THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Infrared Metrics for Fixation-Free Liver Tumor Detection

Zhaomin Chen,[†] Ryan Butke,[‡] Barrie Miller,[‡] Charles L. Hitchcock,[‡] Heather C. Allen,*[,†,‡] Stephen P. Povoski,[§] Edward W. Martin, Jr.,[§] and James V. Coe*[,†]

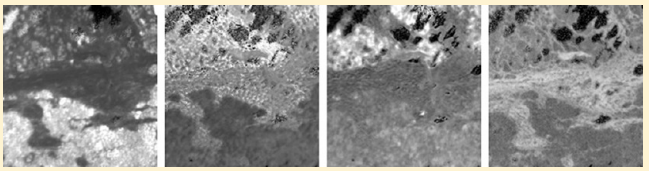
[†]Department of Chemistry and Biochemistry, The Ohio State University, 100 West 18th Avenue, Columbus, Ohio 43210-1173, United States

[‡]Department of Pathology, The Ohio State University, 4132 Graves Hall, 333 West 10th Avenue, Columbus, Ohio 43210, United States

[§]Department of Surgery, Division of Surgical Oncology, The Ohio State University, 410 West 10th Avenue, Columbus, Ohio 43210, United States

Supporting Information

ABSTRACT: Infrared (IR) spectroscopic imaging of human liver tissue slices has been used to identify and characterize liver tumors. Liver tissue, containing a liver metastasis of breast origin (mucinous carcinoma), was surgically removed from a consenting patient and frozen without formalin fixation or dehydration procedures, so that lipids and water remained in the tissues. A set of IR metrics (ratios of various IR peaks) was determined for tumors in fixation-free liver tissues. K-means cluster analysis was used to tell tumor from nontumor. In this case, there was a large reduction in lipid content upon going from nontumor to tumor tissue, and a well-resolved IR spectrum of nontumor liver lipid was obtained and analyzed. These IR metrics may someday guide work on IR spectroscopic diagnostics on patients in the operating room. This work also suggests utility for these methods beyond the identification of liver tumors, perhaps in the study of liver lipids.



INTRODUCTION

It is possible that infrared (IR) spectroscopy could prove useful as a real-time intraoperative diagnostic tool,[1,2] without the need for histologic fixation or staining of examined tissues, because IR spectra can be recorded rapidly and with high signal-to-noise ratio. In this regard, one can envision utilizing a fiber optic attenuated total reflection (IR-ATR) probe[3] to evaluate potentially diseased organs or tissues in situ during surgical procedures. Ideally, a surgeon who is attempting to resect a tumor would like to know whether the entire tumor has been successfully removed and whether the surgical resection margins are negative for residual tumor before leaving the operating room. However, the ability to obtain comprehensive real-time information regarding completeness of surgical resection is generally highly labor-intensive, time-consuming, and fraught by a high level of false-negative results. This is secondary to the intrinsic nature in which the evaluation of the surgical resection margin is conducted by the pathologist and is primarily limited by the fact that the pathologist only evaluates a minute fraction of the entire resected circumferential boundary of the tumor from within the operative field. The aim of the current work is to identify the most important and characteristic IR spectral features associated with tumor-bearing tissues within the liver, with the ultimate goal to develop a method for real-time differentiation of specific cell types associated with diseased liver tissue and normal liver tissue that can be used to assist in the surgical resection of liver tumors.

Liver tissue containing a liver metastasis of breast origin (mucinous carcinoma) was surgically removed from a consenting patient at the time of a planned liver resection and was frozen without formalin fixation or dehydration procedures (i.e., lipids and water remained in the tissues). Fresh liver tissue was snap frozen in liquid nitrogen, and cryostat sections were obtained from which imaging IR spectra were recorded. A similar slice was also treated with a hematoxylin and eosin (H&E) stain[4] and processed as a virtual Aperio slide for comparison to the IR results. A method based on k-means cluster analysis[5−7] was used to identify a set of IR metrics that are important for identifying liver tumors. Some changes in chemical composition between the tumor and its surrounding tissue are presented.

In 2012, approximately 1.6 million new cancer cases were expected to be diagnosed in the United States, and approximately 570 000 Americans were expected to die of cancer.[10] Among these 1.6 million new cancer cases, there were approximately 29 000 new cases of primary liver cancer. Yet, more astonishingly, the incidence of secondary liver tumors (i.e., liver metastases) in the United States is estimated to be as many as 20-times greater than the incidence of primary liver cancer. Such liver metastases most commonly originate from colorectal, breast, and lung cancer cases. From a global

Received: July 23, 2013
Revised: September 18, 2013
Published: September 20, 2013

dx.doi.org/10.1021/jp4073087 | J. Phys. Chem. B 2013, 117, 12442−12450

perspective, the ratio of mortality to incidence of liver cancer is roughly 0.93 (www.who.int/mediacentre/factsheets/fs297/en, http://globocan.iarc.fr/factsheets/cancers/liver.asp). Clearly, tumors of the liver are an important, global problem and the motivation for the current study.

There is an extensive history of IR bioimaging,[8−13] including research on cancer with tissues other than liver[14−19] and some IR spectral work on human liver[20−26] and mice liver tissues.[27] Particularly notable to us was the methodology of Fernandez et al.,[28] who recently reported a table of IR metrics for prostate cancer. These metrics serve as a starting point for our work on IR metrics of liver tissue. Several groups[19,29,30] have shown that the functional equivalent of an H&E stained image could be extracted from IR spectroscopic imaging information on tissues slices. In our current work, we evaluate the utility of the prostate tumor IR metrics of Bhargava and co-workers[28] for liver and augment these with metrics important when the tissue is not fixed. All of the potential metrics were merged into a large set, and a quantitative determination revealed those most important for evaluation of tumors within the liver.

The paper continues with an Experimental Section giving a description of the collection of human liver tissue, the recording of IR spectroscopic imaging data, and software that extracts the IR spectra from the commercial software enabling manipulation with statistical programs. The next section introduces the types of IR metrics under consideration, which is followed by a k-means cluster analysis using scaled IR metrics in order to identify the best IR metrics. Finally, the results of k-means cluster analysis with the unscaled subset of best IR metrics are described using both 5 and 25 groups. The paper concludes with a discussion of the k-means cluster results for liver diagnostics.

EXPERIMENTAL SECTION

The liver tissue was collected during surgery at the University Hospital (Ohio State University, Columbus, OH). The collection, reservation, and resecting of the tissue was approved by the Institutional Review Board (IRB, No. 2011C0085). Immediately after collection, the specimen was pretreated by snap freezing in liquid nitrogen, which allowed rapid lowering of the sample to a temperature below −70 °C. Snap freezing of the liver tissue provided specimen integrity for further analysis. A cryostat, set at −20 °C, was used to obtain 2−3 μm thick sections of tissue. Normally, a pathologist will fix the tissue, that is, soak the tissue in a neutral buffered formalin solution and then dehydrate the sample with a sequence of graded ethanols, xylene, and finally parrafin. This preserves the tissue and works well with various staining procedures. However, research presented here concerns the possibility of taking in vivo IR spectra; therefore, the tissue was not fixed. It was sliced to a thickness of 2−3 μm, and the slice was further cropped to fit onto an IR ZnSe window (8.0 mm diameter, 1.0 mm thickness, from Crystran Ltd. U.K.) within a home-designed, sealed sample holder. An optical microscope image of the sample is shown in Figure 1a, where the intense yellow of the ZnSe window has been digitally reduced.

IR spectra were recorded with an imaging FTIR microscope (Perkin-Elmer Spotlight 300) with a computer-controlled microscope stage and a liquid-nitrogen-cooled linear array of 16 mercury cadmium telluride (MCT) detectors. The data were recorded in seven windows, as shown with dotted lines in Figure 1a. Because the imaging region is bigger than the microscope's field of view, the instrument patches together

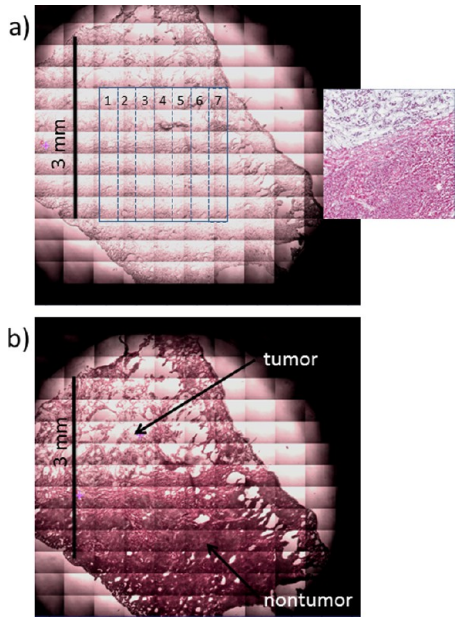


Figure 1. Optical microscope images of a slice of liver tissue with a tumor on ZnSe for IR spectroscopic imaging (scale on left). (a) Seven windows (top) were merged into one window (2.1000 mm wide by 2.2000 mm tall) for IR spectroscopic imaging. An Aperio H&E stain for an optical microscope of a nearby slice at approximately the same place and size is inset at the right. The bottom (b) shows the same tissue as in (a) after H&E staining on ZnSe. The tumor is lighter and at the top half of the images.

multiple optical views, producing a rectangular tiling artifact that is apparently accentuated by the bright yellow color of ZnSe. Despite the tiling artifact, this image can be overlaid, in a pixel-by-pixel fashion with the k-means images of the following sections. Each IR window region (shown in Figure 1a) was 2200.0 μm vertically by 300.0 μm horizontally (352 pixels by 48 pixels). The instrument obtained a full IR spectrum at each image pixel (6.25 μm by 6.25 μm area) in each window (4 cm^{-1} resolution, 2 cm^{-1} steps in the range of 750−4000 cm^{-1}, and 16 scans per pixel). Each of the seven windows required about 3 h of scanning time for a total of 21 h of data collection.

Later, the data were merged with home-written Matlab (version R2013a) routines into one large window that was 2200.0 μm by 2100.0 μm, including 118 272 separate pixels and/or IR spectra. All of the following tissue slice images are of this size. A Matlab routine from Perkin-Elmer called "fsm_load.m" by Ben Peterson facilitated the process (http://www.mathworks.it/matlabcentral/fileexchange/22736-perkinelmer-ir-data-file-import-tools). Our home-written Matlab programs are described in the MS thesis of Z. Chen.[31] The starting point for the analysis is a three-dimensional matrix of spectra transmittances, data(i,j,k'), where i is a pixel index for the image row, j is a pixel index for the image column, and k' is an index stepping through the IR spectrum. After the IR spectroscopic imaging, the sample was taken from the instrument and treated with an H&E stain[4] (as shown in Figure 1b). The intense yellow of the ZnSe window has been digitally reduced in the same manner as with Figure 1a. The staining shows a tumor at the top half of the merged window, and therefore, the bottom half was a region just outside of the tumor.

dx.doi.org/10.1021/jp4073087 | J. Phys. Chem. B 2013, 117, 12442−12450

IR METRICS

Considering that there are 1626 spectral steps (potential metrics from peak intensities) in each IR spectrum and 118 272 spectra, there are more than 192 million pieces of information to be analyzed. Clearly, there is utility in reducing the spectral information to something more manageable and understandable, hence the utility of IR metrics. We started with the IR metrics ($b1-b36$ in Table 1s of the Supporting Information) of Fernandez et al.[28] (for prostate cancer on fixed tissues) and added a set that might be more useful without fixation ($b37-b64$ in Table 1s of the Supporting Information). Matlab routines facilitate the calculation of a value of each metric at each pixel in the image. The result is a three-dimensional matrix of IR metrics, $b(i,j,m)$, where i and j are indices over the rows and columns, respectively, of the image and m is an index over the metrics. The absorbance intensities at each peak were corrected with a two-point baseline specific to each IR peak, eliminating the need for baseline offsetting and flattening before analysis. The metrics reduce the original data, a spectrum of 1626 points at each image pixel, to 64 values at each pixel. The process of k-means clustering was used to evaluate the quality of the metrics leading to a reduced set of IR metrics that are highly appropriate for liver tissue work on fixation-free tissue with lipids.

K-MEANS CLUSTER ANALYSIS WITH SCALED METRICS

Variations apparent in the IR spectra at each image pixel encouraged an attempt to sort the pixels into groups using IR metrics. K-means clustering is one of the simplest clustering techniques, whereby clusters or groups are characterized by centroids (essentially average values of the members of the group). The technique varies the position of the centroids and the membership in the groups to minimize the sum of "distances" between group members and the group centroid. Unlike the normal three-dimensional distances of Euclidean space, these distances are hyperdimensional, having a component contribution from each of the metrics (64 in this part of the evaluation).

In order to use the metrics results with Matlab routines for k-means clustering, the values of each metric, m, in the two-dimensional plane of the image had to be mapped into a single column. Using Matlab's linear mapping convention, the value of each biomarker at each pixel in the two-dimensional image was mapped into a one-dimensional array ($i,j \rightarrow n$), where n is a one-dimensional index of the pixels $[n = (j - 1)i_{max} + i]$. In other words, the three-dimensional metrics data, $b(i,j,m)$, was mapped into a two-dimensional matrix, $X(n,m)$, where n is a single index over the pixels and m is an index over the metrics. Upon defining an index k over the groups, G_k, defined by k-means clustering, the centroids for each group can be denoted as $c(k,m)$. Then, k-means clustering minimizes the quantity

$$\sum_k \left(\sum_{\text{if } n \in G_k} \sqrt{\sum_m [X(n, m) - c(k, m)]^2} \right) \quad (1)$$

by changing group membership and centroid position. The quantity $(\sum_m [X(n,m) - c(k,m)]^2)^{1/2}$ is the hyperdimensional distance between each metric score, $X(n,m)$, for an image pixel in a group and its centroid, $c(k,m)$. This leads to a criterion for evaluating the metrics with regard to distinguishing groups. Noting that the distance between two groups, G_{k1} and G_{k2}, is

$$d_{G_{k1},G_{k2}} = \sqrt{\sum_m [c(k1, m) - c(k2, m)]^2} \quad (2)$$

the contribution of a metric to the separation between groups is

$$d_{G_{k1},G_{k2},m} = |c(k1, m) - c(k2, m)| \quad (3)$$

Therefore, the metrics can be evaluated by the size of $d_{G_{k1},G_{k2},m}$, which determines the importance of each metric to the separation of those groups. However, the metrics have various average values and different distributions that can bias a numerical comparison; therefore, they were all scaled to the same average value (0.5), and the estimated standard deviation of each metric was used to set a range, that is, the scaled value of zero was two standard deviations below the scaled average and the scaled value of one was set to two standard deviations above the average.

K-means cluster analysis was performed with five groups and scaled metrics (an image, Figure 1s, is given is the Supporting Information). There was one nontumor group at the bottom half of the image, three tumor groups, and a group of holes (some holes in the tissue slices were inadvertently created through the microtoming process). Matlab programs produce a black-and-white bitmap image of each metric, a color bitmap image of each of the groups, the centroid component of each metric for each group, the average x and y position of each group, and the IR spectrum of each group. Gray scale images of selected metrics (scaled) are given in Figure 2, showing that

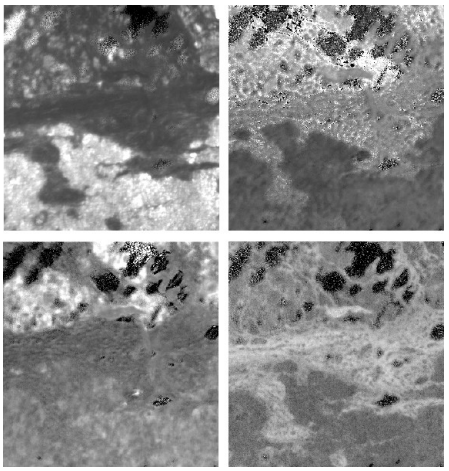


Figure 2. Gray scale bitmap images of selected scaled metrics, L1, L7, L14, and L18 from top left to bottom right. The tumor is largely in the top half of the images.

different IR metrics are sensitive to different aspects of the liver tissue. The contribution of each scaled metric to the difference between groups (eq 3) was calculated for the nontumor group and two of the tumor groups. These contributions were added for each of the 64 initial metrics and sorted from highest to lowest. The top 20 are given in Table 1 in order of importance.

RESULTS WITH UNSCALED IR METRICS

Using the top 20 metrics in Table 1 without scaling, a k-means cluster analysis was performed using five groups producing the image shown in Figure 3. The nontumor is green, and the holes are yellow. Again, there are three groups within the tumor (red, blue, and cyan). The IR spectrum of each group (baseline-flattened and offset to zero) is presented in the bottom part of

dx.doi.org/10.1021/jp4073087 | J. Phys. Chem. B 2013, 117, 12442−12450

Table 1. Top 20 IR Metrics for Fixation-Free Liver Tumor Tissues[a]

name	ratio	name	ratio
*L*1(*b*43)	1744/1548	*L*11(*b*19)	1016/1080
*L*2(*b*37)	1744/1244	*L*12(*b*59)	1252/1544
*L*3(*b*64)	1742/1256	*L*13(*b*39)	1024/1080
*L*4(*b*53)	1160/1548	*L*14(*b*42)	1080/1244
*L*5(*b*28)	1516/1236	*L*15(*b*47)	1516/1582
*L*6(*b*54)	2916/1548	*L*16(*b*7)	1744/1162
*L*7(*b*45)	1120/1020	*L*17(*b*23)	1080/3290
*L*8(*b*46)	2924/1544	*L*18(*b*58)	1556/1548
*L*9(*b*48)	1080/1548	*L*19(*b*2)	1012/1256
*L*10(*b*38)	1744/1162	*L*20(*b*30)	(1144−1182)/1544

[a]Each metric is a baseline-corrected ratio of absorbance at the numerator (in cm^{-1}) to absorbance at the denominator (in cm^{-1}). If a range is given, then the average absorbance over that range is used. The *b* labels are from the set of 64, while the *L* labels are for the new set of 20.

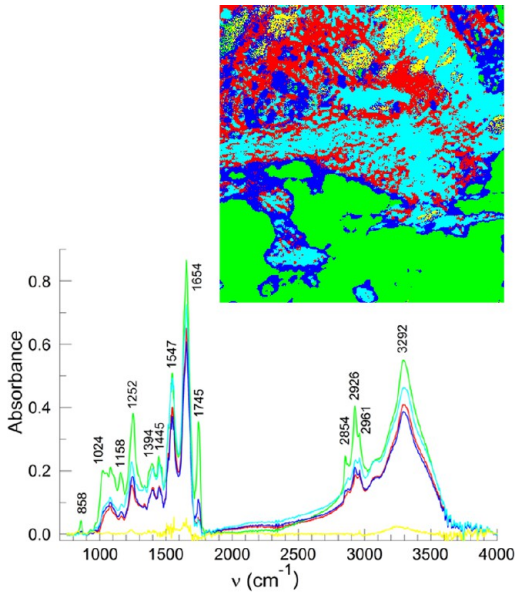


Figure 3. The k-means cluster analysis image with five groups using the *unscaled*, top-20 metrics (top). The nontumor portion is in green, and the holes are yellow. The tumor has red, blue, and cyan groups. The IR spectrum of each group is given at the bottom with the same color coding.

Figure 3 with the same color coding for groups. The IR spectra are normalized by the number of pixels in each group; therefore, the intensities are meaningfully compared. There are significant differences between the tumor and nontumor regions. Most notably, the peak at 1745 cm^{-1}, which is assigned to ester-linked fats, is much smaller in this tumor.

The metric scores for all groups except the holes are presented in Table 2, enabling comparisons between metrics and to other IR work. Many specific, chemical changes are archived in Table 2. As only one example, take the first new metric, *L*1, which is the ratio of the absorbance of ester-linked fat at 1744 cm^{-1} to that of the amide II band of protein at 1548 cm^{-1}. The value is 0.679 in the nontumor, but it falls to 0.288 in the blue tumor region that is closest to the nontumor part, to 0.132 in the red tumor group, and to 0.092 in the cyan tumor group. There is a dramatic reduction of ester-linked fat from the nontumor upon going into the tumor.

The amide I (1656 cm^{-1}) and II (1548 cm^{-1}) protein bands are two of the strongest bands involving protein backbone motions that have many inflections, indicating unresolved structure. There has been much IR work on the secondary structure of proteins[32−35] and, given the high signal-to-noise of the group spectra (they are averages of ∼30 000 individual spectra), second derivatives of the group IR spectra were determined for comparison upon going from tumor to nontumor. The Perkin-Elmer Spectrum program was used to get five-point, finite-difference, second derivatives that have been multiplied by a factor of −200 in order to be displayed on a scale similar to that of the absorption spectra as shown in Figure 4. Second derivatives have large negative-going features at the peak center, which has been reversed and scaled for better comparison to the absorption line shape. Notice how the inflections in the absorption spectra correspond to the peaks of the second derivatives multiplied by −200. There are 7 or 8 discrete peak centers discernible in the amide I band alone. Their peak centers are listed in Figure 4. The second derivative peaks at 1654 cm^{-1} are α-helix, while the peaks at 1637 and 1694 cm^{-1} (or possibly 1683 cm^{-1}) may be β-sheet. The other peaks likely correspond to other helices and various turns in standard structures. There is less α-helix in the tumor and more of the 1683 cm^{-1} amide I group in the tumor. The biggest changes seem to be in the amide II band. The second derivative peak at 1546.9 cm^{-1} is the largest in the nontumor region, while the two surrounding peaks at 1559.3 and 1537.9 cm^{-1} become larger in the tumor. Clearly, there are discernible changes between proteins in and out of the tumor.

It has been useful to make histograms of the IR metric values. In fact, all but one of the metrics (*L*14) show two resolved distributions, one for the tumor and one for the nontumor. It is even more useful to make two-dimensional histograms of one metric against another in order to see if they work together to give even more separation of the tumor and nontumor groups. Such a plot of *L*5 (baseline-corrected absorption at 1516 cm^{-1} divided by that at 1236 cm^{-1}) versus *L*1 (baseline-corrected absorption at 1744 cm^{-1} divided by that at 1548 cm^{-1}) is shown in Figure 5a. Clearly, there are two separate distributions, and the separation is greater than either metric by itself (as is evident by the projections that reveal the one-dimensional histograms of each metric). A plot of a different set of metrics, *L*14 versus *L*1, is presented in Figure 5b. A third distribution is apparent, but there are no groups that identify this region.

One can address this issue by doing the k-means analysis with more groups, but this presents challenges to graphical visualization. A k-means analysis was performed with 25 groups, and the results are given in Figure 6. The IR spectra of all 25 groups are given in the Supporting Information. A matrix of the eq 2 distances between each of the 25 groups was computed, and a hierarchical dendrogram was constructed, which supported classification of the groups into categories of nontumor (green), tumor (purple), and holes (black or dark blue). The images are plotted with shades of green for the nontumor, shades of purple for the tumor, shades of dark blue for the holes, and yellow for the nontumor group closest to the tumor. This exercise emphasizes that it is relatively easy to discern tumor from nontumor. Therefore, some of the most important work ahead lies in trying to extract differences between the groups within a classification.

Noting that groups 3 and 13 are near each other and are neighbors in the dendrogram, the IR spectrum of group 13

dx.doi.org/10.1021/jp4073087 | J. Phys. Chem. B 2013, 117, 12442−12450

Table 2. IR Metric Values for Five Groups from the K-Means Analysis with the Best Unscaled Metrics[a]

IR metric	nontumor (green)	tumor 1 (red)	tumor 2 (blue)	tumor 3 (cyan)
*L*1 (1744/1548)	0.679	0.132	0.288	0.092
*L*2 (1744/1244)	0.986	0.375	0.616	0.216
*L*3(1742/1256)	0.904	0.425	0.643	0.258
*L*4 (1160/1548)	0.353	0.107	0.155	0.167
*L*5 (1516/1236)	1.025	1.741	1.416	1.505
*L*6 (2916/1548)	0.625	0.284	0.373	0.322
*L*7(1120/1020)	0.755	1.816	1.088	1.231
*L*8(2924/1544)	0.786	0.357	0.471	0.391
*L*9 (1080/1548)	0.383	0.166	0.223	0.204
*L*10 (1744/1162)	1.951	1.166	1.858	0.566
*L*11(1016/1080)	0.892	0.255	0.462	0.488
*L*12 (1252/1544)	0.771	0.355	0.485	0.412
*L*13 (1024/1080)	0.962	0.440	0.588	0.631
*L*14 (1080/1244)	0.557	0.450	0.471	0.466
*L*15 (1516/1582)	1.277	1.120	1.165	1.241
*L*16 (1744/1162)	0.372	0.110	0.163	0.172
*L*17 (1080/3290)	0.359	0.179	0.231	0.235
*L*18 (1556/1548)	0.832	0.896	0.878	0.907
*L*19 (1012/1256)	0.433	0.110	0.216	0.251
*L*20 [(1144−1182)/1544]	0.325	0.104	0.150	0.166
pixels	36 476	23 753	21 075	33 143

[a]The associated IR bands (in cm^{-1}) are given as a ratio in parentheses for each metric. If a range is given, then the metric involves an average over that range.

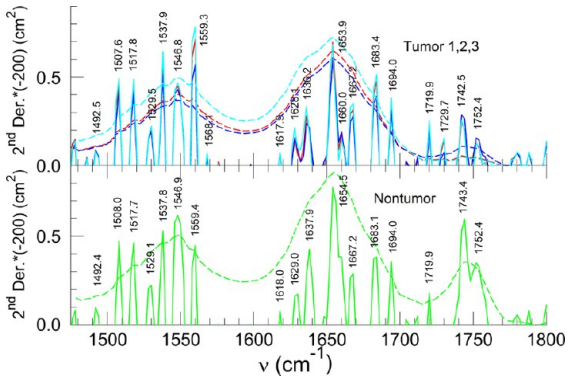


Figure 4. Second derivatives of the IR spectra for each group in Figure 3 (excluding holes) with the same color coding. The dotted traces are absorption spectra, while the solid traces are second derivatives multiplied by −200 in order to put the derivatives on a scale similar to the absorption spectra. The bottom (green) is nontumor, and the top has the three tumor groups.

multiplied by 0.963 was subtracted from the IR spectrum of group 3 in order to cancel the protein (by the amide I band) as best as possible. This leaves a remarkably well resolved IR spectrum of the lipid from the nontumor region, as shown in Figure 7. The nontumor lipid peaks include 3002 (shoulder), 2959 (shoulder), 2925, 2854, 2249, 1746, 1469, 1443, 1369, 1255, 1157, 1107, 1083, 928, and 859 cm^{-1}. The regions from 1500 to 1700 cm^{-1} and above 3100 cm^{-1} are dominated by changes in the protein and are avoided for this discussion about the lipid. Also shown (in red) is a modeled spectrum using FDM's Very Large Bundle (www.fdmspectra.com) of IR libraries based on a weighted sum of 0.6 times glyceryl triacetylricinoleate and 0.4 times ethyl cyanopolyacrylate (where the max peak in each was scaled to one). A Euclidean IR search was performed on the nontumor, ester-linked lipid difference spectrum (black trace in Figure 7) using the Perkin-Elmer Spectrum Search program and FDM's Very Large Bundle of IR libraries. The two top matches were glyceryl triacetylricinoleate and ethyl cyanopolyacrylate. Glyceryl triacetylricinoleate is commonly used in cosmetics. It is a triglyceride with C18:1 chains (18 carbons, 1 double bond at C9), each containing one double bond and an extra ester linkage per fat chain (occurring below the double bond toward the free end of the fat chain at C12), which is somewhat different than textbook triglycerides. This species is an excellent match except for the peaks at 858 and 2248 cm^{-1}, for which cyanopolyacrylate (superglue) does better. The superglue polymer does not match the fat chains at all, but it is a good match in the fingerprint region, and it does get the peaks at 858 and 2248 cm^{-1}. Taken together, this suggests that the average of the nontumor liver lipids (after freezing, slicing, melting, IR exposure, and hours of time) is similar to a glyceryl triacetyl ricinoleate, where polyacrylate polymerization has been occurring perhaps between different fat chains. The average nontumor liver lipid is a triglyceride with roughly twice the ester groups of textbook triglycerides, a prominent backbone vibration at 858 cm^{-1} (not due to fat chains), and some C≡N triple bond groups (less than one per fat chain). Note that the triple bond observation at 2248 cm^{-1} could also be a C≡C, and the only evidence at this point for C≡N is the match to cyanopolyacrylate. Explanations for triple bonds in nontumor, human liver lipid need to consider that the tissue is not fixed. Possible explanations include (1) enzymes that destroy triple bonds are no longer functional, (2) triple bonds exist as intermediates in normal lipid degradation, (3) drug treatments of patients with tumors may damage the liver, and (4) triple bonds are a chemical artifact of mixing cell parts as tissue warms after slicing. There are, after all, good reasons why pathologists fix tissues extracting lipids. Further work is necessary to say more. Despite this, there seems to be potential for analysis of liver lipids by these methods, although caution must be

dx.doi.org/10.1021/jp4073087 | J. Phys. Chem. B 2013, 117, 12442−12450

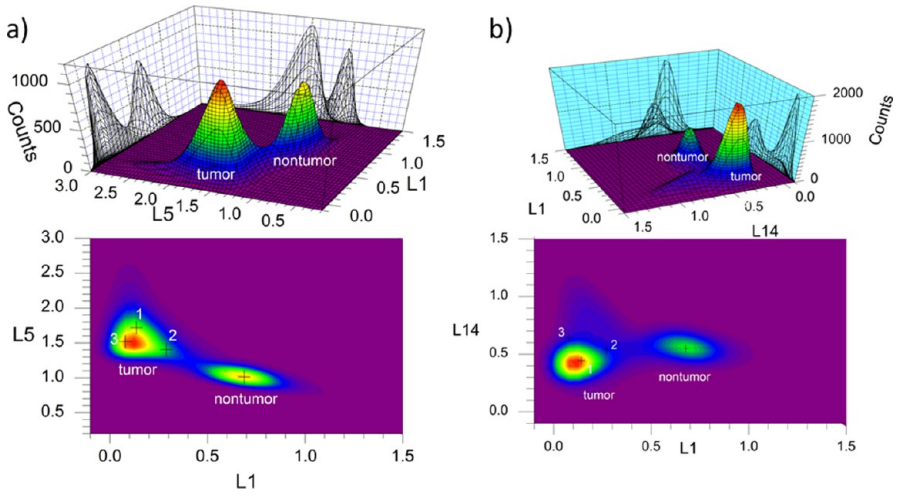


Figure 5. (a) Two-dimensional histogram (top) of IR metrics *L*5 versus *L*1. *L*1 is a ratio of ester-linked fat to protein, while *L*5 is the side of the amide II protein band to phosphate. A contour diagram of the same information is given at the bottom. The centroid values for the k-means groups are given with cross symbols identifying the distributions. There is greater separation of the tumor and nontumor values with two metrics than with either metric by itself (see the projections in the top part). (b) Two-dimensional histogram (top) of the values of IR metric *L*14 versus *L*1. *L*1 is absorbance at 1744 cm^{-1} divided by that at 1458 cm^{-1}, while *L*14 is absorbance at 1080 cm^{-1} divided by that at 1244 cm^{-1}. A contour diagram of the same information is given at the bottom. A third distribution is evident that was not identified by the five-group k-means analysis, suggesting that more k-means groups would be worthwhile.

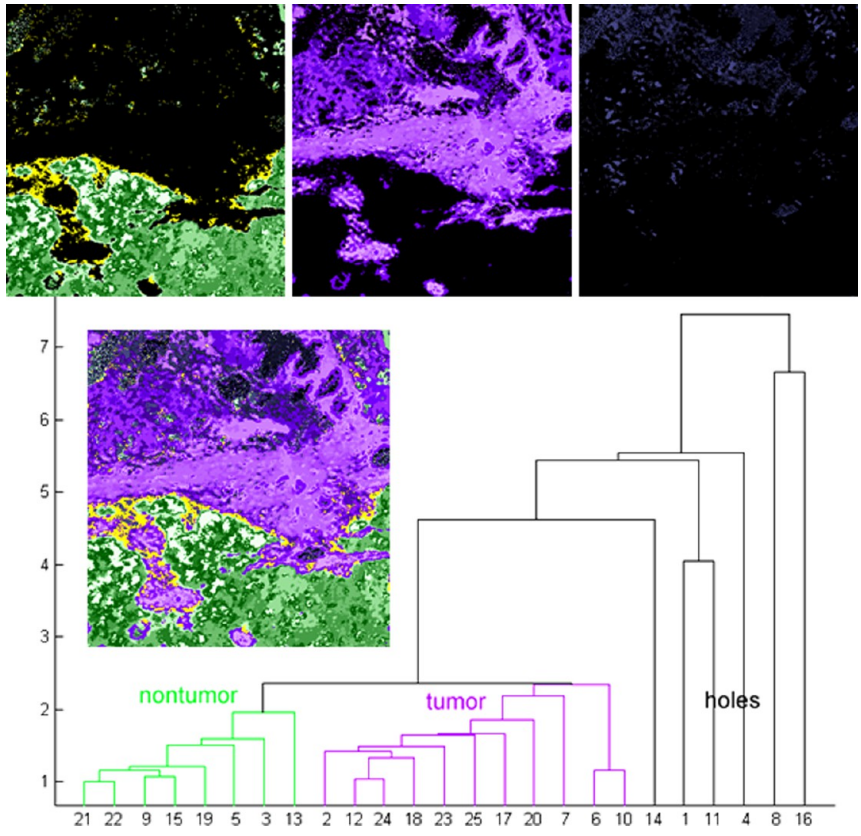


Figure 6. A 25-group k-means analysis with nontumor groups in shades of green (top left), tumor groups in shades of purple (top middle), and holes in dark blue–gray (top right). A dendrogram of the distances between groups (bottom) dictated the color scheme, and the full image is given as an inset to the dendrogram. Yellow indicates the group closest to the interface between the tumor and nontumor.

exercised in extrapolating to the lipid conditions in the living liver. We believe that this merits further study.

With 25 k-means groups, it was informative to re-examine the *L*14 versus *L*1 metric–metric plot in order identify the new distribution from Figure 5b that was not identified by any groups of the five-group analysis. The *L*14 metric is the ratio of absorbance at 1080–1244 cm^{-1} (glycogen and/or lipid to phosphate), and the *L*1 is that at 1744–1548 cm^{-1} (ester-linked lipid to protein). A plot of *L*14 versus *L*1 is given in Figure 8b, with white numbers indicating the centroid values

dx.doi.org/10.1021/jp4073087 | J. Phys. Chem. B 2013, 117, 12442–12450

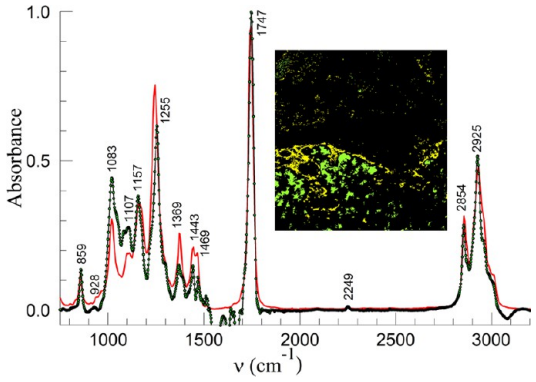


Figure 7. IR difference spectrum of the lipid in the nontumor region (black trace, green dots) obtained by subtracting 0.963 times the IR spectrum of group 13 (yellow) from the IR spectrum of group 3 (light green), where the colors refer to the inset image. The spectrum was arbitrarily scaled to make the max peak equal to one for comparison to IR search results. The red curve is a composite from the FDM library with 60% glyceryl triacetylricinoleate and 40% ethyl cyanopolyacrylate.

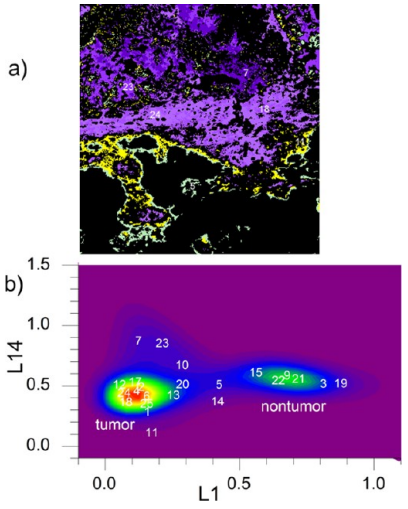


Figure 8. (a) Color-coded image of tumor groups 7 and 23 (darker shades of purple) that comprise inner parts of the tumor, groups 24 and 18, which have fibrous stroma cells (lighter shades of purple). Groups 13 (yellow) and 5 (light green) help to indicate the margin. (b) Contour diagram of the *L*14 IR metric versus *L*1 IR metric. A third distribution is identified for groups 7 and 23, which was not identified by a five-group k-means analysis.

for these metrics from the 25 groups. Clearly, groups 7 and 23 (with *L*1, *L*14 scores of 0.107, 0.868 and 0.197, 0.845, respectively) now identify the third distribution and are imaged with dark purple colors in Figure 8a. These are groups further within the tumor. The lighter purple regions correspond to groups 24 and 18 (with *L*1, *L*14 scores of 0.071, 0.413 and 0.088, 0.367, respectively). By comparison to the standard H&E stains (Figure 1a inset), we believe that there is a concentration of fibrous stroma cells (high protein, low lipid) in these light purple, tumor regions (see Figure 8a) that have H&E staining more like that of the nontumor region. In Figure 8b, it is also interesting that groups 5 and 13 occur in the region between the nontumor and tumor distributions. These groups have been colored light green (5) and yellow (13), as isolated in Figure 8a. We assert that these groups help to define the tumor's margin.

The IR spectra of selected groups are shown in Figure 9. Each group is normalized to the number of image pixels in the group; therefore, intensities are meaningful. There are significant differences in all of the selected groups, which have been identified with a color-coded image inset in Figure 9. As mentioned above, there are many more subtle differences in the IR spectra of these various groups. Therefore, the IR spectra of all 25 groups (and identifying bitmaps of each group) have been provided in the Supporting Information.

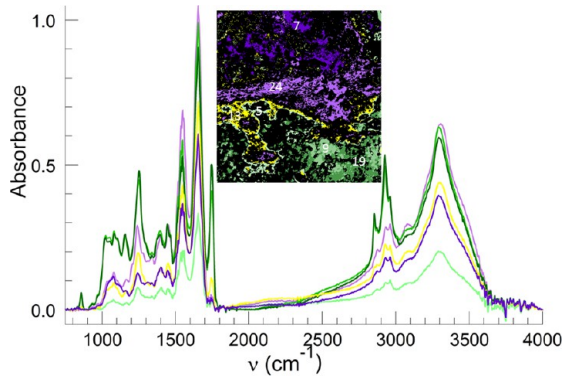


Figure 9. IR spectra of selected groups from the 25-group k-means analysis. The dark (19) and medium green (9) groups are nontumor, the yellow (13) and light green (5) groups are near the margin, the light purple (24) indicates a region with fibrous stroma cells, and purple (7) indicates an inner tumor region.

DISCUSSION

A set of literature IR metrics were augmented by various lipid bands and evaluated giving a set of 20 that are good for detecting fixation-free liver tumors (see Table 1). A scaling of the metrics was introduced to make a level comparison. Only 6 out of 20 were from the previous literature work (on fixed prostate tissue), showing that the molecules removed by fixation are quite important. An ester-linked lipid provides for good tumor distinction, at least in this patient. The top IR metric, *L*1, is the ratio of absorbance at 1744 cm^{-1} to absorbance at 1548 cm^{-1}, that is, a ratio of ester-linked lipid to protein. It was used as the ordinate in several IR metric versus IR metric plots (Figures 5 and 8) that show good distinction between tumor and nontumor. While this work concentrated on distinguishing a tumor in one patient, work on differentiating cell types by the same methods is promising and continues. Work also continues on gathering data from other patients. It is interesting that the ester-linked lipids vary greatly between patients. The lipid-related peaks at 2248 (likely a triple bond) and 858 cm^{-1} are very different than what is seen with IR spectra of fixed liver tissues. So far, these peaks have been seen in two other patients. Future work on more patients will provide a better perspective on the importance of this result.

Diagnostic tools have been presented that could one day be utilized within the operating room, providing vital real-time information on how the surgeon should proceed and for verifying completeness of surgical resection. Noting that fixation is a large perturbation from the tissue of living patients, the freezing and slicing of the current work is a lesser perturbation and a step closer to the ultimate goal of in situ IR spectral evaluation of the liver or on freshly excised liver tissue from biopsy or resectional specimens. The sensitivity of IR spectra to liver lipids suggests utility beyond work on tumors of

dx.doi.org/10.1021/jp4073087 | J. Phys. Chem. B 2013, 117, 12442−12450

the liver. Steatosis or fatty liver disease may be readily assessed with the methods presented herein. Knowledge about the condition of the liver might be important in a variety of health care decisions beyond liver tumors.

■ ASSOCIATED CONTENT

Supporting Information

A table of the original 64 IR metrics, an image of a 5-group k-means analysis using scaled IR metrics, and bitmap images of the groups from a 25-group k-means cluster analysis. Also included are the IR spectra of the groups from a 25-group k-means analysis in an Excel file. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: coe.1@osu.edu (J.V.C.).
*E-mail: allen.697@osu.edu (H.C.A.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the National Cancer Institute and the National Institutes for Health for Grant NIH R21 CA167403.

■ REFERENCES

(1) Wu, J. G.; Xu, Y. Z.; Sun, C. W.; Soloway, R. D.; Xu, D. F.; Wu, Q. G.; Sun, K. H.; Weng, S. F.; Xu, G. X. Distinguishing Malignant from Normal Oral TissueS Using FTIR Fiber-Optic Techniques. *Biopolymers* **2001**, *62*, 185−192.

(2) Diem, M.; Miljkovic, M.; Bird, B.; Chernenko, T.; Schubert, J.; Marcsisin, E.; Mazur, A.; Kingston, E.; Zuser, E.; Papamarkakis, K.; Laver, N. Applications of Infrared and Raman Microspectroscopy of Cells and Tissue in Medical Diagnostics: Present Status and Future Promises. *Spectrosc. Int. J.* **2012**, *27*, 463−496.

(3) Harrington, J. A. *Infrared Fibers and Their Applications*; SPIE Press: Bellingham, WA, 2004.

(4) Prophet, E. B. *Laboratory Methods in Histotechnology*; American Registry of Pathology: Washington, DC, 1992.

(5) Lloyd, S. P. Least Squares Quantization in PCM. *IEEE Trans. Inf. Theory* **1982**, *28*, 129−137.

(6) Kanungo, T.; Netanyahu, N. S.; Wu, A. Y. An Efficient k-Means Clustering Algorithm: Analysis and Implementation. *IEEE Trans. Pattern Anal. Machine Intelligence* **2002**, *24*, 881−892.

(7) Seber, G. A. F. *Multivariate Observations*; John Wiley & Sons: Hoboken, NJ, 1984.

(8) Coates, V. J.; Offner, A.; Siegler, E. H. Design and Performance of an Infrared Microscrope Attachment. *J. Opt. Soc. Am.* **1953**, *43*, 984−989.

(9) Petibois, C.; Deleris, G. Chemical Mapping of Tumor Progression by FT-IR Imaging: Towards Molecular Histopathology. *Trends Biotechnol.* **2006**, *24*, 455−462.

(10) Kohler, A.; Bertrand, D.; Martens, H.; Hannesson, K.; Kirschner, C.; Ofstad, R. Multivariate Image Analysis of a Set of FTIR Microspectroscopy Images of Aged Bovine Muscle Tissue Combining Image and Design Information. *Anal. Bioanal. Chem.* **2007**, *389*, 1143−1153.

(11) Noreen, R.; Moenner, M.; Hwu, Y.; Petibois, C. FTIR Spectro-Imaging of Collagens for Characterization and Grading of Gliomas. *Biotechnol. Adv.* **2012**, *30*, 1432−1446.

(12) Yang, T. T.; Weng, S. F.; Zheng, N.; Pan, Q. H.; Cao, H. L.; Liu, L.; Zhang, H. D.; Mu, D. W. Histopathology Mapping of Biochemical Changes in Myocardial Infarction by Fourier Transform Infrared Spectral Imaging. *Forensic Sci. Int.* **2011**, *207*, E34−E39.

(13) Diem, M.; Papamarkakis, K.; Schubert, J.; Bird, B.; Romeo, M. J.; Miljkovic, M. The Infrared Spectral Signatures of Disease: Extracting the Distinguishing Spectral Features Between Normal and Diseased States. *Appl. Spectrosc.* **2009**, *63*, 307A−318A.

(14) Maziak, D. E.; Do, M. T.; Shamji, F. M.; Sundaresan, S. R.; Perkins, D. G.; Wong, P. T. T. Fourier-Transform Infrared Spectroscopic Study of Characteristic Molecular Structure in Cancer Cells of Esophagus: An Exploratory Study. *Cancer Detect. Prev.* **2007**, *31*, 244−253.

(15) Bird, B.; Miljkovic, M.; Remiszewski, S.; Akalin, A.; Kon, M.; Diem, M. Infrared Spectral Histopathology (SHP): A Novel Diagnostic Tool for the Accurate Classification of Lung Cancer. *Lab. Invest.* **2012**, *92*, 1358−1373.

(16) Schubert, J. M.; Bird, B.; Papamarkakis, K.; Miljkovic, M.; Bedrossian, K.; Laver, N.; Diem, M. Spectral Cytopathology of Cervical Samples: Detecting Cellular Abnormalities in Cytologically Normal Cells. *Lab. Invest.* **2010**, *90*, 1068−1077.

(17) Wood, B. R.; Chiriboga, L.; Yee, H.; Quinn, M. A.; McNaughton, D.; Diem, M. Fourier Transform Infrared (FTIR) Spectral Mapping of the Cervical Transformation Zone, And Dysplastic Squamous Epithelium. *Gynecol. Oncol.* **2004**, *93*, 59−68.

(18) Romeo, M.; Matthäus, C.; Miljkovic, M.; Diem, M. Infrared Microspectroscopy of Individual Human Cervical Cancer (HeLa) Cells. *Biopolymers* **2004**, *74*, 168−171.

(19) Pallua, J. D.; Pezzei, C.; Zelger, B.; Schaefer, G.; Bittner, L. K.; Huck-Pezzei, V. A.; Schoenbichler, S. A.; Hahn, H.; Kloss-Brandstaetter, A.; Kloss, F.; Bonn, G. K.; Huck, C. W. Fourier Transform Infrared Imaging Analysis in Discrimination Studies of Squamous Cell Carcinoma. *Analyst* **2012**, *137*, 3965−3974.

(20) Murakata, L. A.; Lewin-Smith, M. R.; Specht, C. S.; Kalasinsky, V. F.; McEvoy, P. L.; Vinh, T. N.; Rabin, L. N.; Mullick, F. G. Characterization of Acrylic Polyamide Plastic Embolization Particles in Vitro and in Human Tissue Sections by Light Microscopy, Infrared Microspectroscopy and Scanning Electron Microscopy with Energy Dispersive X-ray Analysis. *Mod. Pathol.* **2006**, *19*, 922−930.

(21) Hara, A.; Taketomi, T. Isolation and Determination of Cholesterol Glucuronide in Human Liver. *Lipids* **1982**, *17*, 515−518.

(22) Diem, M.; Chiriboga, L.; Lee, H. Infrared Spectroscopy of Human Cells and Tissue. VIII. Strategies for Analysis of Infrared Tissue Mapping Data and Applications to Liver Tissue. *Biopolymers* **2000**, *57*, 282−290.

(23) Chiriboga, L.; Yee, M.; Diem, M. Infrared Spectroscopy of Human Cells and Tissue. Part VI: A Comparative Study of Histopathology and Infrared Microspectroscopy of Normal, Cirrhotic, And Cancerous Liver Tissue. *Appl. Spectrosc.* **2000**, *54*, 1−8.

(24) Chiriboga, L.; Lee, H.; Diem, M. Infrared Spectroscopy of Human Cells and Tissue. Part VII: FT-IR Microspectroscopy of DNase- and RNase-Treated Normal, Cirrhotic, And Neoplastic Liver Tissue. *Appl. Spectrosc.* **2000**, *54*, 480−485.

(25) Susie, B.-W.; Tamara, G.; Sandra, H.; Jill, B.; Max, D. Infrared Spectroscopy of Human Tissue. V. Infrared Spectroscopic Studies of Myeloid Leukemia (ML-1) Cells at Different Phases of the Cell Cycle. *Biospectroscopy* **1999**, *5*, 219−227.

(26) Diem, M.; Boydston-White, S.; Chiriboga, L. Infrared Spectroscopy of Cells and Tissues: Shining Light onto a Novel Subject. *Appl. Spectrosc.* **1999**, *53*, 148A−161A.

(27) Wong, P. T. T.; Cadrin, M.; French, S. W. Distinctive Infrared Spectral Features in Liver-Tumor Tissue of Mice — Evidence of Structural Modifications at the Molecular Level. *Exp. Mol. Pathol.* **1991**, *55*, 269−284.

(28) Fernandez, D. C.; Bhargava, R.; Hewitt, S. M.; Levin, I. W. Infrared Spectroscopic Imaging for Histopathologic Recognition. *Nat. Biotechnol.* **2005**, *23*, 469.

(29) Walsh, M. J.; Reddy, R. K.; Bhargava, R. Label-Free Biomedical Imaging With Mid-IR Spectroscopy. *IEEE J. Sel. Top. Quantum Electron.* **2012**, *18*, 1502−1513.

(30) Fernandez, D. C.; Bhargava, R.; Hewitt, S. M.; Levin, I. W. Infrared Spectroscopic Imaging for Histopathologic Recognition. *Nat. Biotechnol.* **2005**, *23*, 469−474.

(31) Chen, Z. *Imaging Infrared Microscope Analysis of Fixation-Free Liver Tumor Tissue*. M.S. Thesis, The Ohio State University, Columbus, OH, 2013.
(32) Tatulian, S. A. Structural Characterization of Membrane Proteins and Peptides by FTIR and ATR-FTIR Spectroscopy. In *Lipid-Protein Interactions: Methods and Protocol*; Kleinschmidt, J. H., Ed.; Springer Science+Business Media: New York, 2013; Vol. *974*.
(33) Byler, D. M.; Susi, H. Examination of the Secondary Structure of Proteins by Deconvolved FTIR Spectra. *Biopolymers* **1986**, *25*, 469−487.
(34) Dong, A.; Huang, P.; Caughey, W. S. Protein Secondary Structures in Water from Second-Derivative Amide I Infrared Spectra. *Biochemistry* **1990**, *29*, 3303−3308.
(35) Barth, A. Infrared Spectroscopy of Proteins. *Biochim. Biophys. Acta* **2007**, *1767*, 1073−1101.