

Diagnosis of degenerative lesions of supraspinatus rotator cuff tendons by Fourier transform-Raman spectroscopy

Sergio Camargo Godoy Penteado
Bianca Palma Fogazza
Carolina da Silva Carvalho
Emilia Ângela Loschiavo Arisawa
Mario Augusto Martins
Airton Abrahão Martin

Universidade do Vale do Paraíba
Instituto de Pesquisa & Desenvolvimento
Avenida Shishima Hifumi, 2911
12244-000, São José dos Campos
São Paulo, Brazil

Herculano da Silva Martinho

Centro de Ciências Naturais e Humanas
Universidade Federal do ABC
Rua Catequese 242
08090-400 Santo André
São Paulo, Brazil
E-mail: herculano.martinho@ufabc.edu.br

Abstract. The Fourier transform (FT)-Raman spectroscopy technique is used to assess the biochemical alterations that occur in the degenerative process of the rotator cuff supraspinatus tendon. The main alterations observed occur in the glycine, proline, hydroxyproline, cysteine, cistine, phenylalanine, tyrosine, collagen I and III, nucleic acid, lipids, glycosaminoglycans, and metalloproteinases bands. An increasing intensity for these bands is found in degenerated tendons, a finding well correlated with hyaline state and cellular activity. Statistical analysis (principal components analysis and clustering) shows a clear separation of the spectra into nonhyalinized and hyalinized clusters, which enables the construction of a binary diagnosis model based on logistic regression. Best diagnosis provided a sensitivity of 66.0% and a specificity of 74.7% with 79.6% concordant pairs. The discriminating power of the diagnostic test is assessed by computing the area under the receiving-operator characteristic curve (AUC), which indicates good accuracy (AUC=0.81). In principle, these results indicate that Raman spectroscopy can be used as an auxiliary aid to improve shoulder tendon surgery quality by guiding anchoring onto more healthy (nonhyaline) pieces of tendons. This should contribute to a decrease in the current high rerupture rate (13 to 68%) for this procedure.

© 2008 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.2841017]

Keywords: optical biopsy; Raman spectroscopy; rotator cuff; collagen; tendon degeneration; tendonitis; hyalinization.

Paper 07124RR received Apr. 2, 2007; revised manuscript received Sep. 19, 2007; accepted for publication Oct. 11, 2007; published online Feb. 28, 2008.

1 Introduction

Shoulder pain is a common problem in Western society.¹ The prevalence of shoulder pain in the general population may be as high as 6 to 11% under the age of 50 yr, increasing to 16 to 25% in elderly people.² The loss of productivity and the inability to work and carry out household activities can be a considerable burden to both the patient and society.²⁻⁴ Data from social insurance services of many countries indicate that nearly 20% of all job absences and compensations are due to some kind of shoulder pain.⁵ Rotator cuff tendonitis is the most frequently recorded clinical entity, corresponding to nearly 30% of reported cases.^{1,3-6}

The rotator cuff is the name for a group of four muscles—supraspinatus, infraspinatus, teres minor, and subscapularis—that in conjunction with their tendons act as stabilizers for the glenohumeral joint during shoulder movement.^{7,8} Rotator cuff pathologies can vary from tendonitis to complete tear. The supraspinatus tendon, located on the upper part of the cuff, is the most affected by these tendinopathies.⁸

Despite the magnitude and economic impact of rotator cuff injuries, little is known regarding their etiology. Certain factors that can cause cuff lesions are well known,⁹⁻¹² for example, tendon overload, intrinsic degeneration, age, high body mass index, previous shoulder trauma, reduced shoulder strength, cuff impingement, and degenerative changes. However, despite several cellular, immunohistochemical, and histological studies, the origins of these degenerations are poorly understood.^{8,13} From a biochemical viewpoint, it is known that healthy tendons are mainly composed of type I collagen^{8,14} (65 to 80%). Type III collagen is also found but in small amounts (<3%), as well as other components such as water, elastic fibers, proteoglycans, glycoproteins, and metalloproteinases.^{13,14}

In many cases, a conservative treatment regimen (drugs, rest, massage, etc.) is recommended as an initial treatment strategy.¹⁵ However, chronically painful tendon conditions are difficult to treat, and in cases where complete tendon/muscle tear occurs or when failure of other treatments occurs after 6 weeks, surgery is recommended. Currently, a minimally invasive surgical approach represented by miniopen shoulder videoarthroscopy is the most recommended approach for re-

Address all correspondence to Herculano Martinho, Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, Rua Catequese 242, 08090-400 Santo André, São Paulo, Brazil; E-mail: herculano.martinho@ufabc.edu.br

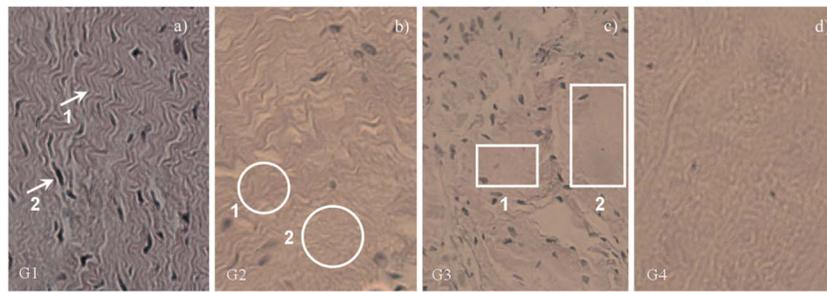


Fig. 1 Histological stains (63 \times objective) representing the four grades of rotator cuff tendon degeneration, as established by Riley et al.¹⁹ (a) Grade G1 (normal), where arrows 1 and 2 represent a single wavy collagen fiber and an elongated tenocyte, respectively; (b) grade G2 (moderate), where circles 1 and 2 represent regions presenting loss of collagen fiber waviness; (c) grade G3 (mild), where squares 1 and 2 represent areas of incipient hyalinization and the high nuclei density is clear; (d) grade G4 (severe), where the uniform hyalinization is clearly present in the stain.

covery of normal cuff state.^{16,17} In this procedure, the inner part of the shoulder joint is accessed via three small incisions. Through these incisions, the surgeon introduces the arthroscope, instrumentation, and a flux of physiological solution. Anchoring of ruptured regions is guided by visual inspection on a monitor.^{16,17} Rerupture rates after rotator cuff repair have been reported to range from 13 to 68%, depending on the size of the original tear.¹⁸ Currently it is impossible to predict the probability of rerupture of a repaired tendon. One possible cause of this high rate is the lack of real-time data concerning the degenerative state of the tendon that is visually chosen by the surgeon during the procedure.

Only in very special cases (e.g., suspicion of cancer or severe degeneration) is a small tendon sample removed for biopsy. One of the most used histological scales is that established by Riley et al.¹⁹ It uses four graduations to characterize the degenerative process of rotator cuff tendons. The characteristic stains of such grading are showed in Fig. 1 (see methodology section for more details). In fact, the quality of shoulder surgery would be substantially increased if the physician possessed an auxiliary aid to discriminate normal tendon tissue from degenerated pieces in real time and performed the anchoring using only normal or less-degenerated tendons. The goal of the work presented here was to determine whether Raman spectroscopy could be used as this auxiliary aid, enabling real-time diagnosis of the degenerative state of rotator cuff supraspinatus tendons. In a pilot study,²⁰ our group previously showed that Raman spectroscopy enabled the classification of normal and degenerative tendon tissues.

2 Methodology

2.1 Tissues

This research was carried out following the ethical principles established by the Brazilian Health Ministry and was approved by the local ethics in research committee (certificate H149/2006/CEP—UniVap). Patients were informed concerning the subject of the research and gave their permission for the collection of tissue samples.

Rotator cuff supraspinatus tendon tissue samples were obtained from 39 patients submitted to shoulder surgery at São Gabriel Clinic, São José dos Campos, São Paulo, Brazil. Immediately after the surgical procedure, the samples were identified, snap frozen and stored in liquid nitrogen (77 K) in cryogenic vials (Nalgene[®]) prior to Fourier transform (FT)-

Raman spectra recording. Each patient's tissue was sectioned into three parts, resulting in 117 samples. Raman spectra were measured at three different points on each sample, resulting in 351 spectra. Next, the samples were fixed in 10% formaldehyde solution for further histopathological analysis. Each measured sample was histopathologically assessed by two pathologists and classified as G1, G2, G3, or G4, according to the histopathological grading proposed by Riley et al.¹⁹

2.2 Reference Compounds

Reference compounds cysteine (Cys), cistine (Cis), proline (Pro), hydroxyproline (Hyp), and collagen types I and III were used to refine the vibrational band assignment. All compounds were of commercial molecular biology grade obtained from Sigma Aldrich (St. Louis, Missouri, USA).

2.3 FT-Raman Spectroscopy

An FT-Raman spectrometer (Bruker RFS 100/S; Bruker Optics GmbH, Ettlingen, Germany) was used with an Nd:YAG laser at 1064 nm as the excitation light source. The laser power at the sample was maintained at 110 mW, while the resolution was set to 4 cm^{-1} . The spectra were recorded using 300 scans (nearly 10 min of acquisition time). For FT-Raman data collection, the samples were brought to room temperature and maintained moistened in 0.9% physiological solution to preserve their structural characteristics, then they were placed in a windowless aluminum holder for Raman spectra collection. Observation revealed that the chemical species in the physiological solution (Ca^{2+} , Na^+ , K^+ , Cl^- , water) presented no measurable Raman signal and their presence did not affect the tissue spectral signal.

2.4 Histopathological Analysis

The stains used on the tendon samples were read by two pathologists and compared with Riley's et al. scale,¹⁹ which comprises four levels of degeneration. Grade 1 [G1, Fig. 1(a)] corresponds to normal tendons and is characterized by a wavy outline of collagen fibers, easily discernible individual fibers [arrow 1, Fig. 1(a)], tenocyte nuclei that are elongated and parallel to the collagen bundles [arrow 2, Fig. 1(a)]. Grade 2 [G2, Fig. 2(b)] corresponds to mild degeneration and it is characterized by relatively well aligned collagen fibers presenting irregular waviness [circles 3 and 4, Fig. 1(b)]. In this case, the individual fibers are not readily identified, tenocyte

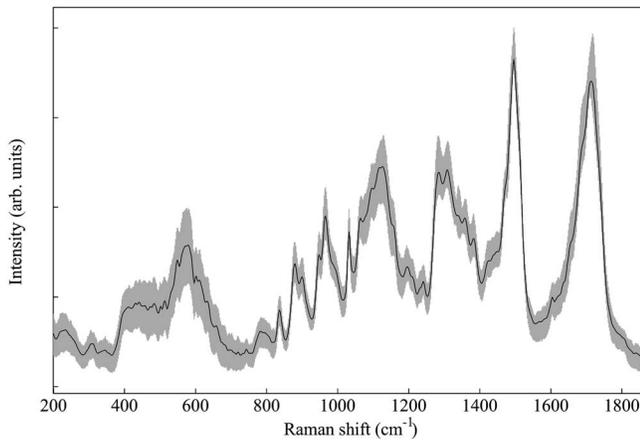


Fig. 2 Box plot representing all tendon spectral data. Each spectrum was baseline corrected and vector normalized. The solid line is the mean Raman spectra. The vertical gray lines indicate the region between the first and third quartiles.

cell nuclei are shorter, and the nuclei are displayed in Indian file. Grade 3 [G3, Fig. 1(c)] is moderate degeneration. The onset of the process of collagen hyalinization occurs in this grade. The hyaline state is characterized by the deposition of collagen peptide fragments in the extracellular matrix. Increasing quantities of hydroxylysine, hydroxyglysilypyridine, and lysylpyridioniline due to incomplete healing, result in greater collagen solubility, thus favoring hyalinization.¹³ When stained, this state is characterized by eosinophilic pink regions [squares 1 and 2, Fig. 1(c)]. Loss of orientation among the fibers and tenocyte nuclei and increases in cell nuclei density are also features of G3. A relative increase in tenocyte numbers [Fig. 1(c)], compared to lower grades, is clearly seen. Finally, grade 4 [G4, Fig. 1(d)] is designated as severe degeneration. In this grade, the collagen fibers present a diffuse hyalinization of homogeneous appearance, the tenocytes present a rounded shape, complete loss fiber orientation occurs, and the number of nuclei is reduced.

3 Data Analysis

All the spectra were baseline corrected and vector normalized. The spectral differences were analyzed using multivariate Principal components analysis (PCA). PCA was performed over the range 365 to 1800 cm^{-1} by computing the covariance matrix. The underlying data structure was summarized by clustering PC2, PC3, and PC4 scores at a 95% level of similarity using correlation distance measurement. The results are presented as a dendrogram. The set of PCs that provided the best classification after visual inspection of scattering plots were fed into the logistic regression (LR) algorithm²¹ to determine the parameter equation that best differentiated the pathologic states. LR provides a method for modeling a binary response variable, considering values of 0 or 1, and is based on the linear dependence between the logit function of the probability of response 1 and the parameters of diagnosis. In this case, these parameters are the PCs. Thus, the LR model equation is

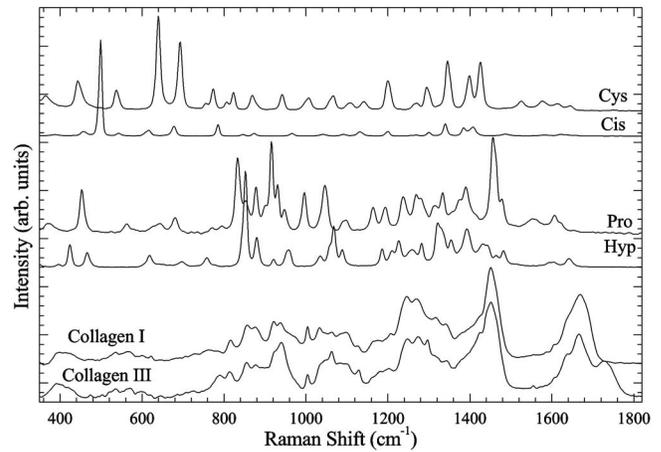


Fig. 3 FT-Raman spectra of cysteine (Cys), cistine (Cis), proline (Pro), hydroxyproline (Hyp), and collagen types I and III.

$$\ln\left(\frac{p}{1-p}\right) = a + \sum_i b_i PC_i, \quad (1)$$

where p is the probability of obtaining response 1, and a and b_i are the model parameters.

All these steps were performed using the statistical software Minitab, version 14.20 (Minitab Inc., State College, Pennsylvania, USA). The model's predictive ability was estimated by measuring the association between the response variable and predictive probabilities. To test the accuracy of the diagnosis, the receiving-operator characteristics (ROC) curves were calculated using a routine constructed using the Mathematica 5.2 software (Wolfram Research, Champaign, Illinois, USA). In these curves, sensitivity (S) is plotted against 1 minus specificity ($1-E$) and is not restricted to single values of S and E , which largely depend on the threshold curve chosen.^{22,23} The more accurately the method separates the data classes, the closer the corresponding area under the ROC curve (AUC) approximates^{22,23} 1. Both the LR model and the ROC curves were built using the leave-one-out and cross-validation procedures.

4 Results

4.1 Biochemical Analysis

Figure 2 shows the box plot of the 351 measured spectra. The black line is the mean spectrum, while the gray vertical lines represent the region between the first and third quartiles. Major spectral variability was found in the low-frequency region, $\omega < 800 \text{ cm}^{-1}$, while the high-frequency region, $\omega > 1450 \text{ cm}^{-1}$, presented more reproducible spectra.

Figure 3 shows the FT-Raman spectra of Cys, Cis, Pro, Hyp, and collagens types I and III. Together with glycine (Gly), these are the major tendon components.¹³ It was possible establish complete assignment for the bands displayed in Fig. 2 with help of these spectra and Refs. 24–29. The assignment is summarized in Table 1. Observation of this table verified that the Raman spectrum was dominated by amino acids and collagen protein bands, as expected from the tendon compositional profile.^{8,14,24} Figure 4 compares the mean FT-Raman spectra for each degenerative grade.

Table 1 Vibrational assignment for the mean tendon spectra of Fig. 3 based on Refs. 24–29.

Band Number	Position (cm ⁻¹)	Assignment	
1	430–550	s(S—S) bridges	
2	568	Pro—skeletal	
3	590	unassigned	
4,5	600, 607	w(CO ₂)—Gly	
6	618	Phe—skeletal	
7	646	Tyr—skeletal	
8	664	G—skeletal	
9	680	Pro—skeletal	
10,11	689, 703	b (CO ₂)—Gly	
12	727	collagens	
13	760	collagen I	
14	775	Pro—skeletal	
15	786	collagen III	
16	815	s (C—O—C; C—C); residue backbone	
17	857	(C—C); Pro ring	
18	877	s (C—C) Hyp	
19	922	s (C—C) Pro ring	
20	940	s (C—C) collagens backbone	
21	961	s (C—C) of residue	
22	970	s (S—S); Cis	
23	1004	s (C—C) Phe aromatic ring	
24	1034	s (C—N); alicyclic	lipids; nucleic acids
25	1063	b(OH)	
26	1085	s (C—O)	
27	1094	s (C—N); alicyclic	
28	1106	s (PO ₂ ⁻¹)	
29	1123	s (C—N; C—C)	
30	1160	NH ₃	
31	1175	s (C—C)	
32	1245–1345	Amide III—collagen	
33–35	1381, 1395, 1427	d (CH ₂)—residue	
36	1450	s (CH ₂)—lipids, glycosaminoglycans, metalloproteinases, collagens, residues	
37	1500–1572	s (C=N; C=C); DNA/RNA	
38	1604	Phe, Tyr	
39	1635–1665	Amide I	
40	1730–1740	collagen III	
41	1746, 1767, 1784	unassigned	

s; stretching; w; wagging; b; bending; d; distortion.

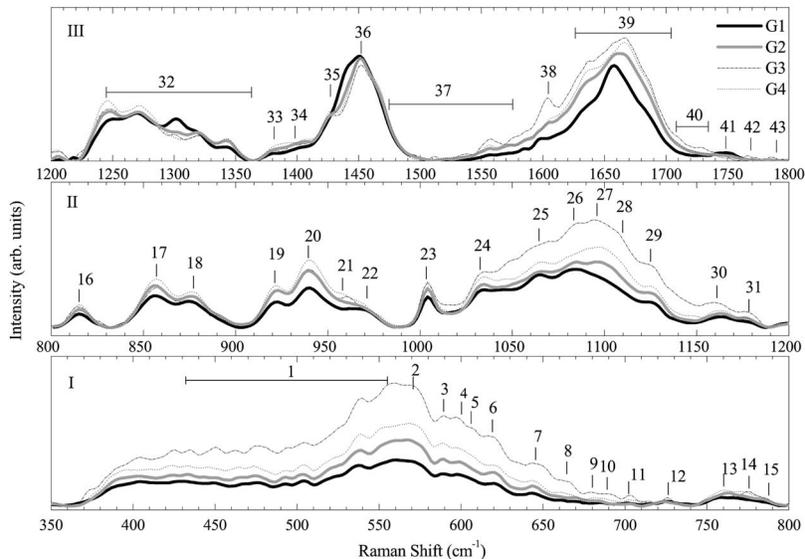


Fig. 4 Mean FT-Raman spectra for grades G1 to G4 between 350 and 800 cm^{-1} (panel I), 800 and 1200 cm^{-1} (panel II), and 800 and 1800 cm^{-1} (panel III). Each vibrational band is numbered according assignment shown in Table 1.

Figure 4 (panel I) shows the mean spectra between 350 and 800 cm^{-1} . At first glance, an increasing spectral intensity in the 350- to 710- cm^{-1} region is identifiable for the pathological groups G2, G3, and G4. This region is dominated by the stretching modes of Cys/Cis (disulphide bridges), skeletal modes of Pro, Phe, Tyr, and guanine (G), and CO_2 wagging/bending mode of Gly. An unidentified band at 590 cm^{-1} also occurred. The increased intensity was notably greater for G3 with $I(\text{G3})/I(\text{G1}) \sim 3$, in comparison with $I(\text{G4})/I(\text{G1}) \sim 1, 4$ and $I(\text{G2})/I(\text{G1}) \sim 1, 3$.

Characteristic collagen and Pro bands dominated the remaining region from 710 to 800 cm^{-1} . The band at 727 cm^{-1} is present in both type I and type III collagens and its intensity was almost grade independent. However, this fact cannot be considered conclusive, because the peak intensity itself is close to the background noise level. The collagen I band at 760 cm^{-1} presented a similar intensity for grades G1, G2, and G3, when considering the data noise level. G4 presented the greatest intensity.

Figure 4 (panel II) shows the mean spectra between 800 and 1200 cm^{-1} . The spectral region of 800 to 1030 cm^{-1} was dominated by C—O—C, C—C, and S—S stretching bands of residue backbone; collagen backbone; and amino acids Pro, Hyp, Phe, and Cis. G1 normal tendons presented a lower spectral intensity in this region. The intensity of G2 and G3 were almost the same, while G4 presented the greatest intensity.

The spectral region of 1030 to 1200 cm^{-1} is dominated by modes of C—N, C—O, C—C, NH_3 , DNA, and PO_2 OH of lipids and nucleic acids. The greatest intensity was displayed by the G3 spectrum.

Figure 4 (panel III) shows the mean Raman spectra between 1200 and 1800 cm^{-1} . The region between 1200 and 1500 cm^{-1} is dominated by amide III and distortion bands of collagens; CH_2 and CH_3 amino acid residues, lipids, glycosaminoglycans (GAGs), and metalloproteinases (MMPs). No marked intensity differentiation occurred, besides those

discussed previously, due to the fact that the variations in the relative amounts of GAGs and MMPs are very small and their bands are always superposed. Exception was the amide III band at 1300 cm^{-1} since it was very intense in the G1 spectra. The region between 1500 and 1700 cm^{-1} is dominated by C=N, C=C, C—N, and amide I vibrational bands of DNA, RNA, Phe, Tyr, and collagens. The remaining three bands at 1746, 1767, and 1784 cm^{-1} were not identified because their intensity variation occurred within background noise levels. Thus, they are not discussed or interpreted further.

To check whether all these spectral differences were sufficient to determine classification and assist coherent diagnosis for all degenerative grades, a complete statistical analysis was performed based on PCA and LR models.

4.2 Statistical Analysis for Diagnostic Purposes

To extract the essential correlation information from the data observed, PCA analysis was performed on the spectra presented in Fig. 2. The corresponding eigenvalues and contribution percentages of the first 10 PCs are shown in Table 2. From these data, it was verified that 96.9% of the data variability is retained up to the fourth PC. It is possible that higher components account for data noise; thus, further analysis will be based on PC1, PC2, PC3, and PC4 data.

Figure 5 shows the PC loadings for the first four PCs. The corresponding band assignment could be seen on Table 1, as presented on previous section. Figure 5(a) shows the first score (PC1), which represented 89.1% of the spectral characteristics of the data set. By comparing Fig. 2 to Fig. 5(a), it was possible to confirm the well-known fact that the first PC strongly resembled the overall spectral mean.

Figure 5(b) shows the second score (PC2), which accounted for 4.0% of the overall spectral characteristics. Compared to the first score, the second presented spectral negative variations in the 850- to 900-, 1050- to 1150-, 1257-, 1270-, 1300-, 1450-, 1660-, and 1751- cm^{-1} regions. In general,

Table 2 Eigenvalues and percentage contribution of the first 10 PCs.

PC	Eigenvalue	Percentage Contribution
1	196.89	89.1
2	8.91	4.0
3	5.35	2.4
4	2.99	1.4
5	1.14	0.5
6	0.89	0.4
7	0.39	0.2
8	0.28	0.1
9	0.26	0.1
10	0.20	0.1

negative or positive variation in a PC indicates that it represents these spectra with increasing or decreasing intensity compared to the mean.

Figure 5(c) shows the third score (PC3), responsible for 2.4% of the spectral characteristics. The behavior of the third score up to $\sim 800\text{ cm}^{-1}$ and within the $1060\text{--}1180\text{ cm}^{-1}$ region was very similar to the second score. Excluding these regions, the variations observed were positive and in the same spectral regions as those observed in the second score.

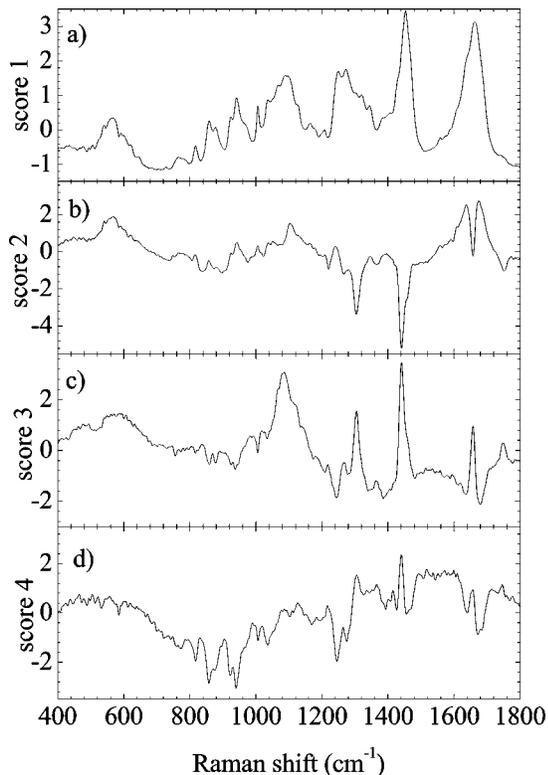


Fig. 5 PC-loadings of the first four PCs.

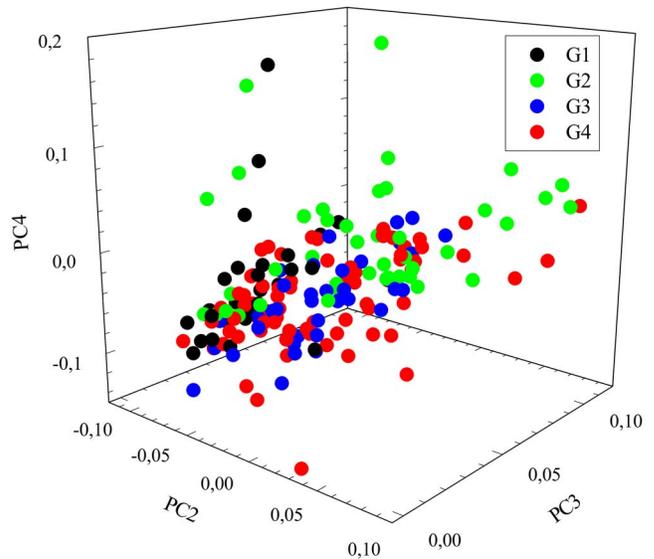


Fig. 6 Scattering plot for PC2, PC3 and PC4.

Figure 5(d) shows the fourth score (PC4), which was close to background noise levels and was responsible for 1.4% of the data spectral variability. Comparing its behavior with those displayed by Figs. 5(b) and 5(c), it is clear that PC4 represented a linear combination of PC2 and PC3 with some peaks resembling the behavior of PC2 and others, those of PC3. However, after a peak-to-peak comparison, it was inferred that the fourth score showed greater similarity to the third score than to the second.

Thus, it is reasonable to expect that classification between tendon degenerative states could be obtained by some combination of PC2, PC3, and PC4.

Figure 6 shows the scattering plot of PC2, PC3, and PC4. Each PC was discriminated by its histopathological grade G1 to G4, following Riley's scale. Unfortunately, direct visual inspection clearly shows that no data separation occurred.

To more clearly discriminate between spectral data groupings, clustering of the PC2, PC3, and PC4 variables was performed considering a threshold of 95% similarity. The results are shown in a dendrogram (Fig. 7, panel I). As can be observed, the data grouped into two large clusters (A and B), with cluster B further divided into two subclusters (B1 and B2). Cluster A was composed of 95 spectra: 40% from samples diagnosed as G1, 47% from G2 samples, 9% from G3, and 4% from G4. Cluster B1 was larger, composed of 150 spectra and of these, 16, 13, 47, and 24% were G1, G2, G3, and G4 samples, respectively. Cluster B2 was composed of spectra from 106 samples, 2, 47, 27, and 24% from G1, G2, G3, and G4, respectively. Considering clusters B1 and B2 together (cluster B), the percentages of samples were 11, 26, 39, and 24% of G1, G2, G3, and G4, respectively. Summarizing, it was possible to infer that G1 and G2 samples were predominant in cluster A (87%), while those of grades G3 and G4 predominated in cluster B (63%). According to the main histological and biochemical characteristics of the degenerative grades, note that the presence of hyalinization is a shared characteristic of grades G3 and G4. In contrast, hyaline regions are absent in both grades G1 and G2. Thus, cluster A is,

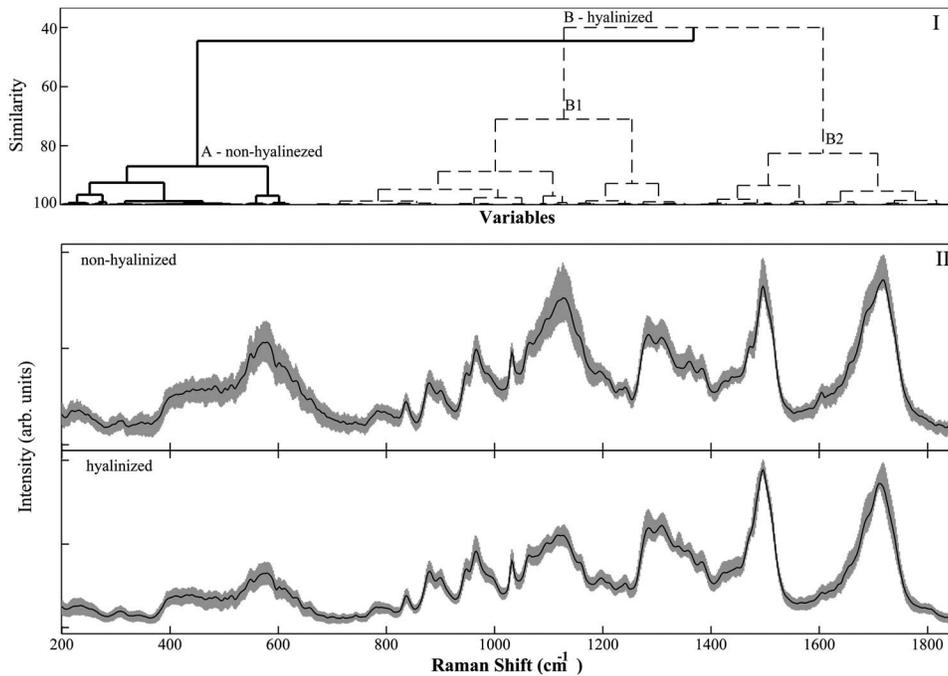


Fig. 7 Panel I, dendrogram based on the PC2, PC3, and PC4, showing the separation of the spectra in two clusters (A and B). Group B presents two important subgroups (B1 and B2). Spectra from nonhyaline samples (G1 and G2) predominate in cluster A (87%), while the hyaline samples (G3 and G4) predominate in cluster B (63%). Panel II, box plot for hyalinized (upper) and nonhyalinized (bottom) clusters.

to a large extent, composed of nonhyaline tendons from G1 and G2, while cluster B is mainly composed of those tendons presenting hyalinization from G3 and G4 grades. Thus, we labeled clusters A and B as nonhyalinized and hyalinized clusters, respectively.

Figure 7 (panel II) shows the box plot for the hyalinized (upper plot) and nonhyalinized (bottom plot) groups. Nearly all spectral regions in the hyalinized groups were more intense than in the nonhyalinized group. This fact is in agreement with data previously presented in Fig. 4, which showed that the mean spectra of G3 and G4 were generally more intense than those of G1 and G2.

Figure 8(a) shows the scattering plot PC4 versus PC2, considering these two classes of tendons. Points arising from the hyalinized tissues presented a relatively clear separation from the nonhyalinized points. It is easy to discern a nonhyalinized region to the left and a hyalinized region to the right. The LR model was applied to this case and the best fit was found to be

$$\ln\left(\frac{p}{1-p}\right) = 0.023 + 1.005PC2 + 16.713PC4, \quad (2)$$

where p is the probability of belonging to the hyalinized class. The model's predictive ability was estimated by measuring the association between the response variable and predictive probabilities using cross-validation. In this case, 72.2% concordant pairs were found, indicating good predictive ability. The solid line in Fig. 8(a) represents the diagnosis line threshold $p=0.50$. The calculated sensitivity was 54.0%, while specificity was 96.0% when considering the threshold line $p=0.50$. A similar classification level was found by analyzing the scattering plot PC3 versus PC2 [Fig. 8(b)]. In this case, the best fit for the LR model was found to be

$$\ln\left(\frac{p}{1-p}\right) = 0.029 + 1.285PC2 - 12.583PC3. \quad (3)$$

The percentage of concordant pairs was 76.4%. Sensitivity and specificity were 54.1 and 76.0%, respectively, when considering the diagnosis line $p=0.50$ depicted in Fig. 8(b).

Figure 8(c) shows the scattering plot PC4 versus PC3. In this case, the best fit for the LR model was found to be

$$\ln\left(\frac{p}{1-p}\right) = 0.0069 - 14.96PC3 + 16.978PC4. \quad (4)$$

The percentage of concordant pairs was 79.1%. Sensitivity and specificity were 65.5 and 75.0%, respectively, when considering the diagnosis line depicted in Fig. 8(c).

The LR model was also tested for PC2, PC3, and PC4 combined, and the best fit was found to be

$$\ln\left(\frac{p}{1-p}\right) = -0.017 + 2.119PC2 - 15.393PC3 + 17.231PC4 \quad (5)$$

The percentage of concordant pairs was 79.6%. Sensitivity and specificity were 66.0 and 74.7%, respectively. Thus, comparatively this combination provided slightly better discrimination than the remainder.

Figure 9 shows the ROC curves for the combinations (PC2, PC3); (PC2, PC4); (PC3, PC4); and (PC2, PC3, PC4). The straight diagonal line is the expected curve for a completely random diagnostic test. The AUCs were computed and are shown in Table 3. The most accurate test was found using the combination (PC2, PC3, PC4), which resulted in an AUC

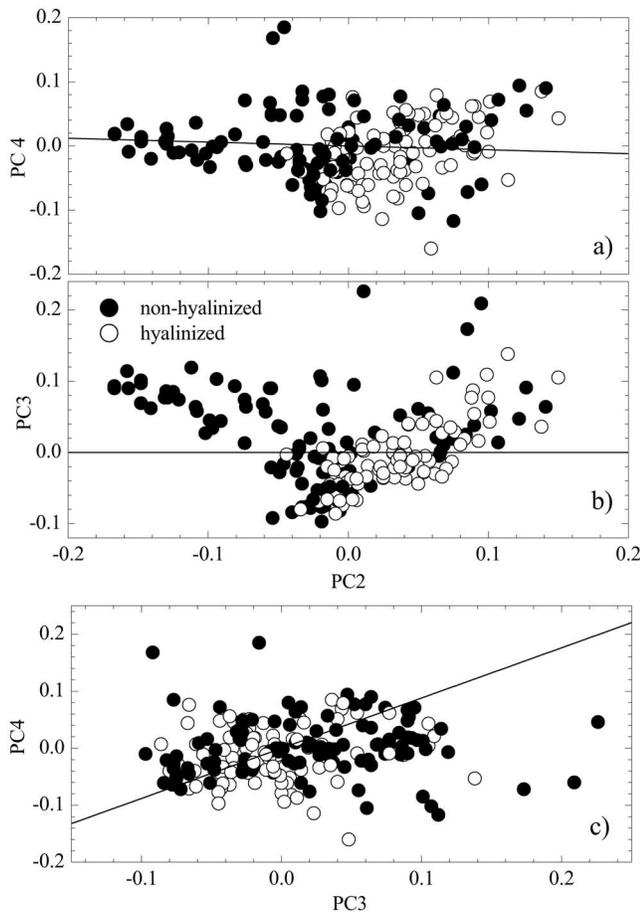


Fig. 8 Scattering plots of (a) PC4 versus (b) PC2, PC3 versus PC2, (c) PC4 versus PC3 for the nonhyalinized (solid circle symbols) and hyalinized (open circle symbols) clusters. The dotted line is a diagnosis line given by the LR model [Eq. (1)] with $p=0.50$.

of 0.81, classified as a good level of accuracy. The combinations (PC2, PC4) and (PC3, PC4) resulted in fair accuracy levels and the worst test was found using the combination (PC2, PC3) (poor accuracy). These results are in agreement with the predictive ability estimation of the LR presented previously, which provided the highest percentage of concordant pairs for the combinations (PC2, PC3, PC4) and (PC2, PC3).

5 Discussion

The biochemical data concerning the tendon degenerative grade probed by Raman spectroscopy (see Fig. 4) can be summarized in three groups. The first group represents changes observed in Gly, Pro, Hyp, Cys, Cis, collagen I, and collagen III bands. An increasing intensity for these bands in degenerated tissues was found, mainly in grades G3 and G4. Gly, Pro, and Hyp are the main constituents of the repeated amino acid motif Gly-X-Y (X, Y=Pro, Hyp) of collagen structure.³⁰ Cys and Cis are common residues (X or Y) found in type III collagen.³⁰ Type III collagen is usually found in degenerated tendons.³¹ The increased intensity of these bands is explained by the hyalinization process that occurs in degenerated G3 and G4 tendons, where peptidic fragments of soluble collagen accumulate in the extracellular matrix.¹³ In healthy tendons,

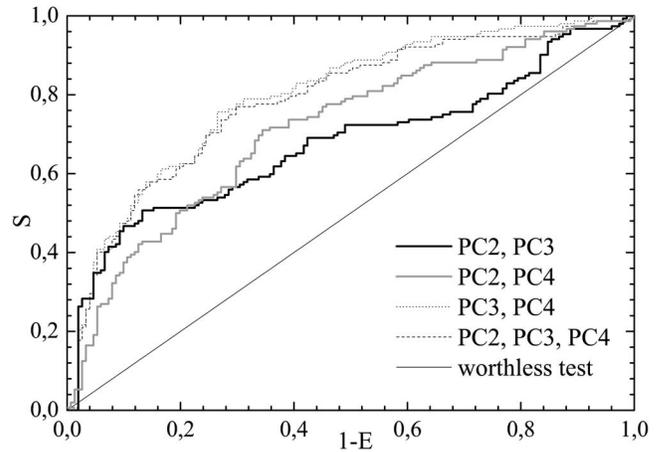


Fig. 9 ROC curves for combinations of PC 2 and PC 3 (solid black line); PC 2 and PC 4 (solid gray line); PC 3 and PC 4 (dotted line); PC 2, PC 3, and PC 4 (dashed line); and the worthless test (thin solid line).

no collagen accumulation occurs in the extracellular matrix, because it is degraded by MMP. One of the characteristics of the degenerative process in tendons is the decrease in MMP content in the matrix.³⁰

The second group is related to Phe, Tyr, G, DNA, and RNA bands. Their Raman bands experienced an enhanced intensity in grades G3 and G4. However, this was more expressive for G3. The amino acids Phe and Tyr participate in the DNA transcription process³² and are related to high cellular activity and reproduction.³² They are also present at low levels in collagen structure.³⁰ Their enhanced intensity, and that of the G, DNA, and RNA bands, correlates with the high cellular nuclei density, which is characteristic¹³ of grade G3.

The third group is related to lipids bands. These bands were displayed in the 1030- to 1200- cm^{-1} region and at 1450 cm^{-1} , in conjunction with other biochemical components. This fact makes it difficult to evaluate lipid behavior in isolation. The first bands appear in conjunction with nucleic acid bands and present an intensity enhancement that correlates with fatty degeneration of tendons.^{8,13} However, the band at 1450 cm^{-1} displayed no intensity alteration. In fact, this band is related to lipid, GAG, MMP, and collagen residues. Lipids and collagens were expected to present enhanced concentrations in degenerated tissues, while GAG and MMP were expected to show diminished levels. No marked intensity differentiation occurred, besides those discussed previ-

Table 3 Area under curve (AUC) of the ROC curves presented in Fig. 9.

PCs	AUC	Accuracy
2, 3	0.67	poor
2, 4	0.71	fair
3, 4	0.78	fair
2, 3, 4	0.81	good

ously, due to the fact that the variations in the relative amounts of GAGs and MMPs are very small and their bands are always superposed.

The just discussed band intensity variations were also observed in the PC loadings plots of Fig. 5. However, despite the fact that Raman spectroscopy is capable of displaying the spectral characteristics of all the degenerative grades no data classification in G1, G2, G3, or G4 grades was observed (Fig. 6).

Observation revealed that the main features probed were those related to the absence (G1 and G2) or presence (G3 and G4) of hyalinization and the distinctive cell proliferating activity of G3. This fact was strongly supported by the statistical analysis performed on the data, which showed a clear separation of the spectra into nonhyalinized and hyalinized clusters (Fig. 7), enabling the construction of a binary diagnosis model based on LR (Fig. 8). The best diagnosis was found when using a combination of PC2, PC3, and PC4. The diagnosis line threshold $p=0.50$ resulted in sensitivity of 66.0% and specificity of 74.7%. The percentage of concordant pairs was 79.6%. The AUC (Fig. 9) indicated that a diagnosis based on these PCs presented good accuracy (AUC=0.81).

6 Conclusions

The results presented in this paper indicated that the FT-Raman spectroscopy technique is capable of classifying rotator cuff supraspinatus tendons as either nonhyalinized (G1 and G2) or hyalinized (G3 and G4). Notwithstanding this fact, it was not possible discriminate all degenerative grades, the identification of hyaline (presenting fragmented and soluble collagens) and nonhyaline (structured waviness fibril collagens) regions in these tendons is of great clinical relevance to the surgeon, who would otherwise have no way to visually discriminate these tendon regions during surgery. In principle, Raman spectroscopy could be used as an auxiliary aid to improve tendon shoulder surgery quality, thus contributing to a decrease in the current high rerupture rate of this procedure, by assisting in determining the process of anchoring onto more healthy tendons. However, more experimental efforts must be devoted to this issue. For *in vivo* applications it is essential to establish whether the resolution level of actually available Raman probes are adequate to obtain a classification similar to those obtained by FT-Raman spectroscopy. Another possible improvement is the validation of the LR models by an independent data set.

Acknowledgments

The authors would like to thank the Brazilian agencies FAPESP (2001/14384-8; 2007/01622-4) and CNPq (302393/2003-0; 301018/2006-5) for providing financial support and Claudio Menezes, MD, of the CIPAX Medical Diagnosis, for fruitful discussions and the careful reading of the slides.

References

1. J. J. Luime, B. W. Koes, I. J. M. Hendriksen, A. Burdorf, A. P. Verhagen, H. S. Miedema, and J. A. N. Verhaar "Prevalence and incidence of shoulder pain in the general population; a systematic review," *Scand. J. Rheumatol.* **33**, 73–81 (2004).
2. D. A. W. Windt, E. Thomas, D. P. Pope, A. F. de Winter, G. J. Macfarlane, L. M. Bouter, and A. J. Silman, "Occupational risk factors for shoulder pain: a systematic review," *Occup. Environ. Med.* **57**, 433–442 (2000).
3. P. M. Bongers, C. R. de Winter, M. A. Kompier, and V. H. Hildebrandt, "Psychosocial factors at work and musculoskeletal disease," *Scand. J. Work Environ. Health* **19**, 297–312 (1993).
4. E. F. Harkness, G. J. Macfarlane, E. S. Nahit, A. J. Silman, and J. McBeth, "Mechanical and psychosocial factors predict new onset shoulder pain: a prospective cohort study of newly employed workers," *Occup. Environ. Med.* **60**, 850–857 (2003).
5. H. P. Mendonça Jr. and A. A. Assunção, "Association between shoulder diseases and work: a brief review," *Rev. Bras. Epidemiol.* **8**, 167–176 (2005).
6. P. M. Bongers, A. M. Kremer, and J. Laak, "Are psychosocial factors, risk factors for symptoms and signs of the shoulder, elbow, or hand/wrist?: a review of the epidemiological literature," *Am. J. Ind. Med.* **41**, 315–342 (2002).
7. B. G. Frieman, T. J. Albert, and J. M. Fenlin Jr., "Rotator cuff disease: a review of diagnosis, pathophysiology, and current trends in treatment," *Arch. Phys. Med. Rehabil.* **75**, 604–609 (1994).
8. P. Sharma and N. Maffulli, "Basic biology of tendon injury and healing," *Surgeon* **3**, 309–316 (2005).
9. L. J. Soslowky, S. Thomopoulos, A. Esmail, C. L. Flanagan, J. P. Iannotti, J. D. Williamson, and J. E. Carpenter, "Rotator cuff tendinosis in an animal model: role of extrinsic and overuse factors," *Ann. Biomed. Eng.* **30**, 1057–1063 (2002).
10. C. Guckel and A. Nidecker, "Diagnosis of tears in rotator-cuff-injuries," *Eur. J. Radiol.* **25**, 168–176 (1997).
11. J. P. Bonde et al., "Prognosis of shoulder tendonitis in repetitive work: a follow up study in a cohort of Danish industrial and service workers," *Occup. Environ. Med.* **60**, E8 (2003).
12. C. S. Neer, "Impingement lesions," *Clin. Orthop. Relat. Res.* **173**, 70–77 (1983).
13. G. P. Riley, "Chronic tendon pathology: molecular basis and therapeutic implications," *Expert Rev. Mol. Med.* **7**, 1 (2005).
14. H. Alfredson and R. Lorentzon, "Chronic tendon pain: no signs of chemical inflammation but high concentrations of the neurotransmitter glutamate. Implications for treatment?" *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **3**, 43–54 (2002).
15. C. W. Goodmurphy, J. Osborn, E. J. Akesson, S. Johnson, V. Stanescu, and W. regan, "An immunocytochemical analysis of torn rotator cuff tendon taken at the time of repair," *J. Shoulder Elbow Surg.* **12**, 368–374 (2003).
16. K. Yamaguchi, C. M. Ball, and L. M. Galatz, "Arthroscopic rotator cuff repair: transition from mini-open to all-arthroscopic," *Clin. Orthop. Relat. Res.* **390**, 83–94 (2001).
17. K. Yamaguchi, "Mini-open rotator cuff repair: an updated perspective," *Instr Course Lect* **50**, 53–61 (2001).
18. B. Jost, C. W. Pfirrmann, C. Gerber, and Z. Switzerland, "Clinical outcome after structural failure of rotator cuff repairs," *J. Bone Jt. Surg., Am. Vol.* **82**, 304–314 (2000).
19. G. P. Riley, M. J. Goddard, and B. L. Hazleman, "Histopathological assessment and pathological significance of matrix degeneration in supraspinatus tendons," *Rheumatology (Oxford)* **40**, 229–230 (2001).
20. B. P. Fogazza, C. da Silva Carvalho, S. G. Penteado, C. S. Menezes, A. A. Martin, and H. da Silva Martinho, "Classification of degenerative grading of lesions of supraspinatus rotator cuff tendons by FT-Raman spectroscopy," *Proc. SPIE* **6445**, 64450S (2007).
21. V. Bewick, L. Cheek, and J. Ball, "Statistics review 14: logistic regression," *Crit. Care* **9**, 112–118 (2005).
22. T. A. Lasko, J. G. Bhagwat, K. H. Zou, and L. Ohno-Machado, "The use of receiver operating characteristic curves in biomedical informatics," *J. Biomed. Inf.* **38**, 404–415 (2005).
23. L. E. Dodd and M. S. Pepe, "Partial AUC estimation and regression," *Biometrics* **59**, 614–623 (2003).
24. T. Ikoma, H. Kobayashia, J. Tanaka, D. Walsh, and S. Mann, "Physical properties of type I collagen extracted from fish scales of Pagrus major and Oreochromis niloticas," *Int. J. Biol. Macromol.* **32**, 199 (2003).
25. A. Mahadevan-Jansen and R. Richards-Kortum, "Raman spectroscopy for the detection of cancers and precancers," *J. Biomed. Opt.* **1**, 31–70 (1996).
26. A. Pawlukoic, J. Leciejewicz, A. J. Ramirez-Cuesta, and J. Nowicka-Scheibe, "l-cysteine: neutron spectroscopy, Raman, IR and ab initio study," *Spectrochim. Acta, Part A* **61**, 2474–2481 (2005).
27. D. Naumann, "FT-infrared and FT-Raman spectroscopy in biomedical research," *Appl. Spectrosc. Rev.* **36**, 239–298 (2001).

28. Y. Shi and L. Wang, "Collective vibrational spectra of α - and γ -glycine studied by terahertz and Raman spectroscopy," *J. Phys. D* **38**, 3741–3745 (2005).
29. R. A. Bitar, H. Martinho, C. J. Tierra-Criollo, L. N. Z. Ramalho, M. M. Netto, and A. A. Martin, "Biochemical analysis of human breast tissues using Fourier-transform Raman spectroscopy," *J. Biomed. Opt.* **11**, 54001 (2006).
30. M. van der Rest and R. Garrone, "Collagen family of proteins," *Collagens* **5**, 2814–2823 (1991).
31. G. P. Riley et al., "Tendon degeneration and chronic shoulder pain: changes in the Collagen composition of the human rotator cuff tendons in rotator cuff tendinitis," *Ann. Rheum. Dis.* **53**, 359–366 (1994).
32. L. Lehninger, D. L. Nelson, and M. Cox, *Lehninger Principles of Biochemistry*, 4th ed., W. H. Freeman & Co., New York (2004).