LAP-MALDI MS analysis of amelogenin from teeth for biological sex estimation

Adair, L. R., Lewis, M. E. ORCID: https://orcid.org/0000-0001-6224-0278, Collins, M. J. and Cramer, R. ORCID: https://orcid.org/0000-0002-8037-2511 (2025) LAP-MALDI MS analysis of amelogenin from teeth for biological sex estimation. Journal of Pharmaceutical and Biomedical Analysis, 255. 116599. ISSN 0731-7085 doi: https://doi.org/10.1016/j.jpba.2024.116599 Available at https://centaur.reading.ac.uk/119753/

Journal of Pharmaceutical and Biomedical Analysis 255 (2025) 116599

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.journals.elsevier.com/journal-of-pharmaceutical-and-biomedical-analysis

LAP-MALDI MS analysis of amelogenin from teeth for biological sex estimation

Lily R. Adair [a], Mary E. Lewis [b], Matthew J. Collins [c,d], Rainer Cramer [a,*]

[a] Department of Chemistry, University of Reading, Whiteknights, Reading RG6 6DX, United Kingdom
[b] Department of Archaeology, University of Reading, Whiteknights, Reading RG6 6AB, United Kingdom
[c] McDonald Institute for Archaeological Research, University of Cambridge, Henry Wellcome Building, Fitzwilliam Street, Cambridge CB2 1QH, United Kingdom
[d] The Globe Institute, Faculty of Health and Medical Sciences, University of Copenhagen, CSS, Oster Farimagsgade 5, Kobenhavn K 1353, Denmark

ARTICLE INFO

Keywords:
Amelogenin
Enamel proteins
LAP-MALDI
MALDI
Mass Spectrometry
Sexing
Sex estimation

ABSTRACT

The biological sex estimation of human individuals can be achieved by extracting fragments of the amelogenin protein from small areas of tooth enamel. The amelogenin gene can be found on both sex chromosomes (X and Y) with chromosome-specific differences in its sequence, and consequently the sequences of the expressed protein in teeth. Virtually all current analytical techniques used to identify the occurrence of the male Y chromosome-specific proteoform employ proteoform-specific peptide analysis by LC-ESI MS/MS, which typically results in longer analytical times due to the LC separation step, despite recent efforts of shortening the LC step. We report a rapid analytical workflow for biological sex estimation by combining minimal acid extraction of amelogenin peptides, including the Y chromosome-specific SM_{ox}IRPPY peptide, with LAP-MALDI (liquid atmospheric pressure matrix-assisted laser desorption/ionization) MS and MS/MS analysis but without the use of an LC system. A total of 27 peptides from amelogenin and ameloblastin were characterized by MS/MS, revealing oxidation and deamidation as chemical modifications and information on the maturation of amelogenin. The entire sample preparation and analysis time for biological sex estimation using the applied workflow is $\leq$ 10 minutes, of which only 1 minute is needed for the MS and MS/MS data acquisition. The sample preparation is minimally hazardous, requiring 10 % HCl for peptide extraction, and can be undertaken in non-specialized labs before being submitted to MS and MS/MS analysis. The developed workflow can also facilitate the MS/MS analysis of many other amelogenin peptides without LC separation, providing further proteomic information on protein expression and mRNA transcription. It was applied to the teeth of five males and five females, whose biological sex had been estimated using osteological techniques, from three archaeological sites.

1. Introduction

Sequencing of the amelogenin gene on both sex chromosomes was accomplished more than three decades ago [1]. Since then, PCR testing has been applied for biological sex estimation by exploiting the differences between the X- and Y-chromosomal versions of the amelogenin gene [2,3]. However, DNA analysis is time-consuming and requires sufficient DNA preservation, which is not always guaranteed, especially in archaeology and forensics.

Amelogenin protein analysis methods, in particular mass spectrometry (MS) methods based on proteomic analysis workflows, have been employed as an alternative for biological sex estimation only in the last few years [4–7]. These MS-based proteomic methods rely on liquid chromatography (LC) separation and electrospray ionisation (ESI) as applied in large-scale proteome analyses. Typically, amelogenin peptides are extracted from the tooth's enamel by treatment with HCl acid and various pre-extraction cleaning and post-extraction sample preparation steps before LC-ESI tandem mass spectrometry (MS/MS) analysis is performed. Most methods sequence the doubly charged SM_{ox}IRPPY peptide ($[M+2H]^{2+}$ at m/z 440.22) from the Y-chromosomal amelogenin for positive identification of biological males. The LC-ESI MS/MS analysis alone typically requires more than one hour of time for each sample.

Recently, two new methods with shorter sample preparation and MS analysis times were published [8]. However, both methods still use ESI and an LC system to infuse the enamel extracts into the mass

* Corresponding author.
E-mail address: r.k.cramer@reading.ac.uk (R. Cramer).

https://doi.org/10.1016/j.jpba.2024.116599
Received 9 September 2024; Received in revised form 28 November 2024; Accepted 28 November 2024
Available online 30 November 2024

spectrometer, and the MS analysis times for subsequent sample analyses were > 2 min per sample.

Matrix-assisted laser desorption/ionization (MALDI) MS and MS/MS analysis of amelogenin peptides obtained from enamel has also been attempted but with lesser success for biological sex estimation as no Y chromosome-specific peptide has so far been detected [9–11]. Only a handful of human amelogenin peptides, all from the X chromosomal gene, were reported to be identified by MALDI TOF/TOF MS, with even fewer by MALDI-TOF/TOF MS/MS using collision-induced dissociation [9–11].

As a fundamentally new soft ionization technique, liquid atmospheric pressure (LAP)-MALDI facilitates the detection of multiply charged ions at low sample consumption and high ion yield stability [12]. It has been applied to various areas of biological mass spectrometry such as disease detection [13,14], microbial biotyping [15], lipid analysis [15,16], high-throughput analysis [17–19] as well as classical bottom-up proteomic workflows [20] and top-down proteomics [14,21,22].

Here, we present an MS and MS/MS workflow that utilizes the laser-based LAP-MALDI technique, enabling highly stable ion yields as well as the production of both singly and multiply charged amelogenin peptide ions. This workflow employs minimal sample preparation of sonication (~1 min), acid extraction (~5–6 min), micro-SPE clean-up using pipette tips (~1–2 min) and LAP-MALDI sample spotting (≤1 min). The MS data acquisition time is negligible while MS/MS data acquisition time per peptide is ~1 minute (see Fig. 1a). The advantage of this workflow is the simultaneous MS detection of all extracted amelogenin peptides and the absence of lengthy LC separation as is predominately used in current MS/MS analyses of amelogenin peptides for biological sex estimation. Potential applications can now include large-scale biological sex estimation without the hyphenation of analytical techniques at much shorter time scales. An additional advantage is the possibility to prepare the samples at remote sites without the need for specialized labs, only requiring small and inexpensive equipment, i.e. a sonicator and a pipettor.

2. Materials and methods

2.1. Reagents and chemicals

LC-MS-grade water, acetonitrile, and trifluoroacetic acid (TFA) and aqueous 37 % (v/v) hydrochloric acid (HCl), which was diluted with LC-MS-grade water to 10 %, were purchased from Fisher Scientific (Loughborough, UK). The matrix compounds α-cyano-4-hydroxycinnamic acid (CHCA) and ethylene glycol were purchased from Merck (Gillingham, UK).

2.2. Samples and osteological analysis

Human teeth were obtained from three sites of varying date (700–1850 CE) from individuals curated by the Department of Archaeology at the University of Reading. One individual (skeleton 2000) from St John's Church in Redhill, Surrey (UK) was a known male who died aged 22 years, based on information preserved on their coffin plate. All destructive sampling was carried out following the guidelines produced by the Advisory Panel on the Archaeology of Burials in England on destructive sampling of archaeological human remains (historicengland.org.uk/content/docs/advice/science-and-the-dead-2nd-ed/), and the British Association for Biological Anthropology and Osteoarchaeology (BABAO) Code of Ethics (babao.org.uk/wp-content/uploads/2024/01/BABAO-Code-of-Ethics.pdf). Table 1 provides the metadata for all teeth investigated. The age at death for each individual was estimated using standard osteological techniques for the skull, ribs and pelvis [23]. Biological sex was determined using an assessment of the morphology of the skull, humerus and pelvis, and measurements of the long bones and joint surfaces as previously described [23,24]. The accuracy of these methods depends on the experience of the osteologist but is estimated to be between 95 % and 100 % accurate when the skeleton is well-preserved [25].

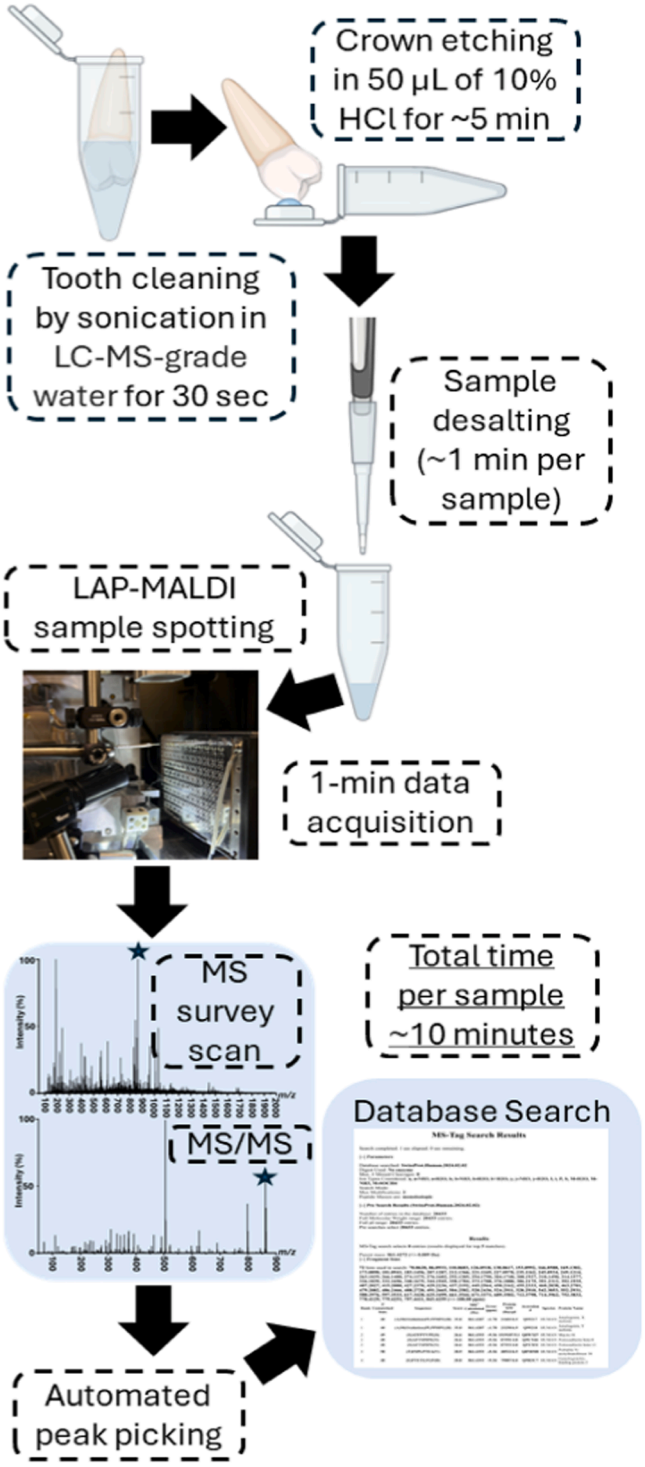

Fig. 1. Tooth sample preparation and peptide analysis workflow for biological sex estimation using LAP-MALDI MS.

2.3. Peptide extraction and purification

First, each tooth was individually cleaned in a 1.5-mL micro-centrifuge tube by sonication for 30 seconds in approximately 300 µL of LC-MS-grade water, using just enough volume to cover the tooth's

Table 1
Metadata and biological sex estimation data of the teeth analyzed in this study.

Skeleton Number	Site location	Period[a]	Sampled Tooth	Age (Osteology)	Sex (Osteology)	Sex (Enamel MS Analysis)
73	St Oswald's Priory, Gloucester, UK	1120–1230	Md L 1 pm	Adult	Male	Male
80	St Oswald's Priory, Gloucester, UK	1120–1230	Mx L 1 pm	17–25 years	Female	Female
109	St Oswald's Priory, Gloucester, UK	1120–1230	Md L 1 pm	36–45 years	Male	Male
369	St Oswald's Priory, Gloucester, UK	1250–1540	Md L 1 pm	26–35 years	Female	Female
467	St Oswald's Priory, Gloucester, UK	1540–1650	Md R 1 pm	26–25 years	Male	Male
1208	St John's Church, Redhill, Surrey, UK	1750–1850	Md R 1 pm	26–35 years	Female	Female
1648	St John's Church, Redhill, Surrey, UK	1750–1850	Md L 2 pm	26–35 years	Female	Female
2000[b]	St John's Church, Redhill, Surrey, UK	1893	Md R 1 pm	22 years	Male	Male
2720	Bishopstone, Sussex, UK	700–1020	Md L 1 pm	26–35 years	Male	Male
2792	Bishopstone, Sussex, UK	700–1020	Md R 2 pm	36–45 years	Female	Female

[a] Time period (CE) in which the individual died.
[b] The biological sex and age of this individual was determined from their coffin plate.
Md = mandible; Mx = maxillary; L = left; R = right; 1 pm = first premolar; 2 pm = second premolar.

crown. The tooth was then incubated in 50 μL of 10 % HCl (v/v) at room temperature for 1 minute and subsequently washed with LC-MS-grade water. Enamel peptides were then extracted in 50 μL of 10 % HCl, which was pipetted into the cap of a 1.5-mL microcentrifuge tube, by incubating the crown of the tooth in the 50-μL droplet for 5 minutes. The acid-etch solution was transferred to a new microcentrifuge tube followed by sample clean-up using C_{18} Zip-Tips (ZTC18S096; EMD Millipore, Merck, Gillingham, UK) according to the manufacturer's protocol. Peptides were eluted from the ZipTips with 10 μL of an acetonitrile/water solution (50/50; v/v).

2.4. MALDI matrix preparation

To prepare the LAP-MALDI matrix, CHCA was dissolved in water/acetonitrile (3/7; v/v) to a concentration of 25 mg/mL. This solution was thoroughly vortexed until CHCA was full dissolved, followed by the addition of ethylene glycol at 70 % by volume and subsequent vortexing to ensure the final liquid MALDI matrix was homogeneous.

2.5. LAP-MALDI MS and MS/MS

For LAP-MALDI MS analyses, 0.5 μL of the liquid MALDI matrix was spotted onto a stainless-steel MALDI sample plate followed by an equal volume of the analyte solution.

An in-house developed LAP-MALDI source coupled with a Synapt G2-Si (Waters Corp., Wilmslow, UK) was used in positive ion sensitivity mode. A detailed description of the in-house developed LAP-MALDI source can be found in a previous publication [26]. The pulsed beam of a 343-nm laser (FlareNX 343–0.2–2; Coherent, Santa Clara, USA) was focused on the centre of the liquid MALDI sample droplet, using pulse energies of approximately 10 μJ at a focus diameter of approximately 50–100 μm with a pulse repetition rate of 50 Hz. An extraction potential of 3.0 kV and a N_2 gas counterflow of 180 L/h was applied to the ion transfer tube. The instrument was manually calibrated over the m/z range of 50−2000 using caesium iodide and Intellistart software (MassLynx 4.2; Waters). Data acquisition was performed within an m/z range of 50−2000, with the scan time set to 1 second/scan. Each MS data acquisition and subsequent MS/MS data acquisitions lasted 1 minute.

For MS/MS analyses, the m/z value for the target precursor ion was selected, and the quadrupole isolation window was adjusted by setting the low-mass (LM) and high-mass (HM) resolution values around the m/z value selected for collision-induced dissociation (CID) MS/MS. The LM resolution value was set in the range of 4.5–4.8 and the HM resolution value was set in the range of 15–18. The CID collision voltage was set between 20 and 70 V, dependent on the precursor ion and its m/z value. With respect to the precursor ion selection for the MS/MS analysis of the enamel peptides of skeleton 73 (Table 2), a minimum signal-to-noise (S/N) ratio of 3 was applied, and all doubly and triply charged ions but only singly charged ions above m/z 600 were selected.

2.6. Data processing and analysis

Raw data files from MassLynx 4.2 (Waters) were processed with Mascot Distiller (Version 2.8.5, 64-bit; Matrix Science, London, England) for automated peak picking. A minimum S/N of 5 with baseline correction (isotope distribution with 500 maximum iterations per scan) and a correlation threshold (Rho) of 0.7 under "MS Peak Picking" as well as "MS/MS Peak Picking" were selected. The peak list was then exported as a mascot generic format (.mgf) containing the monoisotopic masses of the singly charged equivalents of the multiply charged fragment ions detected.

For peptide identification, the .mgf files were searched against 'SwissProt.Human.2024.02.02' (20433 entries) using the MS-Tag search software of ProteinProspector (https://prospector.ucsf.edu/). 'No enzyme' was selected and 'Oxidation (M)' was chosen as variable modification in the search parameters. 'MALDI-QUAD-TOF' was selected as instrument type with M-$SOCH_4$ as an additional ion type to be considered. The number of missed cleavages was set to 0, and the mass tolerances were set at 10 ppm and 100 ppm for the precursor and product ions, respectively. If these searches were unsuccessful, the precursor ion mass tolerance was increased to 20 ppm. In some cases, where the isotopologue distribution indicated the possibility of Q or N deamidation, 'Deamidated (N)' and 'Deamidated (Q)' were also chosen as variable modifications, and the 'PEPMASS' value in the .mgf file was changed by taking into account the potential addition of 0.9840 Da to the precursor ion mass as a result of deamidation.

The Mascot Distiller exported peak lists for the peptides SM(ox)IRPPY (m/z 879.4393) and SIRPPYP (m/z 829.4567) obtained from all teeth were further processed using a script described previously [21]. In brief, fragment ion signals from each MS/MS analysis acquired from the same peptide were grouped within a tolerance of $m/z \pm 0.05$ in a 'Male' and 'Female' group, according to the osteology results. Fragment ions in each group were filtered by using a minimum intensity threshold of 10 and being present in the MS/MS spectra of all 5 peak lists, resulting in a single peak list containing only common fragment ions to all 5 MS/MS analyses for each peptide and each sex, i.e. a total of four peak lists of common fragment ions.

3. Results and discussion

3.1. LAP-MALDI MS and MS/MS analysis of enamel extracts

In total, ten human teeth (Table 1) were each subject to a simple and quick peptide extraction workflow (Fig. 1) using less destructive and time-consuming steps than those in previously described methods [4,7,27,28]. LAP-MALDI MS analysis of all teeth revealed a plethora of extracted peptides (Fig. 2), which were further analyzed by MS/MS. Of the 39 MS/MS analyses performed on the peptide extract from the tooth of skeleton 73 (cf. Table 1), the MS/MS data acquisitions of 27 precursor

Table 2
Identifications of amelogenin and ameloblastin peptides extracted from the male tooth of skeleton 73 (cf. Table 1).

Observed *m/z*	Theoretical *m/z*	Z[†]	Δ (ppm)	Peptide Sequence	Isoform[§]
423.2232	423.2205	2	6.38	MPLPPHPG (17−24)	XY
431.2203	431.2180	2	5.33	M(oxidized)PLPPHPG (17−24)	XY
558.9468	558.9414	3	9.66	MPLPPHPGHPGYINF (17−31)	XY
564.2783	564.2731	3	9.22	M(oxidized) PLPPHPGHPGYINF (17−31)	XY
617.3414	617.3406	1	1.30	LPPHPG (19−24)*	XY
723.8619	723.8619	2	0.00	LPPHPGHPGYINF (19−31)	XY
742.4177	742.4246	1	−9.29	IRPPYP (45−50)*	X
829.4547	829.4567	1	−2.41	SIRPPYP (44−50)	X
837.9077	837.9085	2	−0.95	MPLPPHPGHPGYINF (17−31)	XY
845.4304	845.4338	1	−4.02	MPLPPHPG (17−24)	XY
861.4272	861.4287	1	−1.74	M(oxidized)PLPPHPG (17−24)	XY
879.4396	879.4393	1	0.34	SM(oxidized)IRPPY (58−64)*	Y
908.4708	908.4737	1	−3.19	LPPHPGHPG (19−27)	XY
951.4391	951.4458	1	−7.04	VPFFPQQS (27−34) (ameloblastin)	N/A
962.5542	962.5557	1	−1.56	YEVLTPLK (33−40)	XY[#]
1008.4545	1008.4673	1	−12.69	VPFFPQQSG (27−35) (ameloblastin)	N/A
1049.5898	1049.5877	1	2.00	SYEVLTPLK (32−40)	XY[#]
1079.5385	1079.5455	1	−6.48	MPLPPHPGHP (17−26)	XY
1136.5565	1136.5670	1	−9.24	MPLPPHPGHPG (17−27)	XY
1152.5471	1152.5619	1	−12.84	M(oxidized) PLPPHPGHPG (17−27)	XY
1236.5809	1236.5796	1	1.05	PHPGHPGYINF (21−31)	XY
1446.7159	1446.7165	1	−0.41	LPPHPGHPGYINF (19−31)	XY
1480.6841	1480.6777	1	4.32	VPFFPQQSGTPGM (oxidized)A (27−40) (ameloblastin)	N/A
1527.7131	1527.7413	1	−18.46	MPLPPHPGHPGYIN (17−30)*	XY
1543.7577	1543.7692	1	−7.45	PLPPHPGHPGYINF (18−31)	XY
1674.7982	1674.8097	1	−6.87	MPLPPHPGHPGYINF (17−31)	XY
1690.8020	1690.8046	1	−1.54	M(oxidized) PLPPHPGHPGYINF (17−31)	XY

ProteinProspector MS-Tag was used for identification (for search parameters see Experimental Section). All identified peptides are from amelogenin (Q99217 or Q99218), or where stated ameloblastin (Q9NP70), and are top matches apart from the matches labelled with *, which were within the top 5 matches. Some peptides showed Q/N deamidation (not detailed in this table) – see text for more details. [†] Detected ion charge state. [§] Amelogenin isoform details with 'X' indicating X-chromosome-specific isoform, 'Y' indicating Y-chromosome-specific isoform and 'XY' indicating lack of specificity due to sequence occurrence in isoforms from both sex chromosomes. [#] Note that the search results returned only the X isoform as the Y-isoform database entry did not include the loss of NSHSQAINVDRIAL in the mature Y isoform.

ions were of sufficient quality for peptide identification and could be assigned to peptides from human amelogenin or ameloblastin (see Table 2). In most cases, the precursor ion mass measurement accuracy was within ± 10 ppm.

As expected from acid hydrolysis, several peptides that were detected cover the same protein sequence stretches and show no specific cleavage sites. Most peptide precursor ions were singly charged, in particular the shorter peptide ions, while some longer peptide ions were also detected as doubly and triply charged ions. Multiply charged peptide ion production is facilitated by LAP-MALDI as reported previously [12,17,29]. Both amelogenin and ameloblastin have proline-rich regions, and peptides from these regions are well represented in the set of identified peptides (see Table 2).

3.2. Amelogenin maturation analysis using LAP-MALDI MS and MS/MS data

The amelogenin peptides identified by LAP-MALDI MS/MS indicate the presence of the mature amelogenin protein in teeth. None of the peptides identified covers the signal region (1−16) of amelogenin, while several peptides were obtained starting with the first amino acid after the signal region. Intriguingly, the amelogenin peptide data also supports previously published mRNA data [30], showing that exon 4 is spliced out as can be seen by the occurrence of the peptides YEVLTPLK and SYEVLTPLK. Both peptides cover amino acids of exons 3 and 5 but are missing the 14 amino acids that exon 4 codes for. While the SwissProt database entry for the Y-chromosome amelogenin includes the amino acids from exon 4, the entry for the X-chromosome amelogenin does not. Therefore, the search results in Table 2 were accordingly adjusted for these two peptides, indicating a match for both the X- and Y-chromosomal amelogenin.

Taking into account all processing steps and degradation of the secreted amelogenin protein as reported by Brookes et al., maturation in the enamel typically leads to the small protein fragment called tyrosine-rich amelogenin peptide (TRAP), comprising 44 amino acids, i.e. 16 (Met)-60(Trp), and 45 amino acids, i.e. 16(Met)-61(Trp), for the X-chromosome and Y-chromosome version, respectively. All amelogenin peptides identified by the presented workflow originate from TRAP, covering 31 amino acids (~70 %) of TRAP for both X- and Y-chromosomal amelogenin.

3.3. Oxidation and deamidation of enamel peptides

Owing to the antiquity of the teeth and the environmental conditions they have likely been exposed to, modifications such as oxidation of methionine and deamidation of asparagine and glutamine can be observed. Panels A) and B) in Fig. 3 show the LAP-MALDI MS/MS spectra of the singly charged and doubly charged M(ox)PLPPHPG peptide, respectively. The characteristic fragment ions of methionine oxidation due to the neutral loss of 64 Da (-$SOCH_4$) can be easily assigned.

Deamidation was evident from the MS survey scans of all samples through the 0.9840-Da mass shift associated with deamidation of N and Q (resulting in D and E, respectively), with isotopologue distributions characteristic of partial deamidation. Fragment ion signals further supported the observation of deamidation by revealing the same characteristic isotopologue distribution for all fragment ions that included N or Q as a residue but being absent for all other fragment ions. Panel C) and D) in Fig. 3 show the LAP-MALDI MS/MS spectra of the non-oxidized and oxidized N-deamidated MPLPPHPGHPGYIN(deamid)F and M(ox) PLPPHPGHPGYIN(deamid)F, respectively.

If deamidation was likely as judged by the peptide's isotopologue pattern, the precursor ion masses used for database searching were in some cases adjusted by the 0.9840-Da mass difference because in these cases the automated peak picking could not account for (partial) deamidation. Fig. 4 and its insets demonstrate the effect of glutamine deamidation for the two ameloblastin peptides VPFFPQQS and VPFFPQQSG, where the insets show the isotopologue patterns for the precursor ions. No glutamine was present in any amelogenin peptides that were identified, and therefore only asparagine deamidation was observed. Asparagine deamidation was detected in all asparagine-containing peptides but to a lower extent than glutamine deamidation. Further studies are necessary to elucidate whether the sample preparation promotes deamidation and whether deamidation analysis by LAP-

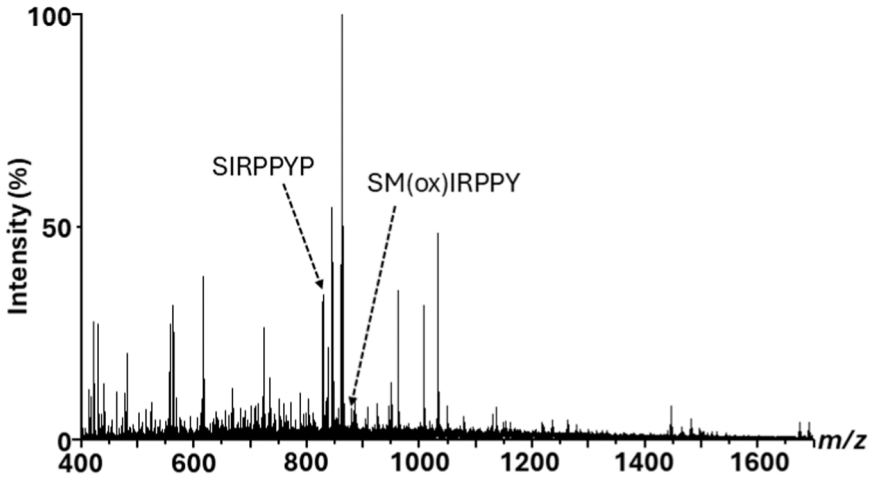

Fig. 2. LAP-MALDI mass spectrum of the extracted peptides from the male tooth of skeleton 73 (cf. Table 1 and Table 2).

MALDI could also be used for sample dating.

3.4. Biological sex estimation

For amelogenin, a notable dimorphic difference between the X- and Y-chromosomal amelogenin can be found in the sequence region where LAP-MALDI MS(/MS) analysis detected the peptides SIRPPYP and SM(ox)IRPPY, respectively. Both peptide sequences are coded for by exon 5 of the amelogenin gene. However, the Y-chromosomal isoform possesses an additional methionine, and the presence of methionine-containing peptides covering this sequence region is an unambiguous indicator of male sex. Here, SM(ox)IRPPY with a theoretical *m/z* value of 879.44 for the $[M+H]^+$ ion and SIRPPYP with a theoretical *m/z* value of 829.46 for the $[M+H]^+$ ion were used to assess biological sex. From the MS survey scan, the ion signal intensity of the SIRPPYP peptide was much greater than that of SM(ox)IRPPY (>4 for the S/N values; see Fig. 2), which is likely due to the differences in expression of the two amelogenin isoforms. It was reported that human Y-chromosomal amelogenin transcription can be as low as 10 % of the total amelogenin transcription [31].

The LAP-MALDI MS/MS spectra of SM(ox)IRPPY and SIRPPYP are shown in Fig. 5. In all cases, teeth previously assigned as belonging to biological males through standard osteological methods showed the presence of SM(ox)IRPPY, confirming the assignment and demonstrating a 100 % success rate for male identification. Even in cases where the LAP-MALDI MS ion signal intensity for the SM(ox)IRPPY ion peak at *m/z* 879.4393 was just above the noise level, when isolated and fragmented using CID, a range of specific fragment ions were observed at high ion signal intensities. When isolating the same *m/z*-value window for MS/MS analysis of previously assigned female teeth, no fragment ions associated to SM(ox)IRPPY were obtained by MS/MS.

The MS/MS fragment ion data of both SM(ox)IRPPY and SIRPPYP for both male and female were subject to sex- and peptide-specific peak list filtering, resulting in four peak lists containing only the common fragment ions found in all five MS/MS analyses for each peptide for each sex (Table 3). For the MS/MS data of *m/z* 829 (SIRPPYP), a total of 29 ions (including immonium and internal fragment ions) were found to be in all MS/MS spectra for the analyses performed on the female and male samples. For the MS/MS data of *m/z* 879 (SM(ox)IRPPY), a total of 13 ions (including immonium and internal fragment ions) were common in all male teeth, with no common fragment ions appearing in the spectra of the female samples as expected.

Based on the data of Table 3, several SM(ox)IRPPY-specific fragment ions can be used to provide high accuracy in assigning male sex. For female sex assignment, it seems to be reasonable to include the absence (defined by a specific S/N threshold value) of the precursor ion signal at *m/z* 879 and SM(ox)IRPPY-specific fragment ions after MS/MS analysis in addition to the presence of *m/z* 829 ions as well as some of the SIRPPYP-specific fragment ions at well-defined ion signal intensities. From the presented data, we suggest using the presence of the MH-$SOCH_4$, b_4-$SOCH_4$ and y_3 fragment ions (at the MS/MS analysis of *m/z* 879) with an S/N value above 10 for male sex assignment. For female sex assignment, we suggest using the absence of all of the above fragment ion signals (apart from y_3) as defined above and include the presence of the *m/z* 829 precursor ion as well as its a_2, b_3, y_3 and y_4 fragment ions using the same S/N value as above. The latter requirements are related to a minimum abundance level of SIRPPYP, making sure that overall amelogenin peptide recovery is given. The above suggestion for assigning biological male and female sex provides 100 % classification accuracy for the investigated sample set but will obviously benefit from further validation by a larger sample set.

It is worth noting that deletion of the Y-chromosomal amelogenin can occur in almost all populations around the world, albeit at a level significantly below 1 % (in many cases well below 0.1 %) but with the frequency of deletion being much higher in biological males of some populations of the Indian subcontinent [32]. In these cases, it would not be possible to determine biological male from biological female. In this dataset, no individuals have been identified as having Indian ancestry and all of the teeth from the five biological males (as characterized by osteology) were confirmed as biological male.

4. Conclusions

The detection of sex chromosome-specific amelogenin peptides from the enamel of archaeological and forensic tooth samples enables accurate biological sex estimation. In this study, we demonstrated that direct LAP-MALDI MS(/MS) analysis of enamel peptides extracted by using a simple and rapid workflow is a robust and accurate method to assign the biological sex of archaeological human skeletal remains. The detection of the X-chromosomal SM(ox)IRPPY peptide is an unambiguous marker of the male sex, and the LAP-MALDI MS(/MS) data from this study achieved a 100 % accuracy in biological sex determination using the osteological information for the individuals as benchmark data. The presented analytical workflow takes only 10 minutes per tooth, from sample preparation to complete data analysis, with no LC separation required. Therefore, this methodology facilitates high throughput analysis of a larger number of individuals compared to current methods. It also offers the possibility to prepare samples in remote and basic laboratories as only inexpensive and minimally hazardous consumables (<$2 per sample using 10 % HCl) and equipment (pipettor and sonicator) are required. Owing to the speed and sensitivity of LAP-MALDI MS(/MS), further in-depth rapid analyses are also possible as

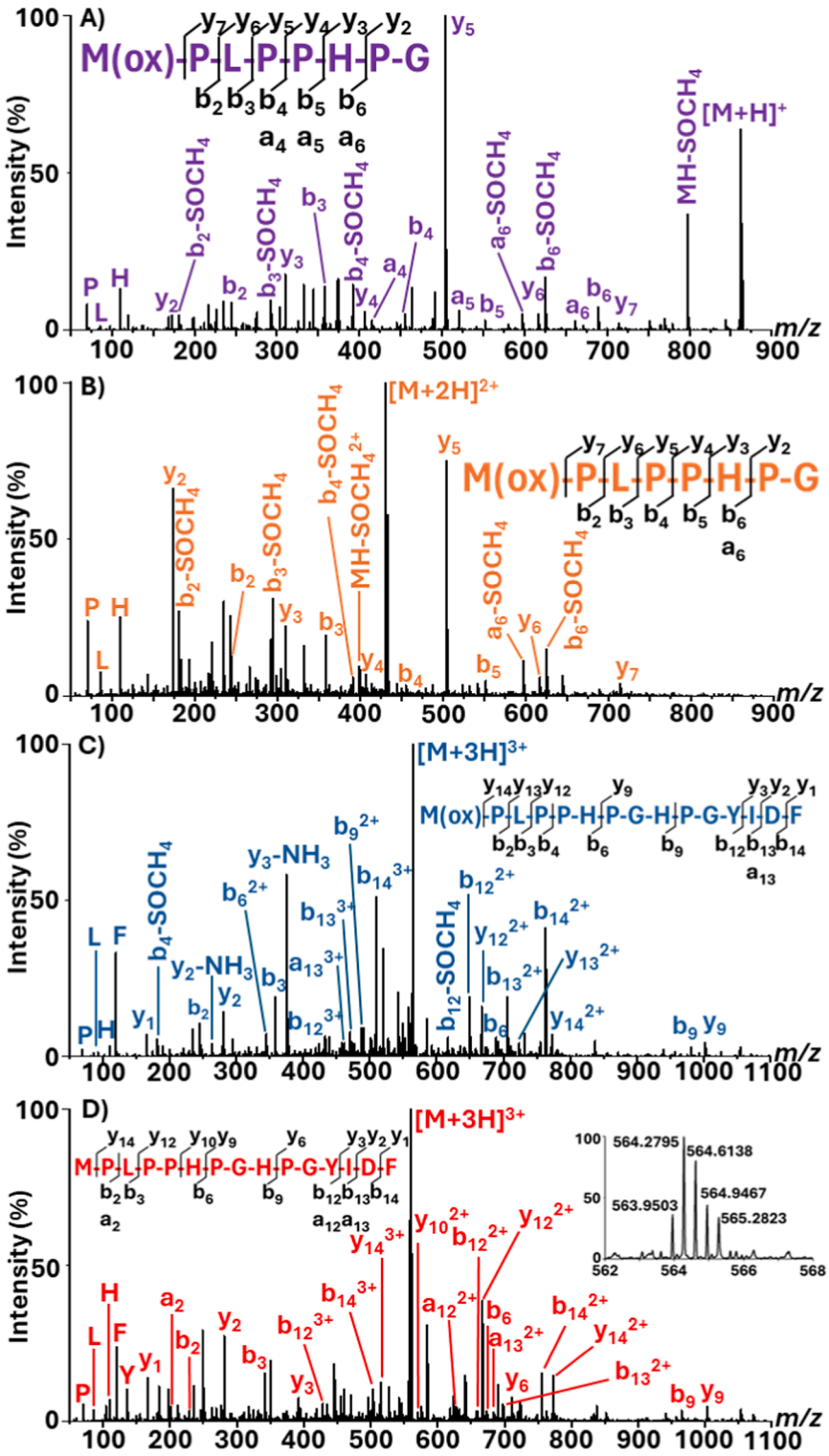

Fig. 3. LAP-MALDI MS/MS spectra of (A) *m/z* 861 (singly charged ion of M(ox)PLPPHPG), (B) *m/z* 431 (doubly charged ion of M(ox)PLPPHPG), (C) *m/z* 558 (triply charged ion of MPLPPHPGHPGYIN(deamid)F) and *m/z* 563 (triply charged ion of M(ox)PLPPHPGHPGYIN(deamid)F). Inset in (D) shows the enlarged *m/z* region of the precursor ion.

demonstrated by the characterization of peptide modifications such as the deamidation of asparagines and glutamines. In the future, these analyses could potentially enable the dating of human remains as well as their biological sex estimation. The presented data also support previous studies of amelogenin, reporting the deletion of exon 4 and the degradation of the mature amelogenin protein in enamel to the tyrosine-rich amelogenin peptide (TRAP). Importantly, the presented method is minimally destructive with only a small area of the tooth being acid-etched, keeping the analyzed tooth intact for long-term heritage preservation.

CRediT authorship contribution statement

Matthew J Collins: Writing – review & editing, Validation. **Rainer**

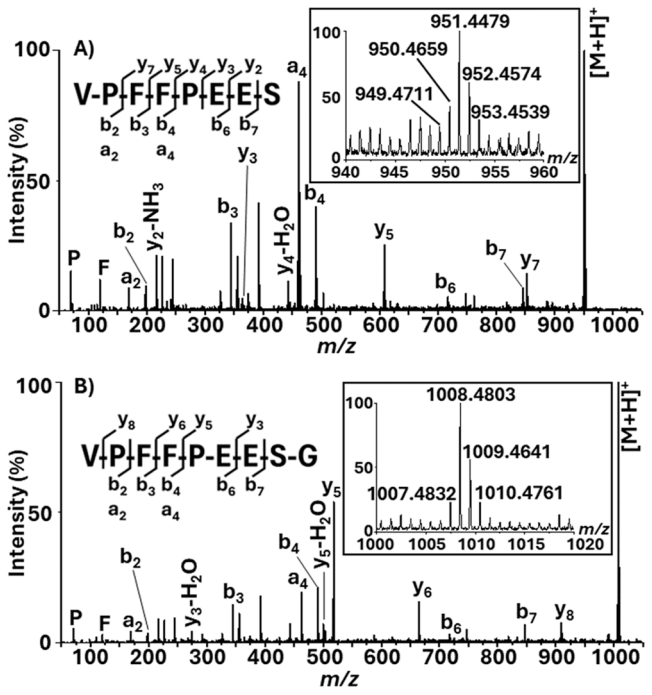

Fig. 4. LAP-MALDI MS/MS spectra of (A) *m/z* 951 (singly charged VPFFPQ (deamid)Q(deamid)S) and (B) *m/z* 1008 (singly charged VPFFPQ(deamid)Q (deamid)SG).

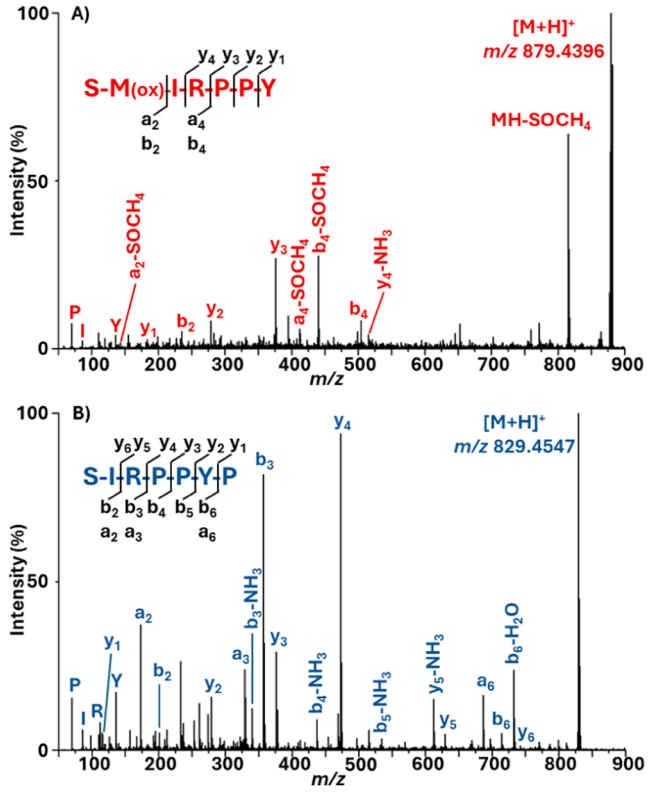

Fig. 5. LAP-MALDI MS/MS spectra of (A) *m/z 879* (SM(ox)IRPPY) and (B) *m/z* 829 (SIRPPYP).

Cramer: Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Mary E Lewis:** Writing – review & editing, Validation, Resources, Methodology. **Lily R Adair:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis.

Table 3
Common fragment ions of *m/z* 829 (SIRPPYP) and *m/z* 879 (SM_{ox}IRPPY) amongst all male (n = 5) and amongst all female (n = 5) teeth.

Theoretical *m/z*	SIRPPYP		Theoretical *m/z*	SM_{ox}IRPPY	
	Male	Female		Male	Female
70.0651	Pro*	Pro*	70.0651	Pro*	-
86.0964	Ile*	Ile*	86.0964	Ile*	-
87.0917	Arg*	Arg*	87.0917	Arg*	-
116.0706	y_1	y_1	136.0757	Tyr*	-
126.0550	Pro*	Pro*	143.0815	a_2-$SOCH_4$	-
136.0757	Tyr*	Tyr*	235.0747	b_2	-
167.1179	ProPro-CO**	ProPro-CO**	279.1339	y_2	-
173.1285	a_2	a_2	376.1867	y_3	-
195.1128	ProPro**	ProPro**	412.2667	a_4-$SOCH_4$	-
201.1234	b_2	b_2	440.2616	b_4-$SOCH_4$	-
233.1285	ProTyr-CO**	ProTyr-CO**	504.2599	b_4	-
237.1346	ArgPro-NH_3**	ArgPro-NH_3**	515.2613	y_4-NH_3	-
253.1659	IleArg-NH_3**	IleArg-NH_3**	815.4410	M+H-$SOCH_4$	-
261.1234	ProTyr**	ProTyr**			
279.1339	y_2	y_2			
329.2296	a_3	a_3			
330.1812	ProProTyr-CO**	ProProTyr-CO**			
340.1979	b_3-NH_3	b_3-NH_3			
357.2245	b_3	b_3			
358.1761	ProProTyr**	ProProTyr**			
376.1867	y_3	y_3			
437.2507	b_4-NH_3	b_4-NH_3			
473.2395	y_4	y_4			
534.3035	b_5-NH_3	-			
612.3140	y_5-NH_3	y_5-NH_3			
629.3406	y_5	y_5			
686.3984	a_6	a_6			
697.3668	-	b_6-NH_3			
714.3933	b_6	b_6			
732.4039	b_6+H_2O	b_6+H_2O			

See text for definition of common fragment ions.
* Immonium or related ions;
** Internal fragment ions.
Fragment ions in this table with the same *m/z* value for both peptides are highlighted in dark grey. Fragment ions (based on theoretical immonium, internal, and a, b and y ions and related ions) specific to each peptide (within *m/z* 0.9) are highlighted in light grey.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Rainer Cramer reports financial support was provided by EPSRC. Rainer Cramer has patent Method for Ion Production issued to University of Reading. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported by the Engineering and Physical Sciences Research Council (EPSRC) through grant EP/V047485/1. The graphical abstract and Fig. 1 were created using BioRender (BioRender.com).

References

[1] Y. Nakahori, O. Takenaka, Y. Nakagome, A Human X-Y homologous region encodes amelogenin, Genomics 9 (2) (1991) 264–269.

[2] A. Akane, H. Shiono, K. Matsubara, Y. Nakahori, S. Seki, S. Nagafuchi, M. Yamada, Y. Nakagome, Sex identification of forensic specimens by polymerase chain-reaction (Pcr) - 2 alternative methods, Forensic Sci. Int. 49 (1) (1991) 81–88.

[3] Y. Nakahori, K. Hamano, M. Iwaya, Y. Nakagome, Sex identification by polymerase chain-reaction using X-Y homologous primer, Am. J. Med. Genet. 39 (4) (1991) 472–473.

[4] R. Gowland, N.A. Stewart, K.D. Crowder, C. Hodson, H. Shaw, K.J. Gron, J. Montgomery, Sex estimation of teeth at different developmental stages using dimorphic enamel peptide analysis, Am. J. Phys. Anthr. 174 (4) (2021) 859–869.

[5] E. Greco, A.M. Gennaro, D. Piombino-Mascali, D. Costanzo, S. Accardo, S. Licen, P. Barbieri, S. Fornasaro, S. Semeraro, E. Marin, S. Signoretti, C. Gabriele, M. Gaspari, Dental proteomic analyses and Raman spectroscopy for the estimation of the biological sex and age of human remains from the Greek cemetery of San Giorgio Extra, Reggio Calabria (Italy), Microchem. J. 195 (2023).

[6] B. Shaw, S. Foggin, P. Hamilton-Stanley, A. Barlow, C. Pickard, L. Fibiger, N. Oldham, P. Tighe, L.M. Kootker, S. Schrader, R. Layfield, Antibody-based sex determination of human skeletal remains, iScience 26 (11) (2023).

[7] N.A. Stewart, R.F. Gerlach, R.L. Gowland, K.J. Gron, J. Montgomery, Sex determination of human remains from peptides in tooth enamel, Proc. Natl. Acad. Sci. USA 114 (52) (2017) 13649–13654.

[8] A.M. Casas-Ferreira, M. del Nogal-Sanchez, A.E. Arroyo, J.V. Vazquez, J.L. Perez-Pavon, Fast methods based on mass spectrometry for peptide identification. Application to sex determination of human remains in tooth enamel, Microchem. J. 181 (2022).

[9] C.M. Nielsen-Marsh, C. Stegemann, R. Hoffmann, T. Smith, R. Feeney, M. Toussaint, K. Harvati, E. Panagopoulou, J.J. Hublin, M.P. Richards, Extraction and sequencing of human and Neanderthal mature enamel proteins using MALDI-TOF/TOF MS, J. Archaeol. Sci. 36 (8) (2009) 1758–1763.

[10] I.M. Porto, H.J. Laure, F.B. de Sousa, J.C. Rosa, R.F. Gerlach, New techniques for the recovery of small amounts of mature enamel proteins, J. Archaeol. Sci. 38 (12) (2011) 3596–3604.

[11] I.M. Porto, H.J. Laure, R.H. Tykot, F.B. de Sousa, J.C. Rosa, R.F. Gerlach, Recovery and identification of mature enamel proteins in ancient teeth, Eur. J. Oral. Sci. 119 (2011) 83–87.

[12] R. Cramer, A. Pirkl, F. Hillenkamp, K. Dreisewerd, Liquid AP-UV-MALDI enables stable ion yields of multiply charged peptide and protein ions for sensitive analysis by mass spectrometry, Angew. Chem. Int. Ed. 52 (8) (2013) 2364–2367.

[13] S.E. Lellman, C.K. Reynolds, A.K.B. Jones, N. Taylor, R. Cramer, LAP-MALDI MS profiling and identification of potential biomarkers for the detection of bovine tuberculosis, J. Agr. Food Chem. 71 (37) (2023) 13899–13905.

[14] C. Piras, O.J. Hale, C.K. Reynolds, A.K. Jones, N. Taylor, M. Morris, R. Cramer, LAP-MALDI MS coupled with machine learning: an ambient mass spectrometry approach for high-throughput diagnostics, Chem. Sci. 13 (6) (2022) 1746–1758.

[15] S.E. Lellman, R. Cramer, Bacterial identification by lipid profiling using liquid atmospheric pressure matrix-assisted laser desorption/ionization mass spectrometry, Clin. Chem. Lab Med. 58 (6) (2020) 930–938.

[16] O.J. Hale, R. Cramer, Collision-induced dissociation of doubly-charged barium-cationized lipids generated from liquid samples by atmospheric pressure matrix-assisted laser desorption/ionization provides structurally diagnostic product ions, Anal. Bioanal. Chem. 410 (5) (2018) 1435–1444.

[17] B. Challen, M. Morris, R. Cramer, Ultra-high-throughput and low-volume analysis of intact proteins with LAP-MALDI MS, J. Am. Soc. Mass Spectr. 34 (6) (2023) 991–994.

[18] H. Krenkel, J. Brown, M. Morris, R. Cramer, The use of salts, buffers and surfactants in LAP-MALDI MS, Int J. Mass Spectrom. 493 (2023).

[19] H. Krenkel, J. Brown, K. Richardson, E. Hoyes, M. Morris, R. Cramer, Ultrahigh-throughput sample analysis using liquid atmospheric pressure matrix-assisted laser desorption/ionization mass spectrometry, Anal. Chem. 94 (10) (2022) 4141–4145.

[20] P. Ryumin, J. Brown, M. Morris, R. Cramer, Protein identification using a nanoUHPLC-AP-MALDI MS/MS workflow with CID of multiply charged proteolytic peptides, Int. J. Mass Spectrom. 416 (2017) 20–28.

[21] L.R. Adair, I. Jones, R. Cramer, Utilizing precursor ion connectivity of different charge states to improve peptide and protein identification in MS/MS analysis, Anal. Chem. 96 (3) (2024) 985–990.

[22] O.J. Hale, M. Morris, B. Jones, C.K. Reynolds, R. Cramer, Liquid atmospheric pressure matrix-assisted laser desorption/ionization mass spectrometry adds enhanced functionalities to MALDI MS profiling for disease diagnostics, ACS Omega 4 (7) (2019) 12759–12765.

[23] T.D. White, M.T. Black, P.A. Folkens. Human Osteology, 3rd Edition ed., Elsevier/AP, Amsterdam; Boston, 2011.

[24] C.G. Falys, H. Schutkowski, D.A. Weston, The distal humerus–a blind test of Rogers' sexing technique using a documented skeletal collection, J. Forensic Sci. 50 (6) (2005) 1289–1293.

[25] J.E. Buikstra, J.H. Mielke, Demography, diet, and health, in: R.I. Gilbert, J. H. Mielke (Eds.), Analysis of Prehistoric Diets, Academic Press, Orlando, 1985, pp. 359–422.

[26] P. Ryumin, J. Brown, M. Morris, R. Cramer, Investigation and optimization of parameters affecting the multiply charged ion yield in AP-MALDI MS, Methods 104 (2016) 11–20.

[27] F. Lugli, G. Di Rocco, A. Vazzana, F. Genovese, D. Pinetti, E. Cilli, M.C. Carile, S. Silvestrini, G. Gabanini, S. Arrighi, L. Buti, E. Bortolini, A. Cipriani, C. Figus, G. Marciani, G. Oxilia, M. Romandini, R. Sorrentino, M. Sola, S. Benazzi, Enamel peptides reveal the sex of the Late Antique 'Lovers of Modena', Sci. Rep. 9 (2019) 13130.

[28] G.J. Parker, J.M. Yip, J.W. Eerkens, M. Salemi, B. Durbin-Johnson, C. Kiesow, R. Haas, J.E. Buikstra, H. Klaus, L.A. Regan, D.M. Rocke, B.S. Phinney, Sex estimation using sexually dimorphic amelogenin protein fragments in human enamel, J. Archaeol. Sci. 101 (2019) 169–180.

[29] O.J. Hale, P. Ryumin, J.M. Brown, M. Morris, R. Cramer, Production and analysis of multiply charged negative ions by liquid atmospheric pressure matrix-assisted laser desorption/ionization mass spectrometry, Rapid Commun. Mass Spectrom. 35 (2021) e8246.

[30] S.J. Brookes, C. Robinson, J. Kirkham, W.A. Bonass, Biochemistry and molecular-biology of amelogenin proteins of developing dental enamel, Arch. Oral. Biol. 40 (1) (1995) 1–14.

[31] E.C. Salido, P.H. Yen, K. Koprivnikar, L.C. Yu, L.J. Shapiro, The human enamel protein gene amelogenin is expressed from both the X-chromosomes and the Y-chromosomes, Am. J. Hum. Genet. 50 (2) (1992) 303–316.

[32] I. Stamfelj, Sex estimation based on the analysis of enamel peptides: false assignments due to AMELY deletion, J. Archaeol. Sci. 130 (2021) 105345.