Anal. Chem. **2008**, *80*, 4590–4597

Identification of Protein Remains in Archaeological Potsherds by Proteomics

Caroline Solazzo,[†,‡] William W. Fitzhugh,[§] Christian Rolando,[†] and Caroline Tokarski*[,†]

Chimie Organique et Macromoléculaire, UMR CNRS 8009, and Protéomique, Modifications Post-traductionnelles et Glycobiologie, IFR 147, Université des Sciences et Technologies de Lille, 59655 Villeneuve d'Ascq Cedex, France, and Department of Anthropology, Smithsonian Institution, MRC 112, Washington, D.C. 20013-7012

We demonstrate here the possibility of identifying proteins trapped in few milligrams of the clay matrix of a 1200−1400 AD Iñupiat potsherd fragment from Point Barrow, Alaska, by a dedicated proteomics approach. The four main steps of a proteomics analysis, (i) protein extraction from biological samples, (ii) protein hydrolysis using a hydrolase enzyme, (iii) nanoLC, nanoESI MS, and MS/MS analysis of the generated peptides, and (iv) protein identification using protein databank proceeded from genomic data, have been optimized for archeological remains. Briefly our procedure starts by grinding the potsherds, extraction with 1% trifluoroacetic acid, digestion with excess of trypsin, nanoLC, nanoESI FT-ICR analysis, and data mining by homology search. The developed conditions were evaluated on protein extracts from remains obtained by heated muscle tissues and blubbers of different seal and whale species, these samples representing the main diet sources of the Eskimo population. Most of the proteins were identified by sequence homology to other species due to the lack of cetacean and pinniped proteins in the databanks. More interestingly, two proteins, myoglobin and hemoglobin, respectively, identified in muscle tissue samples and blubber samples highlight several specific peptides of cetacean and pinniped species; these peptides are significant to prove the presence of these marine species in the analyzed samples. Based on the developed methodology and on protein identification results obtained from the heated seal/whale muscle tissues and blubbers, the analysis of the clay matrix of a 1200−1400 AD Iñupiat potsherd fragment from Point Barrow was investigated. The described method succeeds in identifying four peptides corresponding to the harbor seal myoglobin (species *Phoca vitulina*) with a measured mass accuracy better than 1 ppm (MS and MS/MS experiments) including one specific peptide of the cetacean and pinniped species and one specific peptide of the seal species. These results highlight, for the first time, a methodology able to identify proteins from a few milligrams of archeological potsherd buried for years; the obtained results confirm the presence of a seal muscle tissue protein in this Punuk potsherd.

For years, human activity left material traces of its creative intelligence. Art or archeological remains are found, preserved, and studied revealing us, as real witness of a period, invaluable historical, socioeconomic, and cultural information. However, the precise study of organic material in cultural heritage pieces is a real analytical challenge,[1] and especially in archeological remains, particularly sensible during burial and leaching of hydrophilic molecules. As a consequence, most of the studies deal with lipidic remains.[2−6] The particular analysis of proteins in cultural heritage works was investigated on painting cross sections since the early 1970s using staining methods with organic dyes such as Naphtol Blue Black[7] for the specific detection of proteins, allowing an in situ microcharacterization of egg-based or collagen-based materials. To allow more specific protein bindings, antibody specificity was evaluated and successfully applied to painting binding media[8−10] or some years later, to archeological remains[11] identifying for example milk proteins in ancient potsherds.[12] Overall, immunoassay techniques are informative but they are limited to targeted protein. To determine the molecular structure of the organic matter, gas-phase chromatography and liquid-phase chromatography techniques were investigated for amino acid determination and composition by association with hydrolysis and

* To whom correspondence should be addressed. E-mail: Caroline.Tokarski@univ-lille1.fr. Phone: 333 20 33 64 33. Fax: 333 20 33 61 36.

† Université des Sciences et Technologies de Lille.

‡ Current address: BioArch, University of York, Biology (S Block), P.O. Box 373, York YO10 5YW, UK.

§ Smithsonian Institution.

(1) Tokarski, C.; Cren-Olive, C.; Rolando, C.; Martin, E. In *Molecular Biology and Cultural Heritage*; Saiz-Jimenez, C., Ed.; A.A. Balkema: Exton, PA, 2003; pp 119−130.

(2) Garnier, N.; Cren-Olive, C.; Rolando, C.; Regert, M. *Anal. Chem.* **2002**, *74*, 4868–4877.

(3) Regert, M.; Rolando, C. *Anal. Chem.* **2002**, *74*, 965–975.

(4) Evershed, R. P.; Dudd, S. N.; Copley, M. S.; Berstan, R.; Stott, A. W.; Mottram, H.; Buckley, S. A.; Crossman, Z. *Acc. Chem. Res.* **2002**, *35*, 660–668.

(5) Regert, M.; Garnier, N.; Decavallas, O.; Cren-Olive, C.; Rolando, C. *Meas. Sci. Technol.* **2003**, *14*, 1620–1630.

(6) Regert, M.; Langlois, J.; Laval, E.; Le Ho, A. S.; Pages-Camagna, S. *Anal. Chim. Acta* **2006**, *577*, 140–152.

(7) Martin, E. *Stud. Conserv.* **1977**, *22*, 63–67.

(8) Jones, P. L. *Stud. Conserv.* **1962**, *7*, 10–16.

(9) Johnson, M.; Packard, E. *Stud. Conserv.* **1971**, *16*, 145–164.

(10) Kockaert, L.; Gausset, P.; Dubi-Rucquoy, M. *Stud. Conserv.* **1989**, *34*, 183–188.

(11) Downs, E. F.; Lowenstein, J. M. *J. Archeol. Sci.* **1995**, *22*, 11–16.

(12) Craig, O.; Mulville, J.; Parker Pearson, M.; Sokol, R.; Gelsthorpe, K.; Stacey, R.; Collins, M. *Nature* **2000**, *408*, 312.

10.1021/ac800515v CCC: $40.75
Published on Web 05/22/2008

pyrolysis methods.[13–19] Based on a peak area calculation and ratio comparison of the amino acids from some references to the ancient sample ones, different proteinaceous media could be characterized but without any precise protein identification. To date, protein identification from archeological samples remains difficult. For example, analysis of early medieval potsherds lead to collagen identification,[14] the main protein of connective tissue in animals and the most abundant protein representing 25% of the total protein content. Collagen has been studied in ancient bones since the early 1980s by $\delta^{13}C$ and $\delta^{15}N$ stable isotope ratio measurements (isotope ratio mass spectrometry) reflecting marine and terrestrial components of ancient human diets.[20–22] To obtain more informative results, pioneer analytical work on ancient proteins by mass spectrometry was developed on fossils, allowing the characterization of bone protein osteocalcin by its molecular mass measurement and complementary radioimmunoassay.[23] More recently, peptide mass fingerprinting and peptide sequencing methods were introduced for identification of protein binders in paintings.[1,24] A complete structural approach for protein analysis in Renaissance paintings was even developed using proteomic technologies adapted to the small sample amount available, i.e., soft protein extraction, controlled hydrolysis, online nanochromatography–tandem mass spectrometry, and bioinformatics.[25] This study highlights for the first time the exact identification of egg yolk (vitellogenin) and egg white proteins (ovalbumin, conalbumin, lysozyme) in binders of a few micrograms of two Renaissance paintings. Similar structural approaches were applied on horse fossils identifying, for example the protein osteocalcin,[26] and on fossilized bones[27] and tissues of a 68-million-year-old dinosaur.[28]

Concerning archeological potsherds, protein study and identification are critical. The difficulties result principally from sample burials and leaching of hydrophilic molecules. We propose here a proteomic approach for the study of an Iñupiat potsherd, found in an eroding beach midden of the Punuk period (possibly 1200–1400 AD), from the archeological-rich area of Barrow in Alaska. The first part of the work exposes the methodology development including the soft protein extraction and analytical method validation compatible to small sample amounts. The method was evaluated on different residues from heated cetacean and pinniped muscle tissues and blubbers, which were the main Eskimo feeding source and thus probably represent the main organic remains. The second part of the work presents the results obtained on the archeological sample using the developed methodology. Since one single analysis could be realized on the archeological sample due to the extremely small protein remains, a Fourier transform mass spectrometer was used for MS and MS/MS experiments; its high measure accuracy, its very high resolution, and its possibility of ion accumulation in multipoles and ICR cell have contributed to the first protein identification from this type of archeological samples.

(13) Casoli, A.; Musini, P. C.; Palla, G. *Chromatographia* **1996**, *42*, 421–430.
(14) Evershed, R. P.; Tuross, N. *J. Archeol. Sci.* **1996**, *23*, 429–436.
(15) Colombini, M. P.; Modugno, F. *J. Sep. Sci.* **2004**, *27*, 147–160.
(16) Chen, X.; Fan, Y.; Sinkai, T.; Hu, Z.; Li, Z. *Can. J. Anal. Sci. Spectrosc.* **2005**, *50*, 325–329.
(17) Andreotti, A.; Bonaduce, I.; Colombini, M. P.; Gautier, G.; Modugno, F.; Ribechini, E. *Anal. Chem.* **2006**, *78*, 4490–4500.
(18) Prikryl, P.; Havlickova, L.; Pacakova, V.; Hradilova, J.; Stulik, K.; Hofta, P. *J. Sep. Sci.* **2006**, *29*, 2653–2663.
(19) Peris-Vicente, J.; Gimeno Adelantado, J. V.; Carbo, M. T. D.; Castro, R. M.; Reig, F. B. *Talanta* **2006**, *68*, 1648–1654.
(20) Schoeninger, M. J.; DeNiro, M. J.; Tauber, H. *Science* **1983**, *220*, 1381–1383.
(21) DeNiro, M. J. *Nature* **1985**, *317*, 806–809.
(22) Scarabino, C.; Lubritto, C.; Proto, A.; Rubino, M.; Fiengo, G.; Marzaioli, F.; Passariello, I.; Busiello, G.; Fortunato, A.; Alfano, D.; Sabbarese, C.; Rogalla, D.; De Cesare, N.; D'Onofrio, A.; Terrasi, F. *Isotopes Environ. Health Stud.* **2006**, *42*, 151–158.
(23) Ostrom, P. H.; Schall, M.; Gandhi, H.; Shen, T. L.; Hauschka, P. V.; Strahler, J. R.; Gage, D. A. *Geochim. Cosmochim. Acta* **2000**, *64*, 1043–1050.
(24) Kuckova, S.; Nemec, I.; Hynek, R.; Hradilova, J.; Grygar, T. *Anal. Bioanal. Chem.* **2005**, *382*, 275–282.
(25) Tokarski, C.; Martin, E.; Rolando, C.; Cren-Olive, C. *Anal. Chem.* **2006**, *78*, 1494–1502.
(26) Ostrom, P. H.; Gandhi, H.; Strahler, J. R.; Walker, A. K.; Andrews, P. C.; Leykam, J.; Stafford, T. W.; Kelly, R. L.; Walker, D. N.; Buckley, M.; Humpula, J *Geochim. Cosmochim. Acta* **2006**, *70*, 2034–2044.
(27) Asara, J. M.; Schweitzer, M. H.; Freimark, L. M.; Philips, M.; Cantley, L. C. *Science* **2007**, *316*, 280–285.
(28) Schweitzer, M. H.; Suo, Z.; Avci, R.; Asara, J. M.; Allen, M. A.; Arce, A.; Horner, J. R. *Science* **2007**, *316*, 277–280.

EXPERIMENTAL SECTION

Reagents. Trifluoroacetic acid, ammonium bicarbonate, guanidine hydrochloride, 1,4-dithiothreitol and iodoacetamide were obtained from Sigma-Aldrich (L'Isle d'Abeau Chesnes, France). Formic acid and methanol were purchased from Merck (Whitehouse Station, NJ). Trypsin was obtained from Promega (Charbonnières, France). Deionized water was provided by a Milli-Q Ultrapure water purification system (Millipore, Bedford, MA).

Seal and Whale Muscle Tissues and Blubbers. Harbor seal (*Phoca vitulina*) and gray seal (*Halichoerus grypus*) muscle tissues and blubbers were obtained from the Vertebrate Zoology Department of the National Museum of Natural History (2005 Canadian hunts collection, Smithsonian Institution, Washington, DC). Blubber from ringed seal (*Phoca hispida*) was provided by the National Zoo of Washington, DC. and blubber and muscle tissue from a beluga whale (*Delphinapterus leucas*) were sampled from a male specimen of the Churchill River (Manitoba, Canada) and preserved at the Smithsonian Institution. The samples were conserved frozen at −80 °C.

Archaeological Samples. The Iñupiat potsherd fragment (~3 cm width), from the archeological site of Point Barrow (Arctic coast of Alaska) was obtained from the Department of Anthropology, Smithsonian Institution. Relating to archeologist information, the pottery fragments are believed to come from a 1200–1400 AD cooking vessel used to process seal or whale tissues. The ceramic was heavily tempered with gravel, which gives it a brittle and porous aspect.

Sample Preparation. The sample preparation procedure is divided in three major steps, e.g., protein extraction, enzymatic hydrolysis and sample analysis/database searching for protein identification.

(1) Protein Extraction from Modern Samples. Muscle tissues and blubbers of harbor seal (*P. vitulina*), gray seal (*H. grypus*), ringed seal (*P. hispida*) and beluga whale (*D. leucas*) species were heated at 100 °C for 3 h on modern ceramics. The ceramics, containing apparent residues of heated muscle tissues and blubbers, were crushed in nitrogen, using a pestle and a mortar, into a fine powder. One milliliter of 1% trifluoroacetic acid solution was added to 25 mg of crushed sample. Each sample was subjected to ultrasonic baths (2 × 15 min) followed by centrifugation for 10 min at 1300*g*. The extraction supernatants were concentrated by vacuum centrifuge (Concentrator 5301;

Eppendorf, Hamburg, Germany) and suspended in 10 μL of 50 mM ammonium bicarbonate pH 8.8.

(2) Protein Extraction from Archeological Sample. The fragment of the archeological potsherd was crushed in nitrogen, using a pestle and a mortar, into a fine powder. One milliliter of 1% trifluoroacetic acid solution was added to 250 mg of grinded potsherd. The sample was subjected to ultrasonic baths (2 × 15 min) followed by centrifugation for 10 min at 1300*g*. The supernatant was concentrated by vacuum centrifuge (Concentrator 5301; Eppendorf) and suspended in 10 μL of 50 mM ammonium bicarbonate pH 8.8.

(3) Enzymatic Hydrolysis. The enzymatic hydrolysis was performed on the whole protein extract solution. The proteins were denatured and reduced with 90 μL of 50 mM Tris-HCl, 6 M guanidine, and 4 mM dithiotreitol and incubated for 2 h at 70 °C. Alkylation was realized in the dark with 10 μL of 50 mM iodoacetamide for 2 h. Samples were diluted in 20 mM Tris-HCl (pH 7.6) to reduce the concentration of guanidine to 600 mM. Trypsin was added (10 μL of 1 μg/μL trypsin in ammonium bicarbonate, 50 mM pH 8.8), and the mixture was incubated at 37 °C overnight. After concentration by vacuum centrifuge, the samples were placed in 3 μL of 0.1% trifluoroacetic acid.

(4) Sample Analysis and Database Search. Peptide mass fingerprinting spectra were obtained using a MALDI-TOF Voyager DE-STR mass spectrometer (Perseptive Biosystems, Framingham, MA) equipped with a 337-nm nitrogen laser. All spectra were obtained with the delayed extraction technology in reflectron mode. Peptides were analyzed using a solution of 2,5-dihydroxybenzoic acid (10 mg/mL methanol/water (v/v)); the acceleration voltage was fixed at 20 kV, the grid potential at 70%, and the extraction delay at 550 ns. A total of 200 laser shots were accumulated for peptide spectra. Peptide spectra were calibrated using tryptic hydrolysis ion peaks as internal standard (842.5100; 1045.5642; 2211.1046). All searches using peptide mass fingerprint spectra were carried out using the local MASCOT program with the SwissProt database. No restrictions were placed on species, isoelectric points, or protein mass. The tolerance was set to 50 ppm.

The peptide sequence spectra were obtained using nanochromatography (Ultimate LC system, Dionex, LC-Packings, Amsterdam, The Netherlands) online with an Apex Qe 9.4 T Fourier transform ion cyclotron resonance mass spectrometer (Bruker Daltonics, Bremen, Germany). Starting from a volume of 1 μL of the peptide solution, peptides were desalted and concentrated on a C18 preconcentration column (0.5 cm × 300 μm) and separated on a Pepmap C18 column (15 cm × 75 μm) at 200 nL/min solvent flow. The elution was performed using gradients of solvent A (95% H_2O, 5% ACN, 0.1% HCOOH) and solvent B (20% H_2O, 80% ACN, 0.1% HCOOH): 15 min in 100% solvent A, solvent B was increased to 100% in 130 min, then kept at 100% for 15 min, and finally decreased to 0% in 5 min. The column was allowed to equilibrate for 15 min before another run. The FT-ICR mass spectrometer is equipped with a nanoelectrospray source. Detection was carried out in positive mode. A potential of 1.7 kV was applied on the needle. Time cycle of an experiment (MS or MS/MS), including accumulation, transfer, excitation, detection, and quench, run on ~3 s. In details, ions were accumulated 1 s in the hexapole, 0.01 and 2 s in the quadrupole collision cell, respectively, for MS and MS/MS; 0.0016 s was set for optics transfer and 0.01 s for electronic dwell time. The detection parameters are broadband detection, 512 K acquisition size, start mass at m/z 200 leading to 0.5243-s transient duration, allowing theoretical resolution of 190 000 at m/z 400. Fragmentation was triggered for any multiply charged species with a relative collision energy calculated in function of the selected species charge state by the Bruker AutoMS method in ApexControl software. Once selected and fragmented, eluted species were excluded from the MS/MS selection for 10 min. Between each sample injection, one acetonitrile sample and three water-containing samples were injected; the generated spectra were without any remaining peptides. Concerning model and archeological samples, the peptide monoisotopic mass were extracted from each MS and MS/MS spectrum using DataAnalysis software (Bruker Daltonic) and submitted to local MASCOT software containing SwissProt database (release 0707; 276 256 sequences) for protein identification. The generated mass data in Mascot format were postcalibrated using an in-house program before extensive database search. In detail, a first search was performed using raw data and the most intense identified peptides were used for internal calibration. The error on the measured m/z versus the calculated value was minimized using a quadratic calibration according to mass (offset, slope, and quadratic coefficient) and a linear correction according to intensity, in an independent way for MS and MS/MS spectra. This calibration was then applied to the whole data set. No restrictions were placed on species or mass for the MS/MS experiments. The mass tolerance was set at 5 ppm on precursor ions and 5 mmu on fragment ions. A tolerance of three missed cleavage sites was set. Carbamidomethylation and methionine oxidation were entered as variable modifications. All MS/MS spectra resulting from the archeological sample leading to protein identification were manually checked to verify sequence assignments.

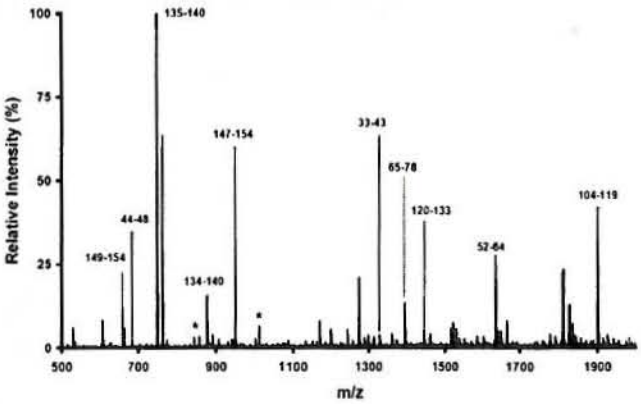

Figure 1. Peptide mass fingerprint spectrum of harbor seal muscle tissue hydrolyzed protein extract. The peptides corresponding to myoglobin are annotated using their position in the sequence. * Tryptic autolysis peaks.

RESULTS AND DISCUSSION

Methodological Development. Considering that seal or whale muscle tissues and blubbers, i.e. thick layer of vascularized fat found under the skin of all cetaceans and pinnipeds, were the principal Eskimo feeding source, and consequently represent a potential part of protein remains in Alaskan archeological potsherds, a preliminary proteomic work on different cetacean and

Table 1. Proteins Identified in Hydrolyzed Extracts from Harbor Seal Muscle Tissue and Blubber Using MALDI-TOF MS

identified proteins	species	sequence coverage (%)	identified peptides	position	$[M + H]^+$ mass	error (ppm)
			Muscle Tissue			
myoglobin	harbor seal	58	ELGFHG	149–154	659.307	12
			FDKFK	44–48	684.359	18
			ALELFR	135–140	748.425	13
			KALELFR	134–140	876.529	1
			YKELGFHG	147–154	950.461	12
			LFKSHPETLEK	33–43	1328.750	22
			HGNTVLTALGGILK	65–78	1393.816	<1
			HPAEFGADAQAAMK	120–133	1443.671	2
			SEDDMRRSEDLRK	52–64	1636.761	6
			YLEFISEAIIHVLHSK	104–119	1898.091	24
			Blubber			
hemoglobinα-chain	harbor seal	46	LRVDPVNFK	91–99	1087.574	47
			IGNHSAEYGAEALER	17–31	1616.774	4
			TYFSHFDLGHGSAQIK	41–56	1807.885	4
			TYFSHFDLGHGSAQIKGHGK	41–60	2187.097	10
			AVGHIDNLPDALSELSDLHAHK	69–90	2352.147	19
			AVGHIDNLPDALSELSDLHAHKLR	69–92	2621.346	12
hemoglobinβ-chain	harbor seal	60	LHVDPENFR	96–104	1129.548	14
			NPNVKKHGQK	57–66	1149.601	40
			LLGNVLVVVLAR	105–116	1265.843	10
			LLVVYPWTQR	31–40	1274.773	37
			VNVEEVGGEALGR	18–30	1328.665	11
			HLDDLKGTFAALSELHCDK	77–95	2169.995	30
			FFESFGDLSTADAVMKNPNVK	41–61	2333.127	6
			GTFAALSELHCDKLHVDPENFR	83–104	2556.235	2

pinniped residues of heated muscle tissues and blubbers was investigated to obtain protein and peptide profiles. Four species were mainly analyzed, i.e., harbor seal (*P. vitulina*), gray seal (*H. grypus*), ringed seal (*P. hispida*), and beluga whale (*D. leucas*).

Inspired by previous proteomic developments for cultural heritage samples, the protocol includes three major steps, (i) soft protein extraction, (ii) controlled protein hydrolysis, and (iii) peptide analysis; for example, a MALDI-TOF mass spectrometer was used for the verification of protein extracting solution efficiency and nanoLC, nanoESI-Qq-FT-ICR MS/MS was used for peptide sequencing and optimized for small amount sample analysis.

The first work difficulty was to find the extraction solution adapted to the protein extraction from the ancient potsherd considering that muscle tissues and blubbers were mixed, heated, and submitted to leaching of hydrophilic molecules during burial. Taking into account that the major part of the organic remains on the Iñupiat potsherd is lipidic, as observed during previous studies,[29,30] aqueous extraction solution was evaluated for extracting soluble proteins avoiding detergents such as sodium dodecyl sulfate or organic solvents that solubilize lipids. Acidified water with 1% trifluoroacetic acid, already tested on ancient paint media[25] and on modern ceramics impregnated with milk (data not shown), proved to be efficient for protein extraction from complex lipidic media and from impregnated ceramics. Consequently, this extraction solution was evaluated on a small amount of muscle tissues and blubbers preliminary heated on ceramics to mimic the Eskimo culinary process. The samples were ground under nitrogen with mortar and pestle in acidic solution, and the resulting supernatant containing denatured proteins was submitted to enzymatic hydrolysis using trypsin. The extracting solution and the enzymatic hydrolysis efficiencies were verified using MALDI-TOF MS. Figure 1 presents the peptide mass fingerprint spectrum of the hydrolyzed extract of harbor seal muscle tissue. The major peaks of the spectrum correspond to myoglobin peptides (Table 1), a muscle tissue protein; accurately 10 peaks annotated on Figure 1 were identified covering 58% of the protein sequence. No other protein could be identified with a significant score despite the presence of numerous peptides on the spectrum; this is probably the consequence of the poor harbor seal protein information in the databases; i.e., only 22 proteins were enumerated in the SwissProt database and 66 in the TrEMBL database. Similar results (not shown) were obtained from the other analyzed cetacean and pinniped samples; e.g., only the myoglobin could be identified. In the same way, the peptide mass fingerprint spectra of the protein extracts from harbor seal, gray seal, ringed seal, or beluga whale blubbers allow identifying major proteins present in the seal databases: the hemoglobin α-chain and β-chain. For example, the hemoglobin α-chain and β-chain were identified in harbor seal sample with six and eight peptides (Table 1) representing 46 and 60% coverage sequence, respectively.

To obtain more information on muscle tissue and blubber extracted proteins, peptide sequencing was investigated using nanoLC, nanoESI-Qq-FT-ICR MS/MS. Considering the results obtained on the harbor seal muscle tissues, it can be pointed out that six proteins could be identified despite the soft extraction conditions voluntarily employed (Table 2). The majority of the proteins was identified by sequence homology to other species as, for example, tropomyosin α-chain was identified with 21% sequence homology to rabbit species; some of them could be identified by sequence homology on the basis of a few long-

(29) Barnard, H.; Ambrosse, S. H.; Beehr, D. E.; Forster, M. D.; Lanehart, R. E.; Malainey, M. E.; Parr, R. E.; Rider, M.; Solazzo, C.; Yohe, R. M. *J. Archaeol. Sci.* **2007**, *34*, 28–37.

(30) Solazzo, C.; Erhardt, D. *Br. Archaeol. Rep.* **2007**, *13*, 161–177.

Table 2. Proteins Identified in Hydrolyzed Extracts from the Harbor Seal and Gray Seal Muscle Tissues and Gray Seal Blubber Using NanoLC, NanoESI-Qq-FT-ICR MS/MS

analyzed sample species	identified proteins	molecular mass (Da)	isoelectric point	score	identified peptides number	sequence coverage (%)	MS error (ppm)	specific peptides to cetacean and pinniped species	identified species: baikal seal (*P. sibirica*)	gray seal (*H. grypus*)[a]	harbor seal (*P. vitulina*)	other marine species	other species	MS/MS mean error (ppm)
								Muscle Tissues						
harbor seal (*P. vitulina*)[a]	myoglobin	17 417	8.72	117	10	59	3	LFKSHPETLEK	X	X	X	[b]		<1
								SHPETLEKFDK	X	X	X			1
								HKIPIKYLEFISE AIIHVLHSK	X	X	X			1
								IPIKYLEFISEAII HVLHSK	X	X	X			<1
								YKELGFHG		X	X	[c]	[d]	1
	myosin 1	20 919	4.96	66	2	10	3						X	
	troponin T slow skeletal muscle	31 265	5.71	57	1	8	1						X	
	troponin T fast skeletal muscle	33 014	5.63	57	1	9	1						X	
	hemoglobin subunit β	15 910	7.96	24	2	10	3						X	
	tropomyosin α-chain	32 661	4.69	282	3	21	3						X	
gray seal (*H. grypus*[a])	tropomyosin β chain	32 817	4.66	215	9	34	2						X	
	tropomyosin α-1 chain	32 675	4.69	169	9	25	2						X	
	actin	41 982	5.23	87	2	5	<1						X	
	myoglobin	17 417	8.72	58	4	16	1	LFKSHPETLEK	X	X	X	[b]		<1
								SHPETLEK	X	X	X	[b]		<1
								SEDLRK		X	X			1
								YKELGFHG		X	X	[c]	[d]	1
	myosin-7	223 091	5.58	49	3	1	<1						X	
	fructose-bisphosphate aldolase A	39 318	8.31	41	3	7	<1						X	
								Blubber						
gray seal (*H. grypus*[a])	hemoglobin subunit α	15 335	8.71	356	11	75	<1	MFINFPSTK				[c]		<1
								IGNHSAEYGAEALER				[c]		1
								VLSPADKTNVKGTWSK				[c]		1
								TYFSHFDLGHGSAQIK				[c]		2
								AVGHIDNLPDALSELS DLHAHKLR				[c]		2
	hemoglobin subunit β	16 025	6.77	331	7	45	<1						X	
	serum albumin	68 615	5.46	278	8	6	2						X	
	histone H1.1	10 359	10.49	269	5	47	<1						X	
	histone H4	11 374	11.36	205	5	39	<1						X	
	histone H2B	13 898	10.31	139	5	37	<1						X	
	histone H2B 1	13 492	10.40	81	3	20	1						X	
	histone H2B.2	15 769	11.86	73	3	18	1						X	
	vimentin	53 676	5.06	41	1	3	<1						X	
	histone H1.0	20 938	10.90	40	1	5	<1						X	
	histone H2A	14 225	10.86	40	1	6	<1						X	

[a] Decoy 0, false positive rate 0%. [b] Sperm whale (*Physeter catodon*) [c] Amazon dolphin (*Inia geoffrensis*), Atlantic bottle-nosed dolphin (*Tursiops truncatus*), beluga whale (*D. leucas*), Black Sea dolphin (*Delphinus delphis*), bridled dolphin (*Stenella attenuata*), Dall's porpoise (*Phocoenoides dalli dalli*), goose-beaked whale (*Ziphius cavirostris*), harbor porpoise (*Phocoenoides phocoena*),Hubb's whale (*Mesoplodon carlhubbsi*), killer whale (*Orcinus orca*), long-finned pilot whale (*Globicephala melaena*), Longman's beaked whale (*Indopacetus pacificus*), melon-headed whale (*Peponocephala electra*), and Stejneger's beaked whale (*Mesoplodon stejnegeri*). [d] Black-lipped pika (*Ochotona curzoniae*). [e] Atlantic bottle-nosed dolphin (*T. truncatus*).

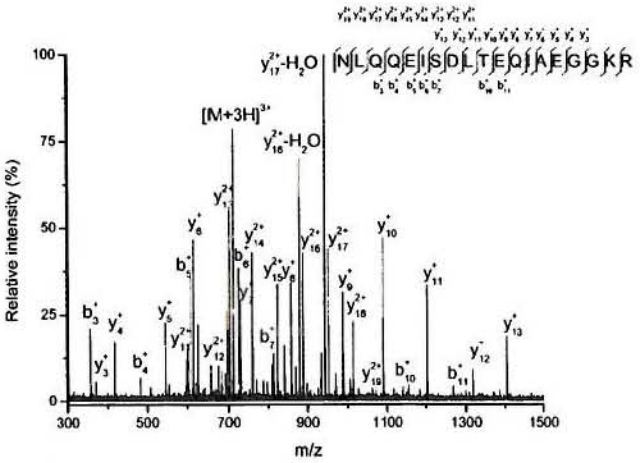

Figure 2. FT-ICR-MS/MS spectrum of the triply charged ion at *m/z* 710.365, from the hydrolyzed extract of harbor seal muscle tissue, presenting the y and b fragments of the myosin heavy chain peptide *m/z* 1507–1525 (NLQQEISDLTEQIAEGGKR) (sequence homology to horse species).

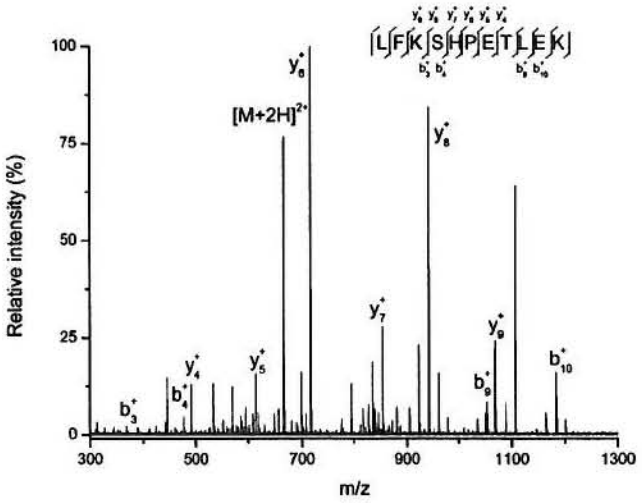

Figure 3. FT-ICR-MS/MS spectrum of the doubly charged ion at *m/z* 664.865 from the hydrolyzed extract of harbor seal muscle tissue, presenting the y and b fragments of the harbor seal myoglobin peptide 33–43 (LFKSHPETLEK).

sequence peptides due to the high mass measurement accuracy obtained with the Fourier transform mass spectrometer. For example, Figure 2 presents triply charged ion at *m/z* 710.365 from the hydrolyzed extract of harbor seal muscle tissue. The y and b fragments (y_3^+–y_{13}^+, y_{11}^{2+}–y_{19}^{2+}, b_3^+–b_7^+, b_{10}^+, b_{11}^+), covering almost the totality of the sequence with a mean mass error less than 1 ppm, obviously allow the peptide 1507–1525 identification from myosin-1 heavy chain (NLQQEISDLTEQIAEGGKR) (sequence homology to horse species).

More interestingly, the myoglobin, identified in the harbor seal muscle tissues with 59% sequence coverage, contains 5 specific peptides of cetacean and pinniped species (see the species details, Table 2) versus 10 peptides identified, i.e., LFKSHPETLEK, SHPETLEKFDK, HKIPIKYLEFISEAIIHVLHSK, IPIKYLEFISEAIIHVLHSK, and YKELGFHG (including black-lipped pika species for the YKELGFHG sequence) with the respective positions 33–43, 36–46, 98–119, 100–119, and 147–154 in the harbor seal myoglobin sequence. For example, Figure 3 presents the MS/MS spectrum of the doubly charged ion at *m/z* 664.865 from the hydrolyzed extract of harbor seal muscle tissue. The spectrum allows identifying the sequence LFKSHPETLEK using the y fragments (y_4^+–y_9^+) and b fragments (b_3^+, b_4^+, b_9^+, b_{10}^+) corresponding to the peptide 33–43 of the harbor seal myoglobin; in addition, this sequence is exclusively found in three other species: the gray seal, the baikal seal, and the sperm whale species. Additionally, it can be pointed out that specific markers of the myoglobin were also identified in the gray seal muscle tissue sample (see the species details, Table 2), particularly in three peptides already identified in the harbor seal muscle tissue sample (LFKSHPETLEK, SHPETLEK, and YKELGFHG with the respective positions 33–43, 36–43, and 147–154 in the gray seal myoglobin sequence) and one additional peptide specific to gray seal and harbor seal species (SEDLRK, position 59–64 in the gray seal myoglobin sequence).

Concerning the blubber analysis, some specific hemoglobin peptides could be also highlighted. For example, the analysis of gray seal blubber allows identifying 5 hemoglobin subunit α-peptides specific to one unique species versus 11 identified peptides: the Atlantic bottle-nosed dolphin (Table 2). It has to be noted that

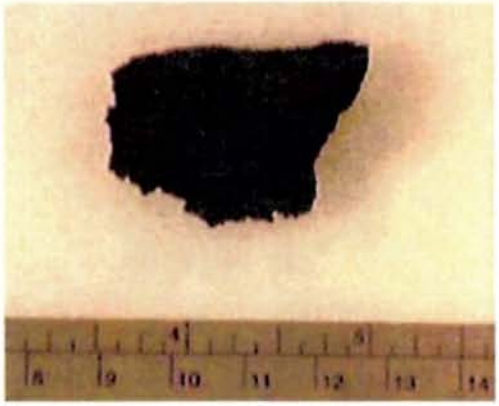

Figure 4. 1200–1400 AD potsherd fragment from the Iñupiat archaeological site of Point Barrow, Alaska, obtained from the Department of Anthropology, Smithsonian Institution (Washington, DC).

the gray seal hemoglobin is not present in the database, but 1659 hemoglobin entries are counted (SwissProt, TrEMBL) including several species among which 6 cetaceans and pinnipeds species, i.e., Galapagoes fur seal, Weddell seal, harbor seal, Atlantic bottle-nosed dolphin, minke whale, and sperm whale.

Globally, these results highlight that the proposed extraction and analysis methodologies succeed in extracting muscle tissue and blubber proteins of cetaceans and pinnipeds and allowed the identification of specific markers of seal species or cetacean and pinniped species.

Application of the Developed Methodology to Archaeological Samples. The described methodology, developed on residues from heated muscle tissues and blubbers, was applied to a 1200–1400 AD potsherd fragment from the Iñupiat archeological site of Point Barrow in North Alaska. The very small archeological fragment of 3-cm width (Figure 4) and, consequently, the extremely small protein remains induce the optimization of the analytical procedure since one single analysis can be realized. The ancient potsherd fragment was crushed with a lot of care into fine powder in the extracting acidified water (1% TFA) solution under nitrogen and enzymatically hydrolyzed. A volume

of 1 μL of the concentrated extract was used for analysis. Peptides generated by tryptic hydrolysis of extracted proteins were separated by nanochromatography and online analyzed with nanoESI-Qq-FT-ICR MS/MS. A previous concentration and desalting step was realized on a C18 nanoprecolumn, directly connected to a C18 analytical column 15 cm length × 75 μm inner diameter. The chromatographic gradient was extended to 3 h to obtain the highest information amount from a single run as well as the best separation of peptides and potentially artifactual extracted molecules. The separated peptides are then online analyzed by the nanoESI-Qq-FT-ICR MS/MS. The time cycle of MS or MS/MS experiments, including accumulation, transfer, excitation, detection, and quench, was adapted to the small sample amounts. In particular, the ion accumulation times in the multipoles and in the ICR cell are important as improving the detection of the very small ion amounts. However, the ratio between small ion amount detection and scan speed has to be balanced. The optimized time cycle run on ~3 s including 1-s ion accumulation in the hexapole and 0.01 and 2 s in the quadrupole collision cell, respectively, for MS and MS/MS; 0.0016 s was set for optics transfer and 0.01 s for electronic dwell time. The high resolving power and the high measurement accuracy of the FT-ICR mass spectrometer allowed obtaining precise peptide and fragment mass measurements, facilitating the protein identification. The identified peptides were generated with measure errors less than 1 ppm on precursor and fragment ions succeeding in the identification of a muscle tissue protein already identified in the model heated muscle tissue, the myoglobin with four peptides (Table 3). For example, Figure 5 presents the ions $y_3^+ - y_{11}^+$ and $b_3^+ - b_{10}^+$, b_{13}^+ allowing the identification of the sequence VETDLAGHGQEVLIR corresponding to the myoglobin peptide 18–32 of different species presented in Table 3. Considering that this spectrum is obtained from a small amount of archeological potsherd, it could be pointed out that high spectrum quality could be obtained with the optimized analytical procedure.

Globally, it appears that the four identified myoglobin peptides represent 32% coverage sequence of harbor seal or gray seal myoglobin, these two species having a totally identical myoglobin sequence. As expected, the sequenced peptides, extracted with aqueous acidified water, correspond mostly to hydrophilic constitutive peptides of the identified proteins. For example, the identified myoglobin peptide gravy scores are negative, i.e., −0.762 for YKELGFHG, −0.421 for HPAEFGADAQAAMK, −1.469 for SHPETLEKFDKFK, and −0.060 for VETDLAGHGQEVLIR, compared to the other nonidentified peptides of harbor/gray seal myoglobin, which are more hydrophobic as for example MGLSDGEWHLVLNVWGK (GRAVY score 0.018), HGNTVLTALGGILK (GRAVY score 0.621), and YLEFISEAIIHVLHSK (GRAVY score 0.606).

Regarding species matching among the 4 identified myoglobin peptides, 2 peptides correspond exclusively to pinniped or cetacean species; i.e., the sequence SHPETLEKFDKFK corresponds to 3 seal species (see Table 3) and the C terminus sequence YKELGFHG corresponds to 17 pinniped and cetacean species (except Black-lipped pika species; see Table 3). These markers were previously found during the analysis of harbor seal muscle tissues. As a result, it can be emphasized that these two peptides give us significant information concerning the presence of marine

Table 3. Harbor Seal Myoglobin Sequences Identified by NanoLC, NanoESI-Qq-FT-ICR-MS/MS from the Protein Extract of the 1200-1400 AD Arctic Potsherd Fragment

identified protein myoglobin[a]	identified sequences	species					peptide position	retention time (min)	precursor ions mass/charge		mean error MS/MS/MS
		baikal seal (*P. sibirica*)	gray seal (*H. grypus*)	harbor seal (*P. vitulina*)	other marine species	other species					
	SHPETLEKFDKFK	X	X	X			36–48	30.3	803.417	2+	<1
	YKELGFHG		X	X	[b]	[c]	147–154	29.6	475.740	2+	<1
	HPAEFGADAQAAMK	X	X	X		[d]	120–133	30.4	722.338	2+	<1
	VETDLAGHGQEVLIR		X	X		[e]	18–32	34.3	818.936	2+	<1

[a] Decoy 0, false positive rate 0%. [b] Amazon dolphin (*I. geoffrensis*), Atlantic bottle-nosed dolphin (*T. truncatus*), beluga whale (*D. leucas*), Black Sea dolphin (*D. delphis*), bridled dolphin (*S. attenuata*), Dall's porpoise (*P. dalli dalli*), goose-beaked whale (*Z. cavirostris*), harbor porpoise (*P. phocoena*),Hubb's whale (*M. carlhubbsi*), Killer whale (*O. orca*), long-finned pilot whale (*G. melaena*), Longman's beaked whale (*I. pacificus*), melon-headed whale (*P. electra*), and Steineger's beaked whale (*M. steinegeri*). [c] Black-lipped pika (*O. curzoniae*). [d] African elephant (*Loxodonta Africana*), weasel lemur (*Lepilemur mustelinus*). [e] African wild dog (*Lycaon pictus*), bat-eared fox (*Otocyon megalotis*), cape fox (*Vulpes chama*), and dog (*Canis familiaris*).

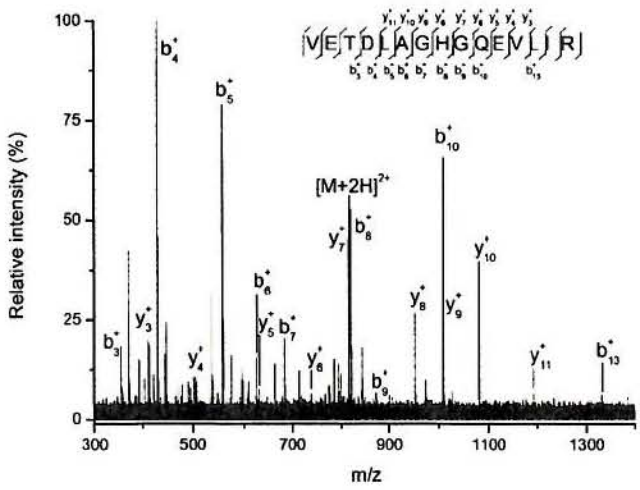

Figure 5. FT-ICR-MS/MS spectrum of the doubly charged ion at *m/z* 818.936, from the hydrolyzed protein extract of 1200−1400 AD potsherd fragment, presenting the y and b fragments of the myoglobin peptide 18−32 (VETDLAGHGQEVLIR).

species in the archeological potsherd and more precisely the presence of seal species.

Thus, the developed methodology allowed the identification of protein remains from small archeological sample amounts. The obtained results point out two peptide sequences specific to pinniped and cetacean species, proving the presence of these marine species in this Inūpiat potsherd. Additionally, this work argues the hypothesis that entrapped proteins were not totally leached during burial; these results are certainly improved by the potsherd conservation conditions, i.e., an Arctic environment, well appropriate to the preservation of organic compounds.

CONCLUSION

This study represents the first identification of proteins by sequencing from archeological potsherd remains. The extraction and the analytical procedures were optimized for small sample amounts. The efficiency of the soft extraction solution composed of acidified water with 1% TFA was verified by MALDI-TOF MS. The nanoLC, nanoESI-Qq-FT-ICR MS/MS procedure was adapted for detecting and sequencing the highest number of peptides from a single analysis starting from 1 μL of the concentrated enzymatically hydrolyzed protein extract. Concerning the chromatographic run, sample preconcentration and desalting prior to separation were realized online with the analytical column; the analytical chromatographic gradient was extended to 3 h to obtain the best separation of peptides and potentially artifactual extracted molecules. The time cycle of the FT-ICR including MS or MS/MS experiments was optimized by the accumulation ion times increase in the multipoles and the ICR cell for detecting low-abundant peptides. Starting from a very small amount of 1200−1400 AD potsherd fragment (3-cm width) from the Iñupiat archeological site of Point Barrow, Alaska, the developed methodology succeeds in identifying a muscle tissue protein, myoglobin with four peptides including one peptide sequence specific to pinniped and cetacean species and one specific to seal species. The high resolving power and the high measurement accuracy of the FT-ICR mass spectrometer allowed obtaining precise peptide and fragment mass measurements as the identified peptides were generated with measure errors less than 1 ppm on precursor and fragment ions. This work clearly confirm the potential and support of such a high-resolution mass spectrometer in the case of difficult proteomic analysis such as art and archeological sample studies. These results underline and confirm archeological and ethnological studies proving the importance of seal hunting for Inūpiat Eskimos.

ACKNOWLEDGMENT

The Proteomics core facility and the FT-ICR spectrometer used in this study were funded by the European Community (FEDER), the Région Nord-Pas de Calais (France), the Génopole of Lille, the CNRS, and the Université des Sciences et Technologies de Lille. We thank Fabrice Cantais for performing the FT-ICR experiments. This work was in part supported by a grant of the Smithsonian's Museum Conservation Institute, project 6001 on Archaeological Food Residues.

Received for review March 11, 2008. Accepted April 24, 2008.

AC800515V