Detection of endogenous biomolecules in Barrett's esophagus by Fourier transform infrared spectroscopy

Thomas D. Wang*†‡, George Triadafilopoulos*†, James M. Crawford§, Lisa R. Dixon§, Tarun Bhandari*†, Peyman Sahbaie*†, Shai Friedland*†, Roy Soetikno*†, and Christopher H. Contag¶

*Division of Gastroenterology, Stanford University School of Medicine, †Gastroenterology Section, Palo Alto Veterans Affairs Health Care System, and ¶Department of Pediatrics, Radiology, and Immunology, 3801 Miranda Avenue, Building 100, Palo Alto, CA 94304; and §Department of Pathology, Immunology, and Laboratory Medicine, University of Florida College of Medicine, P.O. Box 100275, 1600 Southwest Archer Road, MSB 649, Gainesville, FL 32610-0275

Communicated by Britton Chance, University of Pennsylvania School of Medicine, Philadelphia, PA, August 13, 2007 (received for review June 8, 2006)

Fourier transform infrared (FTIR) spectroscopy provides a unique molecular fingerprint of tissue from endogenous sources of light absorption; however, specific molecular components of the overall FTIR signature of precancer have not been characterized. In attenuated total reflectance mode, infrared light penetrates only a few microns of the tissue surface, and the influence of water on the spectra can be minimized, allowing for the analyses of the molecular composition of tissues. Here, spectra were collected from 98 excised specimens of the distal esophagus, including 38 squamous, 38 intestinal metaplasia (Barrett's), and 22 gastric, obtained endoscopically from 32 patients. We show that DNA, protein, glycogen, and glycoprotein comprise the principal sources of infrared absorption in the 950- to 1,800-cm^{-1} regime. The concentrations of these biomolecules can be quantified by using a partial least-squares fit and used to classify disease states with high sensitivity, specificity, and accuracy. Moreover, use of FTIR to detect premalignant (dysplastic) mucosa results in a sensitivity, specificity, positive predictive value, and total accuracy of 92%, 80%, 92%, and 89%, respectively, and leads to a better interobserver agreement between two gastrointestinal pathologists for dysplasia (κ = 0.72) versus histology alone (κ = 0.52). Here, we demonstrate that the concentration of specific biomolecules can be determined from the FTIR spectra collected in attenuated total reflectance mode and can be used for predicting the underlying histopathology, which will contribute to the early detection and rapid staging of many diseases.

dysplasia | pathology | optical biopsy | diagnostics

Transformation to cancer is characterized by molecular changes that correlate with the overall pathology, and measurements of concentrations of specific molecules within tissues can provide diagnostic and prognostic information for aiding in early detection and outcome measures for therapy. Optical methods of disease detection offer the advantages of multiparametric measures of intrinsic molecular changes without need for contrast agents or ionizing radiation. These properties of optical assessment of tissue composition and disease states will have their greatest benefit in diseases where there are predisposing conditions to identify high-risk patients and where the target tissues are accessible for direct evaluation. Adenocarcinoma of the esophagus is the cancer with the greatest recent increase in incidence in the U.S. (1). This tumor arises primarily from Barrett's esophagus, a predisposing metaplastic transformation of the esophageal mucosa that results from chronic acid reflux. Barrett's esophagus is present in 15% or more patients undergoing endoscopy for symptoms of acid reflux (2). Regular endoscopic screening is recommended to identify the presence of dysplasia, a premalignant condition that predicts future malignancy (3), and this offers opportunities for early detection.

However, conventional histopathological evaluation of dysplasia is limited by lack of agreement on diagnostic criteria and human interpretive variability and, hence, the significant inter- and intraobserver variability for the endoscopist and pathologist (4). For example, among experienced gastrointestinal pathologists, the interobserver discordance for low-grade dysplasia is >50%; fortunately, disagreement is only ≈15% for diagnosis of high-grade dysplasia, and improvements in diagnosis are not necessary (5). Additional markers for dysplasia, such as the presence of aneuploidy on flow cytometry, have been used to help predict patient risk for developing adenocarcinoma (6). However, this laboratory assay is prone to sampling error because dysplasia in Barrett's is usually flat, patchy, and endoscopically invisible (7). Unfortunately, any misinterpretation in the presence of dysplasia may lead to catastrophic consequences for the patient, namely, a false-positive identification may lead to an unnecessary esophagectomy, whereas a false-negative histological interpretation may allow a dysplastic lesion to progress to cancer. Thus, a method that can distinguish among tissues in the distal esophagus and between dysplastic and nondysplastic Barrett's esophagus may ultimately be used to guide tissue biopsy, thus increasing the yield of dysplasia detection.

Here, we present the use of Fourier transform infrared (FTIR) spectroscopy in attenuated total reflectance (ATR) mode as a method of performing quantitative analysis of tissue composition and demonstrate its applicability in the analysis of esophageal specimens. Infrared spectroscopy is a powerful tool for studying the molecular structure and composition of cells and tissues without need for exogenous dyes or contrast agents (8, 9). Tissue absorbs infrared light at characteristic wavelengths determined by the vibrational motions of covalently bonded atoms. These vibrational modes include stretching and bending that can occur in either a symmetric or asymmetric fashion (10, 11). FTIR is a fast, efficient technique for collecting an infrared absorption spectrum over a broad spectral band with high sensitivity (12), and this technique has been used to identify neoplasia in breast (13), cervix (14), colon (15), lung (16), stomach (17), and pancreas (18). However, FTIR can be limited by the predominance of the water-absorption spectrum, and therefore, in this study, spectra are collected in ATR mode, which limits the influence of the water band. In ATR mode, an infrared beam passes through an ATR crystal at angles below critical θ_c and undergoes total internal reflection, shown in Fig. 1. An evanes-

Author contributions: T.D.W., G.T., and J.M.C. designed research; T.D.W., G.T., T.B., P.S., S.F., and R.S. performed research; T.D.W., G.T., J.M.C., L.R.D., and C.H.C. analyzed data; and T.D.W. and J.M.C. wrote the paper.

The authors declare no conflict of interest.

Abbreviations: FTIR, Fourier transform infrared; ATR, attenuated total reflectance.

‡To whom correspondence should be sent at the present address: Division of Gastroenterology, University of Michigan Medical School, 109 Zina Pitcher Place, BSRB 1522, Ann Arbor, MI 48109-2200. E-mail: thomaswa@umich.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0707567104/DC1.

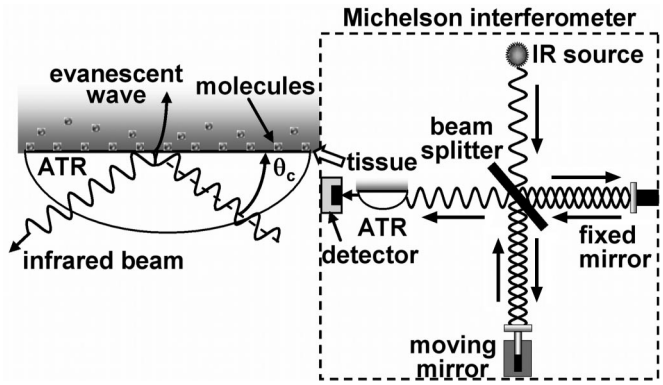

Fig. 1. FTIR spectra are collected from a biopsy specimen placed in contact with the ATR crystal. An infrared (IR) beam from a Michelson interferometer is incident on the tissue–crystal interface at angles below critical θ_c and undergoes total internal reflection. An evanescent wave is created within a thin layer of tissue near the surface and is absorbed by the endogenous molecules present. In the interferometer, a moving mirror varies the length of one optical path relative to the other and creates an interferogram that is converted to an absorbance spectrum by a Fourier transform.

cent wave is generated within the surface of the mucosa (2–3 μm depth) whose energy can be absorbed by endogenous molecules. With ATR, the path length of evanescent light travel in tissue is determined by θ_c and the relative index of refraction between the tissue and crystal (19) and thus is relatively constant among specimens. In the Michelson interferometer, a moving mirror varies the length of one optical path relative to the other and creates an interferogram that is converted to the absorbance spectrum by a Fourier transform. The use of ATR should enable rapid and quantitative assessment of the molecular composition of tissues as an indication of altered physiology that is characteristic of the disease state and linked to the histopathology that is used to make diagnosis. Therefore, we sought to determine the primary molecular components of the FTIR spectra in cancer and correlate quantitative changes in these molecules with the pathology of early dysplasia.

Results

FTIR Spectra of Esophagus. FTIR spectra collected in ATR mode were obtained from 98 excised specimens acquired endoscopically from the distal esophagus. Two gastrointestinal pathologists independently identified each sample as squamous (n = 38), Barrett's (n = 38), or gastric mucosa (n = 22) with complete agreement. Squamous and gastric mucosa are considered normal findings. Screening biopsies, in particular for short-segment (<3 cm) Barrett's, often include gastric tissue that is not identified until the histology is evaluated. Further subclassification of the Barrett's mucosa resulted in pathologist nos. 1 (J.M.C.) and 2 (L.R.D.) identifying 15 and 18 specimens, respectively, as dysplasia, and 23 and 20 specimens, respectively, as nondysplasia. All dysplastic specimens were subclassified as low grade by both pathologists with the exception of one that was initially read as high grade by one pathologist but was later resolved as low grade by consensus. Each specimen was then ranked according to the area-under-the-spectra (AUS) of the measured spectra, by group. The average AUS for squamous, Barrett's, and gastric tissues were found to be 34 ± 10 (range 13.8–55.1), 21 ± 8 (range 10.9–40.9), and 31 ± 12 (range 7.9–47.0) in arbitrary units (A.U.), respectively. Individual spectra with AUS in the 90th, 50th, and 10th percentile for squamous [supporting information (SI) Fig. 6*A*, Barrett's (SI Fig. 6*B*), and gastric (SI Fig. 6*C*) mucosa are shown to illustrate the biovariability observed in the intensity and location of absorbance bands (SI Fig. 6). Next, the spectral peaks are identified and averaged according to histological group, shown in Table 1. To interpret these patterns, FTIR spectra were collected from pure DNA, glycogen, glycoprotein, and protein in the form of dried film, shown in Fig. 2*A*, and band assignments were made by comparison. Water is a significant source of infrared light absorption in this regime, and its spectrum is shown as well.

The infrared absorption of squamous mucosa exhibits DNA peaks at average wave numbers of 970 and 1,051 cm^{-1} (C—O stretch vibration of deoxyribose) and at 1,081 cm^{-1} (symmetric PO_2^- stretch) (20). The lower wave-number regimes of some squamous specimens show prominent glycogen bands at 1,026, 1,081, and 1,154 cm^{-1} (C—O—H bend, C—C and C—O stretch, respectively) (21). For all squamous mucosa, the bands in the upper wave-number regime are produced by molecular bonds in proteins, including 1,240 and 1,297 cm^{-1} (coupled C—N stretch to N—H bend of amide III), 1,398 and 1,454 cm^{-1} (C—H bend in aliphatic side groups of the amino acid residues), 1,544 cm^{-1} (C—N stretch and N—H bend of amide II) and 1,640 cm^{-1} (C═O stretch of amide I) (22). In Barrett's mucosa, DNA peaks are seen at average values of 969, 1,051, and 1,081 cm^{-1} in the lower half of the spectrum, a mean glycoprotein band at 1,164 cm^{-1} (C—C stretch), (23) and little spectral evidence of glycogen. Similar to squamous, the upper wave-number regime of Barrett's mucosa exhibits peaks from protein that were found at average wave numbers nearly identical to that of squamous tissues. In gastric mucosa, the lower wave-number regime reveals DNA peaks at average values of 970, 1,049, and 1,080 cm^{-1} and glycoprotein bands at mean wave numbers of 1,117 cm^{-1} (O—H

Table 1. The average infrared absorption peaks for squamous (n = 38), Barrett's (n = 38), and gastric (n = 22) mucosa along with vibrational band assignments and primary sources of absorption.

Squamous, cm^{-1}	Barrett's, cm^{-1}	Gastric, cm^{-1}	Vibrational band	Primary source
970 ± 2	969 ± 2	970 ± 2	C—O stretch, C—C stretch	DNA
1,026 ± 2	–	–	C—O—H bend	Glycogen
1,051 ± 5	1,051 ± 3	1,049 ± 7	C—O stretch	DNA
1,081 ± 1	1,081 ± 1	1,080 ± 2	PO_2^- stretch, C—C stretch	DNA glycogen
–	–	1,117 ± 1	O—H bend	Glycoprotein
1,154 ± 2	–	–	C—O stretch	Glycogen
	1,164 ± 5	1,162 ± 4	C—C stretch	Glycoprotein
1,240 ± 2	1,240 ± 2	1,238 ± 2	C—N stretch, N—H bend	Protein (amide III)
1,297 ± 3	1,296 ± 9	1,301 ± 5	C—H stretch, N—H bend	Protein (amide III)
1,398 ± 1	1,399 ± 1	1,398 ± 1	C—H bend	Protein
1,454 ± 1	1,455 ± 1	1,453 ± 2	C—H bend	Protein
1,544 ± 1	1,545 ± 1	1,544 ± 2	C—N stretch, N—H bend	Protein (amide II)
1,640 ± 2	1,638 ± 2	1,637 ± 2	C═O stretch	Protein (amide I)

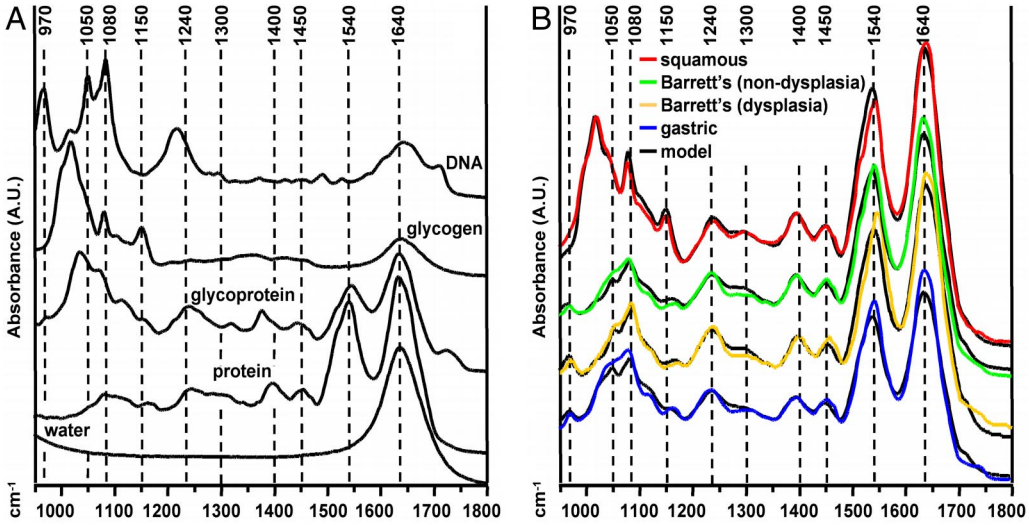

Fig. 2. Basis of FTIR spectra. (*A*) FTIR spectra acquired in ATR mode from dehydrated films of pure DNA, glycogen, glycoprotein, and protein provide the tissue band assignments. Water is a major source of infrared light absorption in this regime. (*B*) Comparison of FTIR model results (black) to measured spectra of representative specimens of squamous (red), nondysplastic Barrett's (green), dysplastic Barrett's (yellow), and gastric (blue) mucosa shows an excellent fit for all spectra.

bend) (23) and 1,162 cm^{-1}. The upper wave-number regime also exhibits peaks primarily associated with protein.

Quantification of FTIR Spectral Components. The FTIR spectra were fit to a linear model based on the tissue concentration of components based on the extinction coefficients derived from dry powder spectra. The average percent concentration of DNA, protein, glycogen, and glycoprotein from the squamous, Barrett's (nondysplasia and dysplasia), and gastric mucosal specimens along with the mean relative error are found (SI Table 2). Several key features of these findings should be noted. Significantly greater glycogen content was observed in squamous mucosa (15 ± 7) than in either Barrett's (3 ± 3, $P < 0.001$) or gastric (4 ± 3, $P < 0.001$), a result supported by the unique presence of glycogen bands at 1,026 and 1,154 cm^{-1} in squamous mucosa. Also, a much higher average glycoprotein concentration was seen in gastric mucosa (12 ± 4) than in either squamous (2 ± 2, $P < 0.001$) or Barrett's (3 ± 4, $P < 0.001$), a finding consistent with the unique appearance of a glycoprotein band at 1,117 cm^{-1} in gastric tissue. Barrett's specimens with dysplasia were found to have higher mean DNA content ($9 \pm 4\%$ versus $4 \pm 3\%$, $P < 0.001$) and ($8 \pm 4\%$ versus $3 \pm 3\%$, $P < 0.001$) and greater average glycoprotein concentration ($5 \pm 5\%$ versus $2 \pm 3\%$, $P = 0.03$) and ($5 \pm 4\%$ versus $2 \pm 3\%$, $P = 0.01$) for pathologists 1 and 2, respectively, compared with that of the Barrett's specimens without dysplasia. Finally, for the three different tissue types, the average relative error is <0.1%, which suggests the model provides an excellent fit to the measured spectra.

In Fig. 2*B*, the FTIR model fits (black) to measured spectra of representative specimens of squamous (red), nondysplastic Barrett's (green), dysplastic Barrett's (orange), and gastric (blue) mucosa are shown. The model parameters and relative error are determined (SI Table 3). In Fig. 3, the histology (H&E) corresponding to the squamous (*A*), nondysplastic Barrett's mucosa (*B*), dysplastic Barrett's mucosa (*C*), and gastric specimens (*D*) is shown. For the squamous spectrum, prominent bands can be seen at 1,026, 1,081 and 1,154 cm^{-1} representing glycogen that is present in the cytoplasm of squamous cells (Fig. 3*A*). DNA is present in the centrally located nuclei. The nondysplastic Barrett's mucosa spectrum reveals less prominent DNA bands at 970, 1,051, and 1,081 cm^{-1} than that of the dysplastic Barrett's, a finding confirmed by a lower model result for DNA. On histology, nuclei are basally located in the well organized glands of nondysplastic Barrett's (Fig. 3*B*), compared with the proliferating dysplastic glands with enlarged, atypical, and pseudostratified nuclei (Fig. 3*C*). For the gastric spectrum, prominent bands can be seen at 1,117 and 1,162 cm^{-1} that correspond to glycoprotein, which is produced by foveolar glands of the gastric cardia (Fig. 3*D*).

Comparison of FTIR Model with Histology. Based on the differences observed in component concentrations, a scatter plot using model coefficients of glycogen and glycoprotein is generated to demonstrate the ability of this method to classify squamous,

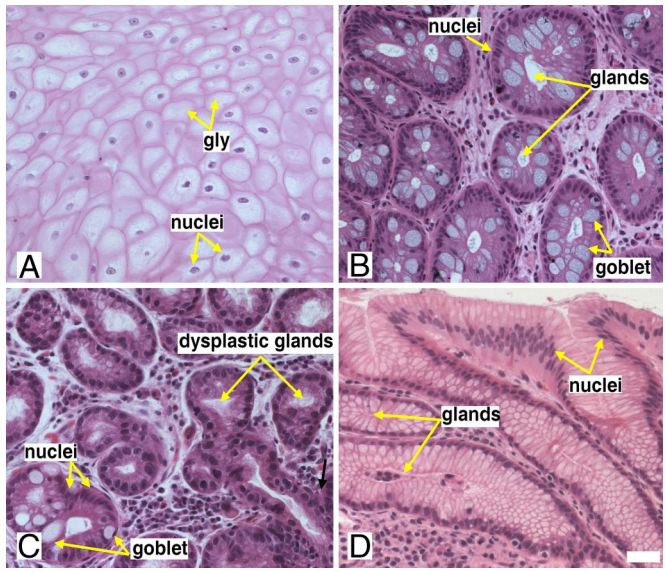

Fig. 3. Histology (H&E) corresponding to the measured FTIR spectra. (*A*) Squamous. Glycogen (gly) is present in the cytoplasm of squamous cells, and DNA is found in the centrally located nuclei. (*B*) Nondysplastic Barrett's. Nuclei are basally located in well organized glands characterized by goblet cells. (*C*) Dysplastic Barrett's. Proliferating dysplastic glands contain enlarged nuclei. (*D*) Gastric. Glycoprotein is produced by foveolar glands of the gastric cardia that have basally located nuclei. (Scale bar, 50 μm.)

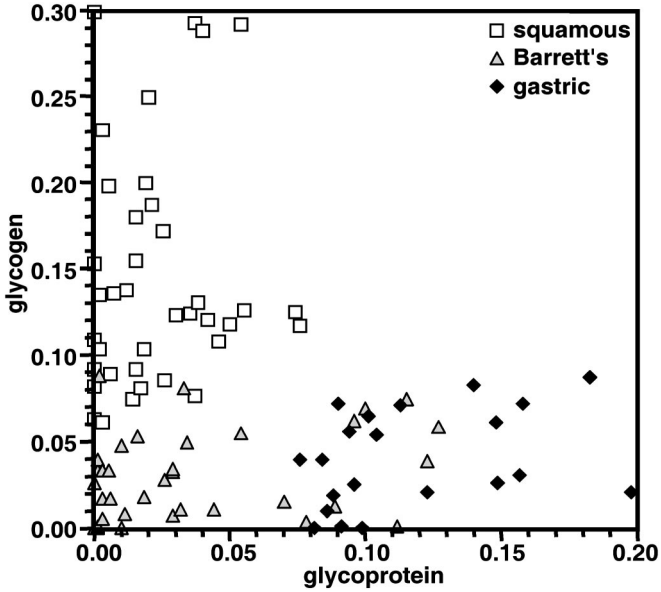

Fig. 4. Scatter plot of biopsy specimens. Glycogen and glycoprotein concentrations by FTIR are used to classify squamous and gastric mucosa with an accuracy of 98% and 92%, respectively.

Barrett's, and gastric mucosa, shown in Fig. 4. By using a linear discriminant function with leave-one-out cross-validation, squamous could be distinguished from columnar mucosa (Barrett's and gastric) with a sensitivity, specificity, positive predictive value (PPV), and accuracy of 100%, 97%, 95%, and 98%, respectively. The total number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) are identified (SI Table 4). Moreover, gastric mucosa could be discriminated from squamous and Barrett's with a sensitivity, specificity, PPV, and accuracy of 89%, 100%, 100%, and 92%, respectively. The FTIR model results for DNA and glycoprotein are shown for the Barrett's specimens subclassified as nondysplasia and dysplasia by pathologists 1 (Fig. 5*A*) and 2 (Fig. 5*B*). Agreement between pathologists for dysplasia by histology alone results in $\kappa = 0.52$ and produced a sensitivity, specificity, PPV, and accuracy of 67%, 85%, 80%, and 76%, with pathologist 1 as the standard. These results are consistent with previously published values (5). However, evaluation of only the data points ($n = 18$) with higher DNA and glycoprotein (gp) concentrations (those above the ellipsoidal curve), led to a better agreement between the two pathologists for dysplasia, $\kappa = 0.72$ and resulted in a sensitivity, specificity, PPV, and total accuracy of 92%, 80%, 92%, and 89%, respectively. In comparison, evaluation of only the data points ($n = 20$) with lower DNA and glycoprotein content (below the ellipsoidal curve), led to a much worse agreement, $\kappa = 0.17$, resulting in a sensitivity, specificity, PPV, and accuracy of 50%, 78%, 20%, and 75%, respectively.

Discussion

FTIR collects a rapid molecular fingerprint of tissue that consists of detailed infrared absorption bands in unprocessed tissue. In the ATR mode, light travels through the mucosa along a constant path and allows for quantitative spectral analysis. The results of this study show that individual absorbance peaks from tissue do not vary independently but, rather, change as a pattern determined by the concentrations of four biomolecules. This study also demonstrates that water can be effectively removed so that the infrared absorption spectra from a thin layer of the mucosal surface can reveal the subsurface histology. For example, increased glycogen and glycoprotein content by FTIR reveals the presence of squamous and gastric tissue, respectively. Although the presence of squamous epithelium within Barrett's can also be identified during endoscopy with iodine staining (24), FTIR is sensitive to small changes in multiple tissue constituents that can be used for discriminating tissue of various pathologies. Furthermore, increased DNA and glycoprotein content predicts the presence of dysplasia in Barrett's mucosa, and is consistent with histological findings of greater DNA activity such as mitoses, enlarged nuclei, and nuclear atypia and of special stains for mucin glycoprotein such as alcian blue (25). These results are consistent with previous studies that have found increased DNA content by image cytometry (26) and Feulgen staining (27). These studies used formalin-fixed rather than fresh tissues and found a smaller increase in DNA content for transformation from metaplasia to dysplasia. A limitation of this technique is that dehydration of the mucosal surface effaces the epithelium and creates artifacts on histology that may affect interpretation. Dehydration is performed to sharpen the absorption bands of amide I and II by removing the obscuring effect of water absorption in this regime. Some improvement in the quality of the histology was observed by rehydrating the tissue after completing the FTIR measurement.

In this study, FTIR provides an average concentration of four biomolecules from a surface area of 2 mm^2, determined by the geometry of the ATR crystal. Vibrational bands outside of this region were omitted to simplify the model. The model provides a very accurate fit despite the presence of molecular interactions and variations that may introduce a shift, distortion, or broadening of individual peaks. This robust feature is particularly

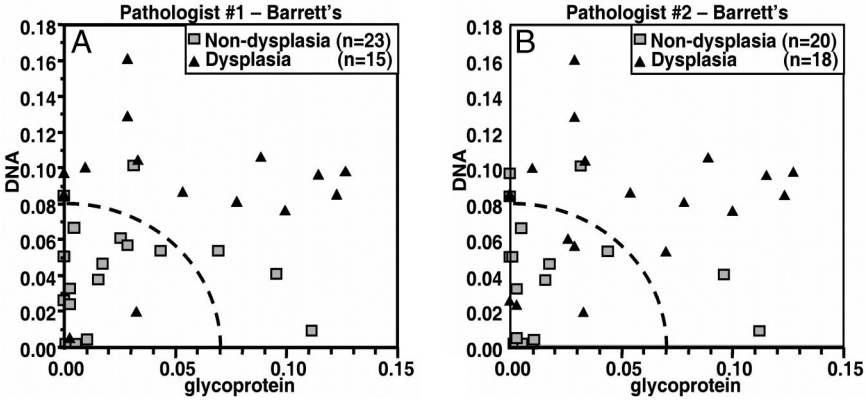

Fig. 5. Subclassification of Barrett's specimens as nondysplastic and dysplastic by pathologists 1 (*A*) and 2 (*B*) show significantly greater agreement for the data points above the dashed ellipsoidal curve ($\kappa = 0.72$), representing specimens with increased DNA and glycoprotein by FTIR, compared with those below ($\kappa = 0.17$).

relevant to components, such as glycogen and mucin glycoprotein, that have many subtypes that produce subtle differences in spectral line shapes. Moreover, the results show that other known tissue sources of infrared absorption, such as phospholipids and porphyrins, have a negligible contribution to the spectra in this regime. Protein had by far the highest concentration in tissue, a finding that is consistent with its proportion in cells (28). Thus, to accurately model its role, we used albumin and chymotrypsin to represent the α-helix and β-sheet conformations and found that a 50:50 mixture best fit the amide I and II peaks. Although other calibration models such as principal component analysis (29) and neural networks (30) may be better suited to analyze the fine structure of the FTIR spectra, a linear model provides an explanation of the sources of infrared absorption in terms of physiological parameters.

In comparison, Raman spectroscopy is an inelastic scattering technique based on unique selection rules that produces weak signals and can become overwhelmed by tissue autofluorescence (31). Consequently, Raman requires expensive detectors and sophisticated spectral processing, conditions that can be difficult to translate to the clinic for screening purposes (32, 33). On the other hand, FTIR is an infrared absorption process that produces large signals that are insensitive to tissue autofluorescence. As a result, relatively inexpensive detectors and simple spectral analysis are needed, conditions that are more compatible with clinical use. Although our FTIR results for identifying dysplasia appear promising, the accuracy of this technique was determined by comparison with the evaluation of the pathologist used as the "gold" standard, which is not ideal for histopathological interpretation. The ultimate determinant of the utility of this adjunct technique would be its ability to improve long-term patient outcomes, both in regards to the subsequent incidence of adenocarcinoma and to patient survival. Furthermore, analysis of the more subtle features may reveal spectral findings that appear well before histological changes occur. These results can be used to better stratify patients at a higher risk of developing adenocarcinoma for increased frequency of surveillance.

Materials and Methods

FTIR Spectra Collection from Esophageal Specimens. A total of 32 patients with a mean age 67.9 years (range, 52–88), history of Barrett's esophagus (intestinal metaplasia on biopsy), and no history of esophageal malignancy, undergoing routine-surveillance endoscopy were recruited for this study. The average length of Barrett's mucosa screened was 4.0 ± 1.6 cm (range, 1–8). The protocol was approved by the human subjects committee at Stanford Hospital and Clinics and the Veterans Affairs Palo Alto Health Care System. After informed consent and completion of the endoscopic evaluation, several pinch biopsies were collected from the distal esophagus with jumbo forceps. The specimens were transported on moist gauze immediately to the spectrometer (Nexus 470; Thermo Electron, Madison, WI). Cool air was gently blown onto the mucosal surface for ≈20 sec to remove ambient moisture and thus reduce the effect of water absorption in the collected spectra. The mucosal surface of the specimens was placed in contact with the ATR crystal, which has a single reflection geometry and is made of zinc selenide (ZnSe). Further dehydration of each sample was performed until an amide I/amide II band ratio of ≈0.85 was observed. Only specimens with sufficient size to completely cover the 2 mm^2 surface area of the ATR crystal were included in the study. A total of eight coadded scans from 950 to 1,800 cm^{-1} with 4 cm^{-1} resolution by using a DTGS detector were collected in <3 sec. The spectrometer was continually purged with dry air to remove water vapor and CO_2. After the spectra were collected, the tissue was rehydrated with normal saline (0.9%) and placed in 10% formalin. The specimens were then cut in 4-μm sections and stained with H&E.

Analysis of Pure Component Spectra. Dehydrated films of DNA, protein, glycoprotein, and glycogen were used without further purification and include (*i*) calf thymus DNA (10 mg/ml; Invitrogen, Carlsbad, CA); (*ii*) 50:50 mixture by weight of human serum albumin (263 mg/ml; GTC Biotherapeutics, Framingham, MA) and type 2 α-chymotrypsin (500 mg/ml; Sigma, St. Louis, MO); (*iii*) type 1-S mucin (100 mg/ml; Sigma); and (*iv*) glycogen (20 mg/ml; Boehringer Ingelheim, Ingelheim, Germany). Approximately 1 ml of each standard was placed onto a CaF_2 window and dehydrated with cool air before spectra collection. Spectral peaks were identified by using Omnic software, version 6.2 (Thermo Electron).

Model of FTIR Spectra. The AUS of the measured spectrum $S(\nu)$ is determined by summing the absorbance values over the spectral regime $950 < \nu < 1{,}800$ cm^{-1}.

$$AUS = \sum_{\nu=950}^{1800} S(\nu) \quad [1]$$

A model $S'(\nu)$ of the measured spectrum $S(\nu)$ assumes a linear (Beer's law) relationship between absorbance and (*i*) DNA, $d(\nu)$; (*ii*) protein, $p(\nu)$; (*iii*) glycogen, $g(\nu)$; and (*iv*) glycoprotein, $gp(\nu)$.

$$S'(\nu) = c_1 d(\nu) + c_2 p(\nu) + c_3 g(\nu) + c_4 gp(\nu) \quad [2]$$

The percent concentration c_i of each component was determined by first subtracting out the spectrum of water from that of tissue by a factor that achieved an amide I/amide II band ratio of 0.85 (34). Then, the error function $E(\nu)$ was minimized by a partial least-squares fit using TQ Analyst Pro Ed software, version 6.2 (Thermo Electron). The four coefficients were constrained to sum to 100%.

$$E(\nu) = (S'(\nu) - S(\nu))^2 \quad [3]$$

The relative error (Rel Err) is defined as the ratio of the sum of the absolute difference between the measured spectrum and the model result $|\, S(\nu) - S'(\nu) \,|$ over the spectral range $950 < \nu < 1{,}800$ cm^{-1} and the AUS of the measured spectrum.

$$Rel\ Err = \sum_{\nu=950}^{1800} |S'(\nu) - S(\nu)| / \sum_{\nu=950}^{1800} S(\nu) \quad [4]$$

Histopathological Evaluation of Tissue Specimens. Histopathological evaluation was performed independently by two gastrointestinal pathologists blinded to the FTIR results. The specimens were classified as follows: (*i*) squamous mucosa with no, minimal, or marked basal zone hyperplasia; (*ii*) intestinal metaplasia (Barrett's esophagus) with no, mild, or moderate hyperplasia, no dysplasia; (*iii*) intestinal metaplasia (Barrett's esophagus) with mild, moderate, or severe dysplasia; and (*iv*) gastric cardia/fundic mucosa with no, minimal, or moderate foveloar cell hyperplasia. Specimens with excessive tissue artifact, chronic inflammation, or insufficient tissue for evaluation were excluded from the analysis.

Statistical Analysis. Statistical significance (*P* value) was calculated by using a two-sided Student's *t* test with unequal variance, and interobserver agreement was determined by using κ-statistics with an equal range of scores. Diagnostic performance is defined by sensitivity = TP/(TP + FN), specificity = TN/(TN + FP), positive predictive value = TP/(TP + FP), and total accuracy = (TP + TN)/(TP + TN + FP + FN), given the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). A linear discriminant function was used to classify the data with STATA, version 9.2 (STATA, College Station, TX), and the performance of the FTIR model to classify squamous and gastric mucosa was determined by using the leave-one-out, cross-validation technique (35).

We thank Harvey Young for technical support and Mark Mackanos and Patrick Wong for insightful discussions. This work was supported, in part, by National Institutes of Health Grants K08 DK67618 (to T.D.W.), R01 DK063624 (to G.T.), and U54 CA105296 (to T.D.W., G.T., J.M.C., L.R.D., T.B., P.S., S.F., R.S., and C.H.C.) and by STI Medical Systems (Honolulu, HI).

1. American Cancer Society (2006) *Cancer Facts and Figures 2006* (Am Cancer Soc, Atlanta).
2. Gerson L, Groeneveld P, Triadafilopoulos G (2004) *Clin Gastroenterol Hepatol* 2:868–879.
3. Spechler S (2002) *N Engl J Med* 346:836–842.
4. Montgomery E, Goldblum J, Greenson J, Haber M, Lamps L, Lauwers G, Lazenby A, Lewin D, Robert M, Washington K, *et al.* (2001) *Hum Pathol* 32:379–388.
5. Reid BJ, Haggitt RC, Rubin CE, Roth G, Surawicz CM, Van Belle G, Lewin K, Weinstein WM, Antonioli DA, Goldman H, *et al.* (1988) *Hum Pathol* 19:166–178.
6. Reid B, Levine D, Longton G, Blount P, Rabinovitch P (2000) *Am J Gastroenterol* 95:1669–1676.
7. Cameron A, Carpenter H (1997) *Am J Gastroenterol* 92:586–591.
8. Bates J (1976) *Science* 191:31–37.
9. Parker F (1971) *Applications of Infrared Spectroscopy in Biochemistry, Biology, and Medicine* (Plenum, New York), 3rd Ed.
10. Jackson M, Mantsch H (1996) *Biomedical Infrared Spectroscopy in Infrared Spectroscopy of Biomolecules* (Wiley–Liss, New York).
11. Fernandez D, Bhargava R, Hewitt S, Levin I (2005) *Nat Biotechnol* 23:469–474.
12. Levin I, Bhargava R (2005) *Annu Rev Phys Chem* 56:429–474.
13. Ci Y, Gao T, Feng J, Guo Z (1999) *Appl Spectrosc* 53:312–315.
14. Wong P, Wong R, Caputo T, Godwin T, Rigas B (1991) *Proc Natl Acad Sci USA* 88:10988–10992.
15. Rigas B, Morgello S, Goldman I, Wong P (1990) *Proc Natl Acad Sci USA* 87:8140–8144.
16. Yano K, Ohoshima S, Gotou Y, Kumaido K, Moriguchi T, Katayama H (2000) *Anal Biochem* 287:218–225.
17. Li Q, Sun X, Xu Y, Yang L, Zhang Y, Weng S, Shi J, Wu J (2005) *Clin Chem* 51:346–350.
18. Kondepati V, Zimmermann J, Keese M, Sturm J, Manegold B, Backhaus J (2005) *J Biomed Opt* 10:054016 1–6.
19. Lu Y, Yalamnchili M, Miller J (1998) *Appl Spectrosc* 52:851–854.
20. Diem, M (1993) *Introduction to Modern Vibrational Spectroscopy* (Wiley–Interscience, New York).
21. Chiriboga L, Xie P, Yee H, Vigorita V, Zarou D, Zakim D, Diem M (1998) *Biospectroscopy* 4:47–53.
22. Pevsner A, Diem M (2001) *Appl Spectrosc* 55:788–793.
23. Patel M, Smart J, Nevell T, Ewen R, Eaton P, Tsibouklis J (2003) *Biomacromolecules* 4:1184–1190.
24. Takubo K, Vieth M, Aryal G, Honma N, Sawabe M, Arai T, Kammori M, Mafune K, Iwakiri K (2005) *Hum Pathol* 36:269–274.
25. Bancroft JD, Gamble, M (2002) *Theory and Practice of Histological Techniques* (Churchill–Livingstone, New York), 5th Ed.
26. Fang M, Lew E, Klein M, Sebo T, Su Y, Goyal R (2004) *Am J Gastroenterol* 99:1887–1894.
27. Khan M, Bui H, del Rosario A, Abdulla M, Ballouk F, Sim Y, Ross J (1994) *Mod Pathol* 7:169–174.
28. Diem M, Boydston-White S, Chiriboga L (1999) *Appl Spectrosc* 53:148A–161A.
29. German MJ, Hammiche A, Ragavan N, Tobin MJ, Cooper LJ, Matanhelia SS, Hindley AC, Nicholson CM, Fullwood NJ, Pollock HM, Martin FL (2006) *Biophys J* 90:3783–3795.
30. Mark S, Sahu R, Kantarovich K, Podshyvalov A, Guterman H, Goldstein J, Jagannathan R, Argov S, Mordechai S (2004) *J Biomed Opt* 9:558–567.
31. Shim M, Song L, Marcon N, Wilson B (2000) *Photochem Photobiol* 72:146–150.
32. Kendall C, Stone N, Shepherd N, Geboes K, Warren B, Bennett R, Barr H (2003) *J Pathol* 200:602–609.
33. Shetty G, Kendall C, Shepherd N, Stone N, Barr H (2006) *Br J Cancer* 94:1460–1464.
34. Lucassen G, van Veen G, Jansen J (1998) *J Biomed Opt* 3:267–280.
35. Vijarnsorn M, Riley C, Ryan D, Rose P, Shaw R (2007) *Am J Vet Res* 68:517–523.