to complete the sample preparation for nanoLC-MS/MS analyses. This workflow presents some key steps with a high risk of material losses, such as protein precipitation and purification, and it, thus, requires important material amounts. Richardin and coworkers, indeed, mentioned that they recovered about 20% in mass of proteins with this workflow and this, from more than 30 mg of raw ancient hair.

Protocol 2 (P2) was adapted from the work of Lee et al. (2006) and Barthélémy (2011). These works were dedicated to the study of important amounts of modern hair—typically

between 40 and 100 mg—to identify a maximum of hair proteins along with their location in the fiber and their molecular modifications.

Protocol 3 (P3) was adapted from the work of Solazzo et al. (2014) for the protein extraction and digestion, and it aimed at analyzing ancient proteins in small amounts of keratinized archeomaterials, typically <1 mg.

In the case of P1 and P2, working on too small hair amounts could lead to technical complications and significantly reduce the extraction rate, but due to the preciousness of our samples it is impossible to work on quantities as large as those described in the literature for these two first workflows.

In the scope of this work, the initial amounts of hair were then set to 15 mg for both P1 and P2, whereas the first experiments with P3 were conducted on 1 mg. This last quantity was not similar to the two others to stay in the mass range of the original study, but this difference will be taken into account for the result interpretations.

Preserving the archeological samples is a crucial issue regarding their preciousness and thus, optimizations of extraction and digestion protocols, of nanoLC-MS/MS parameters, and of MALDI-TOF-MS parameters were done on modern hair. For a proper comparison of modern and ancient samples, each type of hair was cleaned and prepared for analyses by using the same protocols described in the Materials and Methods section. Desalted digests were analyzed by MALDI-TOF-MS to visualize the peptide mass fingerprint of the different samples. Desalted digests were also processed through nanoLC-MS/MS to separate and to analyze the hair peptides. This enabled us to identify the proteins present in the different samples and to study their molecular modifications.

First observations on the three chosen protocols with modern hair

After the cleaning step common to the three protocols, P1 and P2 were first applied to 15 mg of modern hair and P3 was applied to 1 mg of hair. During the extraction step of proteins from modern hair, no clear change was visible with any of the three applied protocols. Hair seems to remain almost intact, hair color is quite similar, and fibers seem to be a little more flexible but keep the same shape, as shown in Supplementary Data (Supplementary Fig. S1). The protein extract—soluble part—shows a very light brown coloration that is almost not visible at first sight.

The more obvious observation is the complexity of P1 and P2 in comparison to P3. They require many stages, are time-consuming—more than 5 days—and are poorly reproducible due to their multiple steps of precipitation rinsing or of dialysis. Important initial amounts of raw hair are, thus, necessary to ensure the recovery of enough digested material to enable protein identification. Moreover, the use of TCA and acetone can induce protein modifications, such as proteolysis due to the high acidity of TCA or glycine modification due to acetone washing (Simpson and Beynon, 2010), and thus, it leads to misinterpretations on the conservation state of the samples. On the contrary, P3 is much easier, requires little time—<2 days—and shows a good reproducibility. The molecules of interest—proteins or peptides—remain solubilized throughout the entire process, from protein extraction to vacuum drying of SPE desalted digests. This highly improves reproducibility, minimizes sample losses, and enables to work on small amounts of raw material, such as 1 mg.

Given these observations, P3 is very promising in the scope of our project since it is crucial to preserve the integrity of the studied archeomaterials. However, the compatibility of P3 with nanoLC-MS/MS analysis and the quality of mass results of P3 versus P1 and P2 remain to be verified.

MALDI-TOF-MS and protocol efficiency on modern hair

As mentioned earlier, MALDI-MS analyses were performed on SPE (C18 cartridges) desalted hair digests to visualize the peptide fingerprint of modern hair and to verify the efficiency of the digestion step for the three different protocols. A first experiment was realized on 1 mg of protein extracts from P1 and P2, which were digested as described in Protocol 1 (data not shown). Regarding peptide distributions and intensities, MALDI-MS spectra show highly similar profiles for both hair digests and also for the control sample (ultrapure water processed through the same workflows than hair samples), which exhibits peaks from trypsin autolysis. The most significant differences compared with the control sample are two main peak groups at mass-to-charge ratios (m/z) 6500 and 8400 and three smaller ones at m/z 6900, 7200, and 9200, but for both protocols, the results are not particularly conclusive.

Digestion in Protocol 1 requires a complex workflow, whereas digestion in Protocol 3 is much easier to execute. In addition, it is more relevant to compare MS results from the digests of the three protocols if the digestion workflow is the same for the three protein extracts. One milligram of P1 and P2 extracts was then prepared by using the Protocol 3 digestion workflow as for the entire protein extract recovered from P3. The three peptide profiles are well resolved but quite different one from another and also from the control sample (Fig. 2). The difference of the hair samples with the control sample comes mostly from the high number of different peptides and from the lower relative intensities of peaks from trypsin autolysis in hair digests. These two points seem to point out the successful and efficient digestion of hair proteins.

NanoLC-MS/MS for identification of modern proteins and molecular degradation

SPE desalted peptide digests from the three protocols were processed through nanoLC-MS/MS analysis to identify modern hair proteins and to study the molecular degradation of modern hair, either coming from natural occurrence or induced by the treatment (Fig. 3 and Supplementary Table S1). With the applied analysis technique—microTOF-Q II mass spectrometer—no peptide or protein is identified in the control samples, even when searching for other taxonomies than "Homo Sapiens," thus showing the absence of detectable modern contaminants in our hair samples.

In P1 digests from the Protocol 1 digestion, more than 600 spectra were detected and the Mascot algorithm was able to match 15% of the spectra, from which it was possible to identify and accept 19 different proteins. Among them, 4 hair keratins were found out of the 18 described in literature (Szeverenyi et al., 2008). On the other hand, 1001 spectra were recovered from P1 samples after the Protocol 3 digestion and 41% of spectra corresponded to database spectra. This enabled to accept 19 proteins, including 9 described human hair keratins.

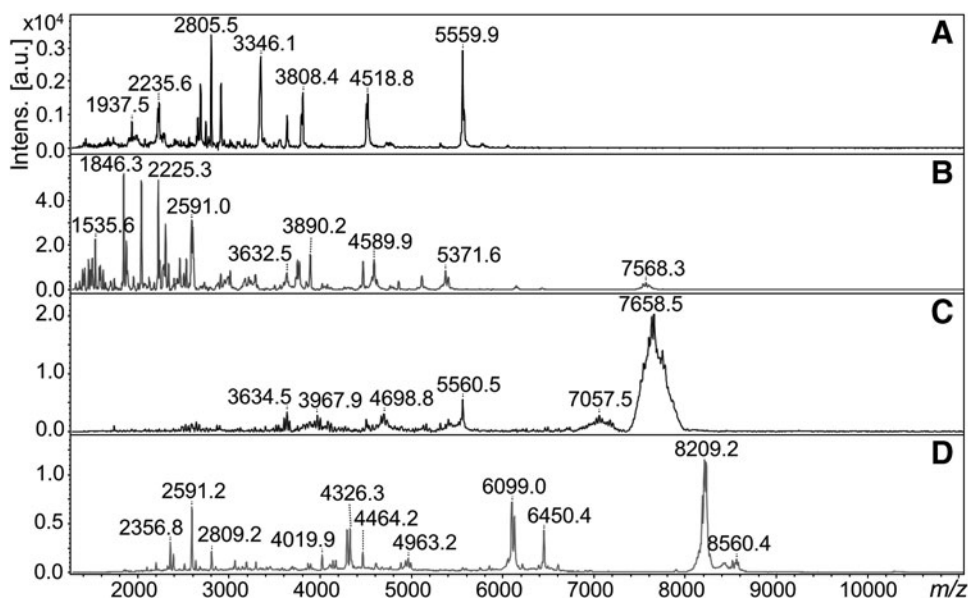

FIG. 2. MALDI-MS spectra of peptide digests of control of Protocol 3 digestion **(A)**, modern hair prepared with P1 **(B)**, P2 **(C),** and P3 **(D)** after Protocol 3 digestion. Matrix: SA 10 mg/mL in ACN/H_20 0.1% TFA (1:1, v/v); Positive ionization mode; Detection in linear mode with high gating suppression under 1000 Da; Laser power at 70.5%, 1500 shots/spectrum at a 50 Hz frequency. ACN, acetonitrile; MALDI-MS, matrix-assisted laser desorption-ionization mass spectrometry; SA, sinapic acid; TFA, trifluoroacetic acid.

With P2 after the Protocol 1 digestion, 395 spectra were detected and only 9 proteins were identified from 6% of matched spectra, including only 2 hair keratins. As for P1, 1001 spectra were recovered from P2 samples after the Protocol 3 digestion, but Mascot research yielded only 5% of matched spectra and 14 identified proteins with 5 described hair keratins. P3 enables to detect about 750 spectra in the analyzed digests and to identify 23 different proteins from 33% of matched spectra, including 8 keratins out of the 18 expected for human hair.

With this technique, it was also possible to study the molecular modifications of the keratins detected in modern hair and more particularly in digested samples from P1 and P3. The main modifications observed are induced by the different treatments, such as carbamylation or carbamidomethylation that come from extraction with urea (Lippincott and Apostol, 1999; Preedy, 2010) and the alkylation with IAM (Rombouts et al., 2013). It should be noted here that the use of NH_4HCO_3 buffer is known to inhibit protein carbamylation induced by urea (Sun et al., 2014); however, the applied NH_4HCO_3 concentrations are too low in this work to completely prevent these modifications that are then still detectable in the different digests.

When carbamylations and carbamidomethylations are not taken into account, an average of 15–35% of matched spectra exhibit different PTMs such as deamidation of glutamines and asparagines—<30% of deamidated sites on average, and oxidation of cysteines—<35% of oxidized sites (Fig. 3). Occasional methylation of aspartic and glutamic acids, oxidation of methionines, N-term cyclization of glutamines, and N-term acetylation were also identified.

Percentages of peptides with no, low (one), or high (two and three) number of miscleavages were also calculated to check the efficiency of the two digestion protocols. Results showed that digestion of Protocol 1 leads to a higher number of miscleavages for P1 and P2 protein extracts with <60% of peptides with no miscleavages, 25–40% with only one, and 19–30% with two or three. On the other hand, P1 and P2 digests from digestion of Protocol 3 exhibit more than 90% of peptides with no miscleavage, <10% with one miscleavage, and <2% with two or three miscleavages. P3 yields also good percentages of peptides with zero or only one miscleavage—respectively 77% and 19%—and a low percentage of peptides with two or three miscleavages—5% (Fig. 4).

These results once again highlight the very promising behavior of P3 for hair protein analysis. Indeed, regarding the better identification results for expected hair keratins, the higher percentages of matched spectra, and the higher rate of low miscleavages (zero or one miscleavage), digestion of Protocol 3 yields a much better efficiency than Protocol 1 digestion (Fig. 3 and Supplementary Table S1). In addition, 1001 spectra were recovered from both P1 and P2 samples digested as described in Protocol 3, whereas <600 and 400 spectra were, respectively, obtained for P1 and P2 samples digested as described in Protocol 1. This difference comes very likely from the low reproducibility of the protein precipitation-rinsing step in Protocol 1 digestion.

Moreover, despite the small initial amount of raw hair, and in comparison with P1 and P2, P3 enables to easily obtain with a good reproducibility, a quite satisfying identification of the major hair proteins with high sequence coverages between 40% and 80% (Supplementary Table S1). P1 yields better identification results than P3 with higher percentages of matched spectra, more identified proteins with similar or higher sequence coverages but from a 15 times higher amount of raw hair than P3. P2, however, shows much poorer results than the two other protocols since <6% of detected spectra could match with the databases. In addition, proteins are not well identified and sequence coverages are lower than 30%, despite the same amount of raw hair as P1 and a 15 times higher initial amount of hair than P3.

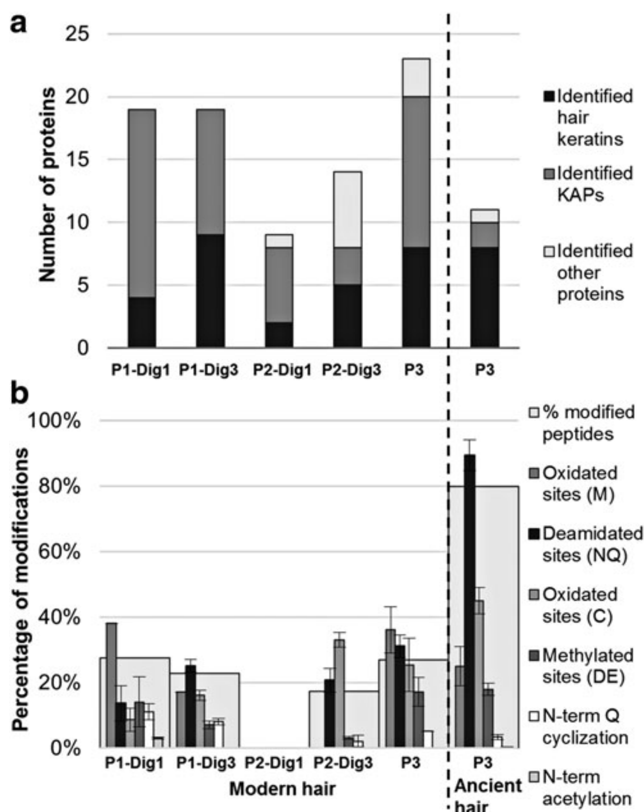

FIG. 3. Identification results of nanoLC-MS/MS data **(a)** and occurrences of PTMs **(b)** for modern hair after P1 with digestion of Protocol 1 (P1-Dig1), P1 with digestion of Protocol 3 (P1-Dig3), P2 with digestion of Protocol 1 (P2-Dig1), P2 with digestion of Protocol 3 (P2-Dig3) and P3, and for ancient hair after P3. Error bars for each specific modification correspond to the standard deviations calculated from the values obtained for the replicates of each experiment. nanoLC-MS/MS, nano liquid chromatography—tandem mass spectrometry; PTMs, post-translational modifications.

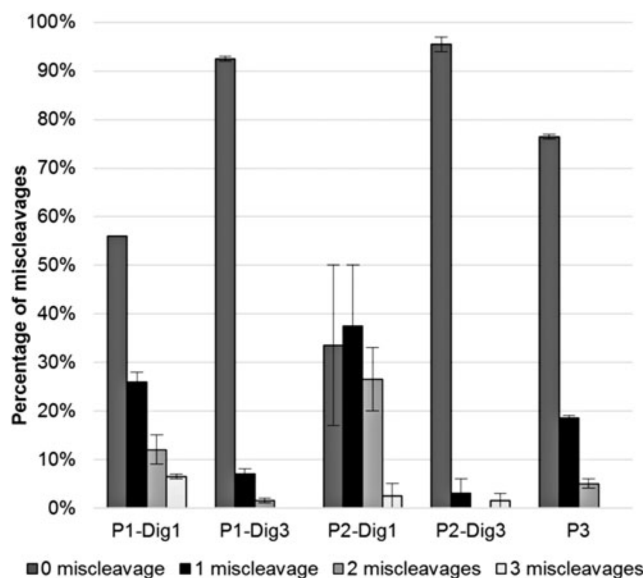

FIG. 4. Repartition of the different numbers of miscleavages for the digests from modern hair obtained after P1 with digestion of Protocol 1 (P1-Dig1), after P1 with digestion of Protocol 3 (P1-Dig3), after Protocol 2 with digestion of Protocol 1 (P2-Dig1), after P2 with digestion of Protocol 3 (P2-Dig3), and after P3. Error bars correspond to the standard deviations calculated from the replicates of each experiment.

In the scope of this project, and regarding the preciousness of archeological samples, P3 was then selected to be further tested on the archeological hair samples, whereas P1 and P2 were set aside for the following of the study.

Analysis of hair from an Andean mummy

After optimization on modern hair, the most promising workflow—P3—was applied to 1 mg of hair from a 3500-year-old Andean mummy. Early global observations pointed out the first indicators of the molecular degradation of ancient hair. Although no clear change was detectable on modern hair during protein extraction, archeological fibers were almost completely solubilized during the same extraction time, giving a yellow solution mostly made of cortical proteins, melanins, and internal lipids, and a tubular colorless insoluble residue composed by the outer layers of the cuticle (Supplementary Fig. S1).

Further testing on P3 showed that archeological sample was almost solubilized after about 1 h, whereas modern hair could be extracted during 48 h without any significant change. This difference of behavior between modern and archeological samples could be explained by molecular modifications, such as the probable alteration of inter- and intramolecular interactions inside hair, thus weakening the cohesion of archeological fibers and fostering the protein extraction process. In particular, disulfide bridges are an abundant molecular interaction in hair that gives its great robustness to the fiber. They are, however, very sensitive to oxidation, a frequently occurring modification in the archeological context, that could dramatically reduce the hair resistance to the protein extraction step.

MALDI-MS spectra of archeological hair digests exhibit similar masses in the low mass range to that of modern hair digests. However, peptides from ancient hair seem less abundant and are distributed on a narrower mass range at low *m/z* ratios (Supplementary Fig. S2). These observations altogether seem to point out the molecular degradation of ancient samples and the probable fragmentation of protein assemblies.

Finally, as with modern samples, hair from the mummy was analyzed by nanoLC-MS/MS (Fig. 3 and Supplementary Table S2). One thousand and one spectra were recovered from P3 and enabled to identify 11 archeological proteins from 25% of matched spectra, including 8 described hair keratins. Major hair keratins are still well identified in ancient hair but only few KAPs and other minor hair proteins were detected, thus pointing out a probably more advanced degradation (Fig. 3). Among these major identified keratins, seven were also found in modern hair: KRT31, KRT33a, KRT33b, KRT 34 (Type I), KRT83, KRT85, and KRT86 (Type II). The study of PTMs was then applied to these seven specific proteins, so that the interpretations and discussions could be more accurate.

As for modern hair, treatment-induced modifications are still present on an important part of peptides from both protocols, but in the case of archeological samples, they are not the major modifications. When these are not taken into account, more than 75% of matched spectra on average show the presence of PTMs. More specifically, P3 digest from ancient hair exhibits more than 85% of modified deamidation

sites against <30% in modern hair, and more than 35% of the identified cysteines are oxidized with 25% of trioxidated cysteines on average against <10% of trioxidation for modern cysteines (Fig. 3). Despite a good identification of ancient proteins in the hair from the studied mummy, with high sequence coverages for the six main keratins in the range 35–75% (Supplementary Table S2), these results seem to highlight an advanced molecular degradation of ancient hair proteins that have now to be studied further.

Minimization of raw hair amounts

In an archeometrical study, preserving the integrity of studied artefact is a crucial concern. It is then primordial to minimize as much as possible the amounts of archeological samples required for the different analyses without creating new technical difficulties or risks of sample losses. P3, which was first applied to 1 mg of raw ancient hair, was then tested on 500 and 200 μg of archeological hair (Supplementary Table S2).

The 500 μg sample gave globally similar MALDI-MS and nanoLC-MS/MS results than for 1 mg of hair without changing the dilution factor. Seven hundred and ninety spectra were, indeed, detected and 22% of spectra could be matched with database spectra, enabling to identify seven proteins. Among them, six were expected hair keratins, all common to digest from 1 mg of ancient hair and to modern hair digest.

On the other hand, workflow was more difficult to apply to 200 μg of hair because of the weighing and cleaning steps on such small samples, and quality of MS and MS/MS results was also poorer due to the lack of material. Seven hundred spectra were detected and six hair keratins were identified but only 15% of matched spectra were obtained and identification results showed lower sequence coverages and Mascot identification scores.

Further considerations

The results presented here gave new insights into the understanding of hair protein alteration processes and could be extended to the study of molecular modifications induced by heavy metals, which were an important environmental pollutant already in ancient times. This work could also be helpful for the study of modern phenomena, such as the degradation and aging processes of modern capillary materials or the impact of our environment on our organism. For further validation of this optimized workflow dedicated to the proteomic analysis of archeological hair, our method can be evaluated further by independent laboratories internationally.

Conclusions

In this article, we described the development of a proteomic approach for a new insight in the characterization of hair from Andean mummies. First, tests and optimizations of the three chosen protocols were done on modern hair and the most promising protocol—P3—was selected for further optimization on archeological hair. With this workflow, it was, indeed, possible to lower the initial hair amounts to 500 μg while maintaining a good quality of MS/MS data, especially for the six main keratins with high sequence coverages and high identification scores in the range 1100–4000. P3 was then selected for the following of this work on hair from Andean mummies with initial hair amounts of 500 μg.

This workflow helps not only to successfully identify main proteins from very small amounts of a quite complex material with high identification scores and high sequence coverages but also to go further in the characterization of this material with the study of PTMs. We were then able to identify main ancient keratins from hair of a prehispanic mummy and to assess for the first time the molecular preservation of ancient hair by a proteomic approach. In terms of methodology, our work enables to apply a proteomic approach to precious and rare samples not only in the context of archeometrical studies but also for any applications that would require the use of very small amounts of sample.

The results presented here collectively pave the way for new insights into the understanding of hair protein alteration processes such as those due to aging and ecological exposures.

Acknowledgments

The authors would like to thank Veronica Silva, curator at Museum National of Natural History, Santiago-Chile and PhD student at the Max Planck Institute for Evolutionary Anthropology (Leipzig, Germany) and at the University of Tarapacá and the Catholic University of Northern Chile (Arica) for her collaboration on the archaeological samples; Pr. Vivien Standen, Head of Physical Anthropology Laboratory from Museum San Miguel de Azapa of the University of Tarapacá; Laura Bom and the Laboratory of Magnetic Resonance and Membrane Biophysics from the University of Strasbourg for their help with the technical materials.

This project was financed by the University of Strasbourg for the PhD thesis scholarship and the Proyecto UTA 2014, 3741-14. The authors also thank Convenio de Desempeño UTA-MINEDUC as well as the National Network of Interdisciplinary Archaeometrical Competencies (CAI-RN) and their ITC program ''Fostering the competency transfer'' for their financial help. Finally, they thank Jean-Claude Dran for the improvements of the English writing.

Author Disclosure Statement

The authors declare that no conflicting financial interests exist.

References

Arriaza B. (2005). Arseniasis as an environmental hypothetical explanation for the origin of the oldest artificial mummification practice in the world. Chungara Rev Antropol Chil 37, 255–260.

Arriaza B, Amarasiriwardena D, Cornejo L, et al. (2010). Exploring chronic arsenic poisoning in pre-Columbian Chilean mummies. J Archaeol Sci 37, 1274–1278.

Aufderheide AC, Kelley MA, Rivera M, et al. (1994). Contributions of chemical dietary reconstruction to the assessment of adaptation by ancient highland immigrants (Alto Ramirez) to coastal conditions at Pisagua, North Chile. J Archaeol Sci 21, 515–524.

Barthélémy N. (2011). Qualitative and quantitative proteomico, a way to link genomic expression to the construction of biologic structures: Application to the human hair molecular structure understanding. PhD Thesis, University of Strasbourg.

Barthélémy NR, Bednarczyk A, Schaeffer-Reiss C, Jullien D, Van Dorsselaer A, and Cavusoglu N. (2012). Proteomic tools

for the investigation of human hair structural proteins and evidence of weakness sites on hair keratin coil segments. Anal Biochem 421, 43–55.

Bartkus L, Amarasiriwardena D, Arriaza B, Bellis D, and Yañez J. (2011). Exploring lead exposure in ancient Chilean mummies using a single strand of hair by laser ablation-inductively coupled plasma-mass spectrometry (LA-ICP-MS). Microchem J 98, 267–274.

Bencko V. (2005). Hair and exposure to environmental pollutants. In: *Hair in Toxicology – An Important Bio-Monitor. Part II.* Tobin DJ, ed. Cambridge: Royal Society of Chemistry, 159–174.

Bertrand L, Doucet J, Dumas P, Simionovici A, Tsoucaris G, and Walter P. (2003). Microbeam synchrotron imaging of hairs from ancient Egyptian mummies. Synchrotron Radiat 10, 387–392.

Bertrand L, Vichi A, Doucet J, Walter P, and Blanchard P. (2014). The fate of archaeological keratin fibres in a temperate burial context: Microtaphonomy study of hairs from Marie de Bretagne (15th c., Orléans, France). J Archaeol Sci 42, 487–499.

Bianucci R, Jeziorska M, Lallo R, et al. (2008). A pre-hispanic head. PLoS One 3, e2053.

Boston CE, and Arriaza BT. (2009). Arseniasis and teratogenic anomalies in the Atacama Desert coast of ancient Chile. Interciencia 34, 338–343.

Buckley M, and Kansa SW. (2011). Collagen fingerprinting of archaeological bone and teeth remains from Domuztepe, South Eastern Turkey. Archaeol Anthropol Sci 3, 271–280.

Buckley M, and Wadsworth C. (2014). Proteome degradation in ancient bone: Diagenesis and phylogenetic potential. Palaeogeogr Palaeoclimatol Palaeoecol 416, 69–79.

Byrne S, Amarasiriwardena D, Bandak B, et al. (2010). Were Chinchorros exposed to arsenic? Arsenic determination in Chinchorro mummies' hair by laser ablation inductively coupled plasma-mass spectrometry (LA-ICP-MS). Microchem J 94, 28–35.

Collins MJ, Cappellini E, Buckley M, Penkman KEH, Griffin RC, and Koon HEC. (2006). Analytical methods to detect ancient proteins. In: *Bio-and Material Cultures at Qumran.* Gunneweg J, Greenblatt C, and Adriaens A, eds. Fraunhofer, IRB Verlag, Stuttgart, 33–40.

Farell J, Amarasiriwardena D, Goodman AH, and Arriaza B. (2013). Bioimaging of trace metals in ancient Chilean mummies and contemporary Egyptian teeth by laser ablation-inductively coupled plasma-mass spectrometry (LA-ICP-MS). Microchem J 106, 340–346.

Focacci G. (1974). Excavaciones en el cementerio Playa Miller 7, Arica (Chile). Chungara Rev Antropol Chil 3, 23–74.

Fresnais M, Richardin P, Gimat A, Sepúlveda M, Leize-Wagner E, and Charrié A. (2015). Recent advances in the characterization of hair of mummies from the Chilean Andean coast. Forensic Sci Int 249, 25–34.

Goodrich JM, Chou H-N, Gruninger SE, Franzblau A, and Basu N. (2016). Exposure of dental professionnals to elemental mercury and methylmercury. J Expo Sci Environ Epidemiol 26, 78–85.

Haglund WD, and Sorg MH. (1996). *Forensic Taphonomy: The Postmortem Fate of Human Remains.* CRC Press, Boca Raton, FL, USA.

Hallégot P, Walter P, Cotte M, et al. (2008). Accumulations of Ca/P in the core of hairs from Taklamakan desert mummies. Appl Surf Sci 255, 1154–1157.

Hardy EM, Duca RC, Salquebre G, and Appenzeller BMR. (2015). Multi-residue analysis of organic pollutants in hair and urine for matrices comparison. Forensic Sci Int 249, 6–19.

Kales SN, and Christiani DC. (2005). Hair and metal toxicity. In: *Hair in Toxicology – An Important Bio-Monitor. Part II.* Tobin DJ, ed. Cambridge: Royal Society of Chemistry, 125–158.

Kamata T, Shima N, Sasaki K, et al. (2015). Time-course mass spectrometry imaging for depicting drug incorporation into hair. Anal Chem 87, 5476–5481.

Kempson IM, and Lombi E. (2011). Hair analysis as a biomonitor for toxicology, disease and health status. Chem Soc Rev 40, 3915.

Kintz P, Ginet M, Marques N, and Cirimele V. (2007). Arsenic speciation of two specimens of Napoleon's hair. Forensic Sci Int 170, 204–206.

Kintz P, and Villain M. (2005). Application of hair biology to environmental assessments. hair in forensic toxicology with a special focus on drug-facilitated crimes. In: *Hair in Toxicology – An Important Bio-Monitor. Part II.* Tobin DJ, ed. Cambridge: Royal Society of Chemistry, 87–104.

Kumtabtim U, Matusch A, Ulhoa Dani S, Siripinyanond A, and Sabine Becker J. (2011). Biomonitoring for arsenic, toxic and essential metals in single hair strands by laser ablation inductively coupled plasma mass spectrometry. Int J Mass Spectrom 307, 185–191.

Lee YJ, Rice RH, and Lee YM. (2006). Proteome analysis of human hair shaft from protein identification to posttranslational modification. Mol Cell Proteomics 5, 789–800.

Lippincott J, and Apostol I. (1999). Carbamylation of cysteine: A potential artifact in peptide mapping of hemoglobins in the presence of urea. Anal Biochem 267, 57–64.

Lubec G, Nauer G, Seifert K, et al. (1987). Structural stability of hair over three thousand years. J Archaeol Sci 14, 113–120.

Lubec G, Zimmerman MR, Teschler-Nicola M, Stocch V, and Aufderheide AC. (1997). Protein oxidation of a hair sample kept in alaskan ice for 800–1000 years. Free Radic Res 26, 457–462.

Preedy VR. (2010). *Adhesion Molecules.* CRC Press, Boca Raton, FL, USA.

Richardin P, Coudert M, Gandolfo N, and Vincent J. (2013). Radiocarbon dating of mummified human remains, Application to a series of Coptic mummies from the Louvre Museum. Radiocarbon 55, 345–352.

Richardin P, Gandolfo N, Carminati P, and Walter P. (2011). A new protocol for radiocarbon dating of hair and keratin type samples—Application to an Andean mummy from the National Museum of Natural History in Paris. Archaeol Anthropol Sci 3, 379–384.

Rombouts I, Lagrain B, Brunnbauer M, Delcour JA, and Koehler P. (2013). Improved identification of wheat gluten proteins through alkylation of cysteine residues and peptide-based mass spectrometry. Sci Rep 3, 2279.

Simpson DM, and Beynon RJ. (2010). Acetone precipitation of proteins and the modification of peptides. J Proteome Res 9, 444–450.

Solazzo C, Rogers PW, Weber L, Beaubien HF, Wilson J, and Collins M. (2014). Species identification by peptide mass fingerprinting (PMF) in fibre products preserved by association with copper-alloy artefacts. J Archaeol Sci 49, 524–535.

Sun S, Zhou J-Y, Yang W, and Zhang H. (2014). Inhibition of protein carbamylation in urea solution using ammonium-containing buffers. Anal Biochem 446, 76–81.

Swift JA, and Bews B. (1974). The chemistry of human hair cuticle – I: A new method for the physical isolation of cuticle. J Soc Cosmet Chem 25, 13–22.

Szeverenyi I, Cassidy AJ, Chung CW, et al. (2008). The Human Intermediate Filament Database: Comprehensive information on a gene family involved in many human diseases. Hum Mutat 29, 351–360.

Tobin DJ. (2005). *Hair in Toxicology*. The Royal Society of Chemistry, RCS Publishing, Cambridge.

Wilson AS. (2005). Hair as a bioresource in archaeological study. In: *Hair in Toxicology - An Important Bio-Monitor. Part IV*. Tobin DJ, ed. RCS Publishing, Cambridge, 321–347.

Wilson AS, Dodson HI, Janaway RC, Pollard AM, and Tobin DJ. (2007). Selective biodegradation in hair shafts derived from archaeological, forensic and experimental contexts. Br J Dermatol 157, 450–457.

Wolfram LJ. (2003). Human hair: A unique physicochemical composite. J Am Acad Dermatol 48, S106–S114.

Address correspondence to:
Armelle Charrié-Duhaut, PhD
Laboratoire de spectrométrie de masse des interactions et des systèmes (LSMIS)
Université de Strasbourg
CNRS, CMC UMR 7140
4, rue Blaise Pascal
F-67008 Strasbourg
France

E-mail: acharrie@unistra.fr

Abbreviations Used

ACN = acetonitrile
BC = before Christ
DCM = dichloromethane
DCO = sodium deoxycholate
DTT = DL-dithiothreitol
EtOH = ethanol
FA = formic acid
FDR = false discovery rates
IAM = iodoacetamide
KAPs = Keratin associated proteins
KIFs = Keratin intermediate filaments
MALDI-MS = matrix-assisted laser desorption-ionization—mass spectrometry
MeOH = methanol
nanoLC-MS/MS = nano liquid chromatography—tandem mass spectrometry
NH_4HCO_3 = ammonium bicarbonate
PTM = post-translational modification
SA = sinapic acid
SDS = sodium dodecyl sulfate
SPE = solid phase extraction
TCA = trichloroacetic acid
TFA = trifluoroacetic acid
TOF = time of flight
Tris-HCl = Trizma® hydrochloride