FT-Raman spectroscopy as a method for screening collagen diagenesis in bone

Christine A.M. France, Daniel B. Thomas, Charlotte R. Doney, Odile Madden

Keywords: Raman; Bone; Collagen; Bioapatite; Diagenesis

This study examines Fourier transform (FT) Raman spectroscopy as a non-destructive screening method to determine collagen quality in archaeological and paleontological bones. Bone samples were characterized for collagen quality using well-established elemental abundance analyses (i.e., percentage nitrogen and C:N) as the primary criteria for classification. FT-Raman spectra were collected from outer surfaces and freshly cut cross-sections of both well-preserved and poorly-preserved historic mammal bones. Peak heights and peak areas were studied visually and with bivariate and multivariate statistics. Raman spectra from cross-sections provided the most accurate determination of collagen quality. A ratio of the 960 cm⁻¹ and 1636 cm⁻¹ peak heights provided the most unambiguous distinction between bones with well-preserved and poorly-preserved collagen. The 960 cm⁻¹ and 1636 cm⁻¹ peaks represent phosphate anion stretching in the bone mineral and amide carbonyl stretching in the collagen, respectively. FT-Raman spectra from bones with well-preserved collagen produced a 960 cm⁻¹:1636 cm⁻¹ ratio of 19.4 or less (after peaks were baseline corrected). This mineral to collagen ratio was typically greater in poorly-preserved samples as organic material tends to be more susceptible to early stages of diagenesis. These criteria now can be used to accurately determine collagen quality in bones before sacrificing samples to the lengthy and destructive chemical extractions necessary for carbon-14 dating, stable isotope analyses, proteomic analyses, and other techniques of archaeological or paleontological interest.

1. Introduction

1.1. Collagen

In recent years numerous applications for examining bone collagen in paleontological and archaeological specimens have emerged. Stable isotope analyses are ubiquitous in studies of past diets, ecosystem reconstruction, past climates, and migrations. Carbon-14 dating often is performed on extracted collagen to determine age of specimens. Additionally amino acid sequencing and structural analyses of larger proteins provide insight into evolutionary trends and alternative dating methods. Inherent in the use of bone collagen for all such analyses is the preservation of intact unaltered protein.

Collagen is a strong abundant protein comprising approximately 10–25% of total bone weight, most of which is type I collagen (Fratzl, 2008; Kucharcz, 1992; Nimni, 1988; Ramachandran and Reddi, 1976). Previous analyses of type I collagen to evaluate degradation have relied primarily on the characteristic elemental abundance and amino acid profile of collagen compared to other vertebrate proteins and organic material found in burial environments (Ajie et al., 1991; DeNiro and Weiner, 1988; Hare, 1980; Ho, 1965; Kucharcz, 1992; Liden et al., 1995; Nimni, 1988; Ostrom et al., 1994). Commonly used indicators of collagen quality include ~31–37 mol% glycine, ~8–14 mol% proline, ~7–11 mol% hydroxyproline, ~11–16% total nitrogen content, ~30–45% total carbon content, and a C:N ratio of 2.8–3.6 (Ajie et al., 1991; Ambrose, 1990; Bocherens et al., 1991, 1994, 1996, 1997; Coltrain et al., 2004; DeNiro, 1985; Drucker et al., 2001, 2003; Jorkov et al., 2007; McNulty et al., 2002). Limited data is available from spectroscopic examination of collagen quality (Chadebols et al., 2009; Lebon et al., 2011; Turner-Walker and Syversen, 2002). However, spectroscopic, amino acid, and elemental abundance analyses require considerable time, destructive sampling, exposure to embedding resins or chemical extractions, and expensive mass spectrometry which may ultimately indicate a poorly-preserved and unusable specimen.

Analysis of bone mineral (i.e., bioapatite) has provided an alternate approach for examining collagen degradation. Bioapatite
grains are nested within the collagen matrix and diagenesis of the former may imply degradation of the latter (Nelson et al., 1986; Person et al., 1996; Tütken et al., 2008; Veis, 2003). Several methods for assessing bone mineral alteration have been proposed (Iacumin et al., 1996; Kohn et al., 1999; Michel et al., 1995; Person et al., 1995, 1996; Shemesh, 1990; Tuross et al., 1989; Zazzo et al., 2004), but many are destructive and arguably are supported less well than the direct elemental and amino acid abundance measures of collagen itself. In regards to spectroscopic methods, destructive mid-infrared spectroscopy has received more attention for tracking inorganic bone diagenesis (Barrick and Showers, 1994; Lee-Thorp and Sponheimer, 2003) than non-destructive Raman spectroscopy (McLaughlin and Lednev, 2011; Thomas et al., 2011).

1.2. Raman spectroscopy of bone

Raman spectroscopy is a non-destructive analytical method that describes functional groups in molecules and polyatomic ions in minerals (Smith and Dent, 2005). The technique uses a laser to scatter light from a sample which allows the analysis to be performed at a distance with little to no sample preparation. For this reason, the technique is considered non-invasive and non-contact. We assessed whether Raman spectroscopy could provide a rapid and non-destructive method of evaluating collagen preservation with particular interest in identifying Raman spectral regions that change as collagen degrades.

Peaks in Raman spectra are counts of scattered photons, and peak intensities can be used to compare the concentrations of specific chemical bonds among similar samples. Characteristic bone collagen peaks will be larger relative to bioapatite peaks when there is more collagen present and vice versa. One limitation of Raman spectroscopy is interference from sample fluorescence which creates a broader strong signal that can overwhelm sharper but less intense Raman peaks. Fourier transform Raman (FT-Raman), used in this study, uses a near infrared excitation laser which induces less fluorescence in teeth and bone (Edwards et al., 2005).

A Raman spectrum of bone (Fig. 1A) contains chemical and structural information about collagen and bioapatite that is consistent across animal species (Frushour and Koenig, 1975; Morris and Mandair, 2011). Two prominent spectral bands, identified as amide I and amide III, inform about the collagen peptide backbone (Barth and Zscherp, 2002). Amide I (1560 and 1725 cm\(^{-1}\)) is due mostly to carbonyl stretching. Amide III (1210 and 1350 cm\(^{-1}\)) describes two vibrational modes: stretching between carbon and nitrogen atoms and bending of a secondary amine.

Bending of methylene groups (\(\delta\text{[CH}_2\text{]}\)) and stretching between carbon and hydrogen atoms (\(\nu\text{[C-H]}\)) are vibrational modes common to several peptides and manifest as peaks at 1365–1500 cm\(^{-1}\) and 2820–3035 cm\(^{-1}\). Benzene ring-breathing is specific to phenylalanine (\(\nu\text{[Phe]}\)) and occurs at \(\sim 1003\) cm\(^{-1}\) (Morris and Mandair, 2011). Bioapatite anions also contribute bands to a Raman spectrum of bone (Fig. 1A). The most prominent is a sharp peak attributed to symmetric stretching of phosphate (\(\nu\text{1-}[\text{PO}_4^{3-}]\)). Four smaller bands describe stretching and bending modes of phosphate (\(\nu_2, \nu_3, \nu_4\), and carbonate (\(\nu_1-\text{[CO}_3^{2-}]\)) (Awonusi et al., 2007; de Aza et al., 1997; Kravitz et al., 1968). Because organic

Fig. 1. Raman spectrum from specimen 558, a historic human metatarsal and an example of a bone with ‘well-preserved’ collagen. A) Peaks in a Raman spectrum of bone identify anions in bioapatite and functional groups in collagen. B) Fluorescence is a source of spectral noise and can be mitigated with baseline correction. Peak intensities used for bivariate ratio calculations are indicated with double-headed arrows. Hatched and shaded regions show peak areas used for bivariate ratio calculations (further detail in the Electronic Supplementary Material).
matter tends to degrade more rapidly than bioapatite in most burial conditions, we anticipated that bands related to structural features of collagen (i.e., amide I and amide III) would be particularly informative about collagen preservation.

The aim of this study was to develop a simple protocol that accurately identifies well-preserved collagen using FT-Raman spectroscopy. Modern and historic bone specimens for which the collagen quality is known were analyzed, and the spectra evaluated visually and quantitatively to provide a distinct numeric value characteristic of well-preserved specimens.

2. Materials and methods

2.1. Bone specimens

We studied 44 modern and historic (19th century) bone specimens from humans and other mammals (Table 1). Thirty-nine historic specimens were excavated from cemeteries and battlefields across disparate locations in the United States (n = 36) and Africa (n = 3) representing a range of taphonomic environments. Five modern specimens were not exposed to burial conditions and served as controlled examples of well-preserved collagen. A solid piece was removed from each bone for elemental analysis. For FT-Raman analysis, samples were prepared to expose the outer surface and a fresh cross-section using a chisel or bone saw. Clinging sediment was removed mechanically; no chemical treatments were applied. Because bone pieces for elemental analysis were removed prior to Raman analysis, the latter could not be performed in the exact location from which collagen was extracted and was performed in an area immediately adjacent.

Collagen quality parameters for each bone specimen were analyzed, including gravimetric measurements of extracted collagen yield, and elemental abundance mass spectrometric

### Table 1

<table>
<thead>
<tr>
<th>Sample</th>
<th>Museum designation</th>
<th>Description</th>
<th>Burial location</th>
<th>Bone % Collagen</th>
<th>Wt% N</th>
<th>Wt% C</th>
<th>C:N Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modern</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BB</td>
<td>cow</td>
<td>Unknown</td>
<td>Femur</td>
<td>25.9</td>
<td>14.7</td>
<td>36.8</td>
<td>2.9 Good</td>
</tr>
<tr>
<td>DDB</td>
<td>dog</td>
<td>Unknown</td>
<td>Femur</td>
<td>10.5</td>
<td>15.0</td>
<td>43.5</td>
<td>3.4 Good</td>
</tr>
<tr>
<td>DBII</td>
<td>deer</td>
<td>Virginia, U.S.</td>
<td>Mandible</td>
<td>31.0</td>
<td>10.8</td>
<td>27.0</td>
<td>2.9 Good</td>
</tr>
<tr>
<td>WR</td>
<td>whale</td>
<td>Unknown</td>
<td>Rib</td>
<td>13.6</td>
<td>13.5</td>
<td>40.6</td>
<td>3.5 Good</td>
</tr>
<tr>
<td>SB</td>
<td>cow</td>
<td>Unknown</td>
<td>Femur</td>
<td>3.2</td>
<td>15.5</td>
<td>44.0</td>
<td>3.3 Good</td>
</tr>
<tr>
<td>Historic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>281</td>
<td>29LA1091-BOR-42</td>
<td>Human</td>
<td>Metatarsal</td>
<td>4.9</td>
<td>14.2</td>
<td>40.4</td>
<td>3.3 Good</td>
</tr>
<tr>
<td>343</td>
<td>51RICHARDS-CC-07</td>
<td>Human</td>
<td>Metacarpal</td>
<td>12.1</td>
<td>14.8</td>
<td>41.8</td>
<td>3.3 Good</td>
</tr>
<tr>
<td>345</td>
<td>GLO-099-2A</td>
<td>Human</td>
<td>Femur</td>
<td>13.3</td>
<td>14.2</td>
<td>39.7</td>
<td>3.3 Good</td>
</tr>
<tr>
<td>350</td>
<td>GLO-099-2C</td>
<td>Human</td>
<td>Femur</td>
<td>6.3</td>
<td>14.2</td>
<td>40.0</td>
<td>3.3 Good</td>
</tr>
<tr>
<td>394</td>
<td>6CT58-S-AMMO3</td>
<td>Human</td>
<td>Femur</td>
<td>21.2</td>
<td>15.0</td>
<td>41.7</td>
<td>3.2 Good</td>
</tr>
<tr>
<td>396</td>
<td>6CT58-S-AMMO5</td>
<td>Human</td>
<td>Femur</td>
<td>17.6</td>
<td>14.6</td>
<td>40.3</td>
<td>3.2 Good</td>
</tr>
<tr>
<td>400</td>
<td>6CT58-S-AMM11</td>
<td>Human</td>
<td>Femur</td>
<td>19.0</td>
<td>14.5</td>
<td>40.2</td>
<td>3.2 Good</td>
</tr>
<tr>
<td>417</td>
<td>51KEYWORTH-CC-04</td>
<td>Human</td>
<td>Metacarpal</td>
<td>14.9</td>
<td>14.7</td>
<td>41.0</td>
<td>3.3 Good</td>
</tr>
<tr>
<td>420</td>
<td>51KEYWORTH-CC-07</td>
<td>Human</td>
<td>Metatarsal</td>
<td>21.5</td>
<td>14.3</td>
<td>40.4</td>
<td>3.3 Good</td>
</tr>
<tr>
<td>426</td>
<td>51WHITE-CC-02</td>
<td>Human</td>
<td>Fibula</td>
<td>16.4</td>
<td>14.5</td>
<td>40.8</td>
<td>3.3 Good</td>
</tr>
<tr>
<td>442</td>
<td>51CAUSTEN-CC-11</td>
<td>Human</td>
<td>Metacarpal</td>
<td>23.5</td>
<td>14.2</td>
<td>40.4</td>
<td>3.3 Good</td>
</tr>
<tr>
<td>443</td>
<td>51CAUSTEN-CC-13</td>
<td>Human</td>
<td>Metacarpal</td>
<td>27.5</td>
<td>14.9</td>
<td>41.8</td>
<td>3.3 Good</td>
</tr>
<tr>
<td>454</td>
<td>18FPP22A-03-300</td>
<td>Human</td>
<td>Femur</td>
<td>9.7</td>
<td>14.3</td>
<td>40.3</td>
<td>3.3 Good</td>
</tr>
<tr>
<td>455</td>
<td>18FPP22A-99-400</td>
<td>Human</td>
<td>Tibia</td>
<td>7.3</td>
<td>12.4</td>
<td>35.3</td>
<td>3.3 Good</td>
</tr>
<tr>
<td>466</td>
<td>51KEYWORTH-CC-08</td>
<td>Human</td>
<td>Tibia</td>
<td>16.8</td>
<td>12.3</td>
<td>34.9</td>
<td>3.3 Good</td>
</tr>
<tr>
<td>471</td>
<td>31FOSICUS-USA-037</td>
<td>Human</td>
<td>Tibia</td>
<td>7.2</td>
<td>14.1</td>
<td>40.7</td>
<td>3.4 Good</td>
</tr>
<tr>
<td>508</td>
<td>ELMINA-A-L49B</td>
<td>Human</td>
<td>Elmina Settlement, Ghana, Africa</td>
<td>5.0</td>
<td>12.1</td>
<td>34.6</td>
<td>3.3 Good</td>
</tr>
<tr>
<td>525</td>
<td>ELMINA-A-S51L</td>
<td>Human</td>
<td>Elmina Settlement, Ghana, Africa</td>
<td>8.0</td>
<td>10.2</td>
<td>28.8</td>
<td>3.3 Good</td>
</tr>
<tr>
<td>556</td>
<td>7NC8EA-WOODVILLE-04</td>
<td>Human</td>
<td>Fibula</td>
<td>5.1</td>
<td>12.3</td>
<td>36.5</td>
<td>3.5 Good</td>
</tr>
<tr>
<td>558</td>
<td>7NC8EA-WOODVILLE-10</td>
<td>Human</td>
<td>Metatarsal</td>
<td>16.1</td>
<td>13.6</td>
<td>38.5</td>
<td>3.3 Good</td>
</tr>
<tr>
<td>560</td>
<td>7NC8EA-WOODVILLE-SLOPEB</td>
<td>Human</td>
<td>Metatarsal</td>
<td>29.7</td>
<td>13.8</td>
<td>39.1</td>
<td>3.3 Good</td>
</tr>
<tr>
<td>566</td>
<td>FABC-08-167A</td>
<td>Human</td>
<td>Metatarsal</td>
<td>15.3</td>
<td>14.5</td>
<td>40.8</td>
<td>3.3 Good</td>
</tr>
<tr>
<td>589</td>
<td>44PWHINSCHELO-0114C</td>
<td>Human</td>
<td>Metatarsal</td>
<td>3.6</td>
<td>12.9</td>
<td>38.8</td>
<td>3.5 Good</td>
</tr>
<tr>
<td>255</td>
<td>GETTYS-NPS-965</td>
<td>Human</td>
<td>Temporal</td>
<td>0.6</td>
<td>8.2</td>
<td>27.9</td>
<td>4.0 Bad</td>
</tr>
<tr>
<td>266</td>
<td>29LA1091-BOR-23B</td>
<td>Human</td>
<td>Temporal</td>
<td>1.0</td>
<td>6.5</td>
<td>34.0</td>
<td>6.1 Bad</td>
</tr>
<tr>
<td>267</td>
<td>29LA1091-BOR-23C</td>
<td>Human</td>
<td>Temporal</td>
<td>0.5</td>
<td>4.3</td>
<td>25.9</td>
<td>7.0 Bad</td>
</tr>
<tr>
<td>268</td>
<td>29LA1091-BOR-23D</td>
<td>Human</td>
<td>Temporal</td>
<td>0.2</td>
<td>5.3</td>
<td>33.9</td>
<td>7.5 Bad</td>
</tr>
<tr>
<td>297</td>
<td>TRINITY-EAST-06</td>
<td>Human</td>
<td>Temporal</td>
<td>0.4</td>
<td>1.9</td>
<td>8.3</td>
<td>5.0 Bad</td>
</tr>
<tr>
<td>303</td>
<td>TRINITY-EAST-14</td>
<td>Human</td>
<td>Temporal</td>
<td>-0.2</td>
<td>7.7</td>
<td>35.2</td>
<td>5.4 Bad</td>
</tr>
<tr>
<td>305</td>
<td>TRINITY-EAST-22</td>
<td>Human</td>
<td>Temporal</td>
<td>0.8</td>
<td>9.7</td>
<td>34.8</td>
<td>4.2 Bad</td>
</tr>
<tr>
<td>307</td>
<td>TRINITY-EAST-25</td>
<td>Human</td>
<td>Temporal</td>
<td>0.5</td>
<td>7.1</td>
<td>25.9</td>
<td>4.3 Bad</td>
</tr>
<tr>
<td>309</td>
<td>TRINITY-EAST-04C</td>
<td>Human</td>
<td>Metatarsal</td>
<td>-0.2</td>
<td>5.2</td>
<td>47.2</td>
<td>10.5 Bad</td>
</tr>
<tr>
<td>329</td>
<td>7BAVVISTA-DHCA-116X</td>
<td>Human</td>
<td>Metatarsal</td>
<td>-0.2</td>
<td>6.3</td>
<td>26.6</td>
<td>5.0 Bad</td>
</tr>
<tr>
<td>330</td>
<td>7BAVVISTA-DHCA-118A</td>
<td>Human</td>
<td>Metatarsal</td>
<td>0.5</td>
<td>7.0</td>
<td>22.0</td>
<td>3.7 Bad</td>
</tr>
<tr>
<td>401</td>
<td>6CT58-S-AMM13</td>
<td>Human</td>
<td>Metatarsal</td>
<td>0.5</td>
<td>12.9</td>
<td>23.3</td>
<td>2.1 Bad</td>
</tr>
<tr>
<td>501</td>
<td>ELMINA-A-E51B</td>
<td>Human</td>
<td>Metatarsal</td>
<td>4.4</td>
<td>10.0</td>
<td>32.5</td>
<td>3.8 Bad</td>
</tr>
<tr>
<td>555</td>
<td>7NC8EA-WOODVILLE-02</td>
<td>Human</td>
<td>Metatarsal</td>
<td>1.6</td>
<td>6.4</td>
<td>27.8</td>
<td>5.0 Bad</td>
</tr>
<tr>
<td>557</td>
<td>7NC8EA-WOODVILLE-06</td>
<td>Human</td>
<td>Metatarsal</td>
<td>6.0</td>
<td>3.5</td>
<td>32.7</td>
<td>5.9 Bad</td>
</tr>
<tr>
<td>571</td>
<td>18ST1-103-2420W</td>
<td>Human</td>
<td>Metatarsal</td>
<td>0.9</td>
<td>7.5</td>
<td>27.7</td>
<td>4.3 Bad</td>
</tr>
</tbody>
</table>

### Notes:

- **Elemental abundance data** from most of the historic specimens has been published previously in France and Owsley (in press) and France et al. (2014) and reappears here.
- **Historic specimens** are accessioned objects in the collection of the Smithsonian National Museum of Natural History.
- **Conditions of good** and **bad** refer to well-preserved and poorly-preserved collagen, respectively.
- **Yields** are noted as 0% when the limit of resolution of our scale and are noted as <0.2% to encompass the inherent error.
assessments of C:N ratio, percentage nitrogen, and percentage carbon. Collagen was extracted from solid bone pieces (~500 mg) by modified methods of Longin (1971). Solid samples were decalcified in 0.5 M HCl at 4 °C for several days and rinsed to neutrality. Humic and fulvic acidic contaminants were removed by soaking in 0.125 M NaOH overnight. Samples were rinsed and heated in 0.03 M HCl at 95 °C overnight to denature and separate the collagen strands. The resulting supernatant was freeze-dried and, in well-preserved specimens, the organic extract contained denatured collagen. The collagen yield was recorded as a percentage of the original bone weight. A portion of each organic extract was weighed (~0.5 mg) and packed into tin cups. Samples were combusted on a Costech 4010 Elemental Analyzer (EA) coupled to a Thermo Delta V Advantage mass spectrometer via a Conflo IV interface. Weight percent nitrogen and weight percent carbon yields were calculated by calibration of peak area with a homogeneous acetanilide standard (associated error is ±0.5%). Weight percent yield values were converted to moles and used to calculate atomic C:N ratios. Bones were assigned a designation of ‘well-preserved’ or ‘poorly-preserved’ based on the previously well-established elemental abundance and C:N ratio parameters indicating collagen quality as outlined in Section 1.1.

2.3. Collection and processing of Raman spectra

One Raman spectrum was collected from the outer surface, and one from a freshly cut cross-section, for each specimen. An additional two spectra were collected from the cross sections of five specimens with relatively small proportions of collagen: 266, 309, 401, 455, 508. These additional spectra were collected from different places and were used to assess compositional heterogeneity across the bone surface. A further two spectra (one from the outer surface and one from the cross-section surface) were collected from specimens 266, 281, 303, 309, 329, 471 and 508 at a lower spectral resolution.

Peak height ratio = Peak height at wavenumber A ÷ Peak height at wavenumber B  
Peak area ratio = Peak area at wavenumber A ÷ Peak area at wavenumber B

Raman spectra were collected with an NXR FT-Raman module coupled to a 6700 Fourier transform infrared spectrometer (Thermo Electron Corporation, Madison, WI, USA). The Raman module was equipped with a continuous wave near infrared Nd:YVO4 excitation laser (1064 nm), CaF2 beam splitter, and thermoelectrically-cooled InGaAs detector. Spectra were collected using a 50 μm laser spot and 1.0 OD neutral density filter that limited laser power to a maximum estimated 0.13 W. Laser power was chosen empirically to maximize signal-to-noise (SNR) ratio without burning the sample. Initially, spectra were a co-addition of 1024 scans across 100–3701 cm⁻¹ (4 cm⁻¹ resolution). As the study progressed it became clear that spectral noise potentially was confounding correct assignment of some specimens, particularly those with low collagen concentrations (i.e., the collagen peaks were small). To mitigate this problem, some spectra were collected again from a different spot free of accretions or surface staining, and/or with more co-added scans (i.e., 2048 or 4096 scans). Spectra for a subset of 7 samples (listed above) were collected at 8 cm⁻¹. Noise reduction through post-processing was explored using three or nine point Savitzky Golay filtering performed with the ‘sgolayfit’ function from the signal 0.7–2 package (Signal developers, 2011) in R 2.15.2 (R Core Team, 2012).

2.4. Data analysis

Raman spectra were evaluated visually and quantitatively to determine which approach and combination of spectral features produced the best indicator of collagen quality. Spectra collected from outer surfaces and freshly cut cross-sections were considered as two separate datasets for all subsequent analyses.

2.4.1. Visual features of Raman spectra of bone collagen

Spectral characteristics of visual interest were peaks typical of bone, the shape of those peaks, fluorescence, spectral noise, and spectral artefacts caused by heating of the sample by the laser. If a spectrum appeared noisy or the baseline was distorted by sample heating, a new spectrum was collected with adjusted parameters or at a new spot to determine the ideal parameters for collection.

2.4.2. Bivariate data analysis: peak ratios

A desired outcome of this research was an unambiguous numeric value to distinguish collagen quality. An ideal statistical indicator would show a range of values that corresponds to well-preserved collagen (as determined by the elemental abundance indicators) that is distinct from the range for poorly-preserved collagen, with no overlap. We calculated simple ratios for all possible pairwise combinations of Raman peaks in the bone spectrum:

The peak ratios grouped into three ranges: 1) ratios from well-preserved specimens only, 2) ratios from poorly-preserved specimens only, and 3) ratios from both well- and poorly-preserved specimens. Ratios were scored according to how well they segregated well-preserved from poorly-preserved specimens, where spectra in groups 1) and 2) were considered correctly assigned:

Classification rate = (# correctly assigned spectra ÷ total number of spectra) × 100

using a 50 μm laser spot and 1.0 OD neutral density filter that limited laser power to a maximum estimated 0.13 W. Laser power was chosen empirically to maximize signal-to-noise (SNR) ratio without burning the sample. Initially, spectra were a co-addition of 1024 scans across 100–3701 cm⁻¹ (4 cm⁻¹ resolution). As the study progressed it became clear that spectral noise potentially was confounding correct assignment of some specimens, particularly those with low collagen concentrations (i.e., the collagen peaks were small). To mitigate this problem, some spectra were collected again from a different spot free of accretions or surface staining, and/or with more co-added scans (i.e., 2048 or 4096 scans). Spectra for a subset of 7 samples (listed above) were collected at 8 cm⁻¹. Noise reduction through post-processing was explored using three or nine point Savitzky Golay filtering performed with the ‘sgolayfit’ function from the signal 0.7–2 package (Signal developers, 2011) in R 2.15.2 (R Core Team, 2012).

Raman spectra were evaluated visually and quantitatively to determine which approach and combination of spectral features produced the best indicator of collagen quality. Spectra collected from outer surfaces and freshly cut cross-sections were considered as two separate datasets for all subsequent analyses.

2.4.1. Visual features of Raman spectra of bone collagen

Spectral characteristics of visual interest were peaks typical of bone, the shape of those peaks, fluorescence, spectral noise, and spectral artefacts caused by heating of the sample by the laser. If a spectrum appeared noisy or the baseline was distorted by sample heating, a new spectrum was collected with adjusted parameters or at a new spot to determine the ideal parameters for collection.

2.4.2. Bivariate data analysis: peak ratios

A desired outcome of this research was an unambiguous numeric value to distinguish collagen quality. An ideal statistical indicator would show a range of values that corresponds to well-preserved collagen (as determined by the elemental abundance indicators) that is distinct from the range for poorly-preserved collagen, with no overlap. We calculated simple ratios for all possible pairwise combinations of Raman peaks in the bone spectrum:

The peak ratios grouped into three ranges: 1) ratios from well-preserved specimens only, 2) ratios from poorly-preserved specimens only, and 3) ratios from both well- and poorly-preserved specimens. Ratios were scored according to how well they segregated well-preserved from poorly-preserved specimens, where spectra in groups 1) and 2) were considered correctly assigned:
the ‘baseline.fillPeaks’ function from the baseline 1.01 package (Liland and Mevik, 2012) in R 2.15.2. For each of the seven samples collected at 8 cm\(^{-1}\) resolution, collagen quality was predicted by substituting a lower resolution spectrum for the corresponding 4 cm\(^{-1}\) resolution spectrum, and the classification rates were recalculated.

2.4.3. Multivariate data analysis

Raman spectra also were analyzed with multivariate methods, which can be useful when information is dispersed across many regions of a spectrum. The wavenumber range of each spectrum was reduced to 3100–360 cm\(^{-1}\) and divided into eight regions: 1) 3100–2800 cm\(^{-1}\), 2) 2800–1750 cm\(^{-1}\), 3) 1750–1550 cm\(^{-1}\), 4) 1550–1530 cm\(^{-1}\), 5) 1530–1150 cm\(^{-1}\), 6) 1150–910 cm\(^{-1}\), 7) 910–700 cm\(^{-1}\), and 8) 700–360 cm\(^{-1}\). Separate baseline corrections were made for each region using the ‘baseline.fillPeaks’ function from the baseline 1.01 package (Liland and Mevik, 2012) in R 2.15.2. A region of the spectrum not informative for collagen or bone mineral subsequently was removed (2800–1750 cm\(^{-1}\)). Each spectrum was minimum–maximum normalized and each dataset was mean-centered (Beebe et al., 1998; Thomas and Chinsamy, 2011). Spectral data were explored with principal components analysis (PCA), calculated using the ‘princomp’ function in R 2.15.12 (Fig. 2).

The outer surface and cross sectioned datasets were studied separately with a single component partial least squares discriminant analysis (PLSDA). The spectra in each dataset were subdivided into well-preserved and poorly-preserved groups according to their previously established elemental abundance collagen quality indicators. One spectrum was removed from the dataset, and the collagen quality for that spectrum was predicted according to the spectra remaining in each group. A Raman spectrum of a well-preserved specimen predicted to belong in the well-preserved group was considered correctly classified, whereas a well-preserved spectrum predicted to be part of the poorly-preserved group was misclassified (and vice versa for poorly-preserved collagen specimens). Predictions were calculated for each spectrum (i.e., full cross validation). PLSDA was performed with the ‘plsda’ function from the caret package (Kuhn, 2012) in R 2.15.2.

3. Results

3.1. Visual features of Raman spectra of bone collagen

Qualitative assessment of Raman spectra indicates several distinct peaks that visually distinguish collagen quality (Fig. 3). For poorly-preserved bones, the organic peaks show a general trend of decreasing size and shape distortion related to decreasing collagen quality in the outer surface dataset (Fig. 3A). The organic peaks of the cross sectioned dataset show a general trend of decreasing size and shape distortion related to decreasing collagen quality in the cross sectioned dataset (Fig. 3B).
yield. The amide peaks disappear first, followed by shrinkage and often distortion and/or disappearance of the C–H peaks. However, some samples with poorly preserved collagen gave spectra with nicely defined C–H peaks at 2820-3035 cm⁻¹. Spectral noise also contributed to decreased peak quality, especially by camouflaging smaller peaks which might otherwise be diagnostic. The collection of good quality spectra required subjective adjustments including choosing clean, lighter colored surfaces for testing, checking multiple spots on the same bone, increasing the number of co-added scans, and decreasing the spectral resolution to 8 cm⁻¹ in some cases. Spectra from outer surfaces often differed from those of cross-sections.

3.2. Statistical data analysis

3.2.1. Outer surface and cross-section comparison

The degree to which Raman spectra of outer bone surfaces differed from cross-sectioned surfaces of the same bone was quantified for each peak ratio. Differences in the peak height ratios from outer surface spectra (Ou) and cross-section spectra (Cs) were recovered as a percentage of the maximum ratio for each pair:

\[
\text{Variation between Ou and Cs} = \left| \frac{\text{abs}(\text{Ou} - \text{Cs})}{\text{max}(\text{Ou}, \text{Cs})} \right| \times 100
\]

The average discrepancy between cross-section and outer surface spectra was 31.7% for mineral to collagen peak ratios (n = 39, \(\sigma = 22.1\%\)). Hence, the outer surfaces of the bones that had been exposed to the burial environment were spectrally different from the freshly cut internal surfaces. Heterogeneity also was observed among spectra collected at different spots along the same cross-sectioned surface. The mineral to collagen ratio showed an average discrepancy of 38.5% (n = 5, \(\sigma = 24.5\%\)) within the same bone.

For samples in which multiple cross-section spectra were collected, a single representative spectrum was chosen for the quality predictions and parameter regressions that follow. These spectra were selected according to two guidelines: 1) the spectrum collected nearest to the collagen extraction sampling site was chosen, or if the extraction sampling site was not apparent, 2) the spectrum with the best SNR was chosen. The full set of spectra is available as electronic supplementary material.

3.2.2. Collagen quality predictions: bivariate analysis

Raman spectra from cross-sectioned surfaces that were collected and analyzed with the same parameters (1024 co-added scans, 4 cm⁻¹ resolution, no spectral smoothing) successfully classified between 38 and 100% of samples using the 960:1636 cm⁻¹ peak height ratio (Fig. 4). The classification rate depended on which spectra were selected to represent specimens 266, 309, 401, 455 and 508. Hence, the classification rate was influenced by surface heterogeneity. The highest rate was achieved when spectra had the highest SNR and were collected closest to the collagen sampling site. This ratio cleanly separated the well- and poorly-preserved samples (Fig. 5A). Cross-section spectra from well-preserved samples featured a 960 cm⁻¹:1636 cm⁻¹ ratio between 9.9 and 19.4 (mean = 14.0, 1σ = 2.6). Cross-section spectra of poorly-preserved specimens featured a 960 cm⁻¹:1636 cm⁻¹ ratio between 20.4 and 156.6 (mean = 59.3, 1σ = 34.9).

The classification rate for outer-surface spectra using the 960:1636 cm⁻¹ peak height ratio was only 16% because the range of well- and poorly-preserved samples overlaps for all but the highest values (Fig. 5B). The 2978:2882 cm⁻¹ ratio was most selective for outer surfaces with 57% correct classifications.

Bivariate analyses based on peak area produced a high percentage of correct classifications. However, ratios most indicative of preservation quality differed for the cross-section and outer surface datasets, and post-processing smoothing, or lack thereof, affected the outcome. Therefore we could not determine a single unambiguous numeric value indicative of collagen quality for both outer surfaces and cross-sections, which rendered this approach less informative. The highest classification rate was 74% for the ratio of the 485–360 cm⁻¹ peak area to the 1500–1365 cm⁻¹ peak area, and used unsmoothed 4 cm⁻¹ spectra collected from cross sections. The classification rate rose to 100% when the spectral resolution of the cross-section spectrum from specimen 309 was reduced to 8 cm⁻¹. For outer surface spectra, the highest classification rate using peak area ratios was 35%. This classification rate was from the ratio of the 635–510 cm⁻¹ peak area to the 1500–1365 cm⁻¹ peak area and the spectra had been nine-point smoothed. However, in most cases post-processing smoothing did not reliably improve, and often worsened, our predictions overall.

3.2.3. Collagen quality predictions: multivariate analysis

Twenty one of 22 well-preserved cross-section spectra and 16 of 17 poorly-preserved cross-section spectra were classified correctly with PLSDA, resulting in a sensitivity of 95%. Two misclassified spectra, specimens 309 and 455, have poorly-preserved and well-preserved collagen, respectively. A principal component (PC) analysis positively loaded collagen bands on the first PC. PC1 score values for specimen 309 clustered among samples with well-
preserved collagen, and the PC1 score value for specimen 455 was intermediate between the cluster of well-preserved and poorly-preserved specimens (Fig. 2). Spectra collected from outer surfaces were less informative about collagen quality than the freshly cut cross-sections (Fig. 5B). Thirteen external surface spectra were misclassified by PLSDA, resulting in a sensitivity of 68%. While PLSDA identified well-preserved specimens quite successfully, we note that this approach requires significant time and effort to establish an internal laboratory dataset for comparison.

4. Discussion

The preservation of collagen in exhumed bone can be evaluated with FT-Raman spectroscopy. Because bone is a composite material of mineral and protein, substances that respond differently to aging and burial conditions, deviations from known proportions are good indicators of preservation. All of our statistical classification methods show promise and improve selectivity over blind selection. The most unambiguous and accurate measure of preservation was achieved through analysis of peak height ratios. Analysis of peak area ratios did not produce a single unambiguous measure of preservation, and PLSDA requires a large internal dataset which is restrictive for smaller laboratories with lower throughput.

The bivariate analysis of peak height ratios showed the highest rates of correct classification using the 960 cm\(^{-1}\) and 1636 cm\(^{-1}\) ratio. A Raman spectrum from a bone with well-preserved collagen will have a 960 cm\(^{-1}\):1636 cm\(^{-1}\) ratio of \(\pm 19.4\) after baseline correction (value reflects an approximate 95% confidence interval). The 960 cm\(^{-1}\) band is attributed to phosphate stretching and is the most distinct mineral band in a spectrum of bone; the 1636 cm\(^{-1}\) band is attributed to amide I stretching and provides insight into the peptide backbone of collagen (Barth and Zscherp, 2002; Kravitz et al., 1968). Hence, the peak height ratio of 960 cm\(^{-1}\):1636 cm\(^{-1}\) informs about collagen fragmentation, an important index of collagen degradation (e.g., Tuross et al., 1988). Despite the C–H stretching region around 2880 cm\(^{-1}\) being the most intense collagen region in a Raman spectrum of bone, bands in this region are not accurate predictors of collagen quality. Indeed, aliphatic C–H functionality occurs in both degraded and intact collagen (Frushour and Koenig, 1975).

Bivariate analysis of mineral and collagen peak heights also predicts the amount of intact collagen in bone. Comparison of the 960 cm\(^{-1}\):1636 cm\(^{-1}\) ratios to the percent collagen extracted from the same bones indicates that decreased collagen quality co-occurs with collagen loss in both outer surfaces and fresh cut cross-sections (Fig. 5). This reiterates the information provided by the ratios because a relatively small 1636 cm\(^{-1}\) peak indicates that less amide I, a defining component of collagen, is present. Given the variability in collagen content on outer bone surfaces, a ratio in the higher range of 19.4–26.8 still offers some chance of finding intact collagen, but the % collagen yield likely will be low (<5%). Further study of bones in this range might benefit from a larger sample set.

The mineral to collagen ratio of a Raman spectrum varied across a bone. Outer bone surfaces tended to have high ratios, indicating...
that collagen has been lost from surfaces directly exposed to burial environments. Well-preserved bones tended to have lower mineral to collagen ratios on freshly cross-sectioned surfaces. Importantly, we observed spectral heterogeneity on cross-sectioned surfaces which indicates heterogeneous preservation within a single bone. This has been observed previously in microscale imaging and spectroscopy of bone composition (Chadefaux et al., 2009; Lebon et al., 2011; Turner-Walker and Syversen, 2002), although our Raman technique identifies such patterns non-destructively. The variation in mineral to collagen ratios, and hence the variation in collagen abundance or preservation, encourages collagen sampling and Raman spectral measurement from the same location on the bone. Collagen would ideally be extracted from the same location that a 960 cm\(^{-1}\):1636 cm\(^{-1}\) peak height ratio of $<19.4$ has been measured. We note that our analyses were conducted on historic bone. A majority of the samples produced good quality spectra using the co-addition of 1024 scans across 100–3700 cm\(^{-1}\). More co-added scans may be needed (i.e., 2048 or 4096) when the amount of intact collagen is low and/or the bone has a darker color. All samples in this study were analyzed initially at 4 cm\(^{-1}\) resolution; collecting spectra at lower resolution (i.e., 8 cm\(^{-1}\)) may reduce the influence of spectral noise on peak heights and improve selectivity for collagen preservation. If the baseline-corrected 960 cm\(^{-1}\):1636 cm\(^{-1}\) peak height ratio is $\geq 26.8$, the bone is unlikely to yield intact collagen. For example, a bone in our study set with an outer surface 960 cm\(^{-1}\):1636 cm\(^{-1}\) peak height ratio $<26.8$ had a 61% chance of bearing well-preserved collagen; an outer surface ratio of $<19.4$ increased that likelihood to 67%. These probabilities should allow investigators to quickly isolate an adequate sample set for further analyses such as radiocarbon dating, stable isotope analyses, or proteomics. If more samples are needed and one is willing to devote more analytical time to each sample, additional spectra can be collected from outer surfaces to identify cases where a bone might have heterogeneous preservation across the surface. Alternatively, collection of additional spectra from fresh cross-sections with one baseline-corrected 960 cm\(^{-1}\):1636 cm\(^{-1}\) peak height ratio $<19.4$ increases the likelihood of extracting well-preserved collagen to 95%.

5. Conclusions

This study examined FT-Raman spectroscopy as a means to screen historic bone samples for the presence of well-preserved collagen. A majority of the samples produced good quality spectra using the co-addition of 1024 scans across 100–3700 cm\(^{-1}\) (4 cm\(^{-1}\) resolution). Although a few samples required variations on these settings to produce quality spectra, we suggest that these parameters offer a quick screening method with a high probability of accurate results. It should be noted that Raman as described here is a spot analysis and a spectrum from one location may not be representative of the entire bone specimen. Variation in collagen quality was observed across a bone with the outer surfaces typically showing greater alteration than internal surfaces. Future work will examine older archaeological and paleontological specimens to determine if this pattern of heterogeneity changes with more exposure to diagenetic influences.

Despite this heterogeneity, collagen quality still was predictable with high accuracy. A bivariate analysis of the 960 cm\(^{-1}\):1636 cm\(^{-1}\) (mineral to collagen) peak height ratio in freshly cut cross-sections produced the most unambiguous and accurate index of collagen quality. Samples with a mineral to collagen ratio of $<19.4$ are considered well-preserved. Samples with a mineral to collagen ratio of $\geq 19.4$ show a systematic decrease in peak height intensities associated with degradation of the collagen peptide backbone while the intensities of peaks associated with the mineral decrease to a lesser degree. This suggests a preferential breakdown of organic material in bone during diagenesis which has implications for the duration of burial beyond which a bone will no longer yield reliable data from its collagen and other organic components.

Collagen preservation was predicted better from cross section surfaces than from outer surfaces. Alteration of outer surfaces most often resulted in well-preserved samples being misclassified as...
poorly-preserved. This would result in several usable specimens being eliminated from further destructive analyses, but would still allow one to select a subset of well-preserved specimens that is highly likely to produce quality results in further testing. Laboratories can now use this pre-screening method before outsourcing analyses and spending additional time or money on radiocarbon dating, stable isotope analyses, or protein sequencing.

Acknowledgments

The authors acknowledge K. Bruwelheide, S. McGuire, S. Mills, W. Miller, D. Owsley, and A. Warmack for assistance with sample procurement and preparation. Financial support for D. Thomas provided by Smithsonian Peter Buck Post-doctoral Fellowship. Financial support for C. Doney provided by National Science Foundation Research Experience for Undergraduates Award #SMA-1156360. All analyses performed at the Smithsonian OSUSS/SMI Stable Isotope Mass Spectrometry Laboratory and the Smithsonian MCI Modern Materials Laboratory at the Smithsonian Museum Conservation Institute, Suitland, MD.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jas.2013.11.020

References