Biochemical differences in tumorigenic and nontumorigenic cells measured by Raman and infrared spectroscopy

Judith R. Mourant
Kurt W. Short
Susan Carpenter
Nagapratima Kunapareddy
Leslie Coburn
Tamara M. Powers
James P. Freyer
Los Alamos National Laboratory
MS E535, Bioscience Division
Los Alamos, New Mexico 87544

Abstract. Both infrared and Raman spectroscopies have the potential to noninvasively estimate the biochemical composition of mammalian cells, although this cannot be unambiguously determined from analysis approaches such as peak assignment or multivariate classification methods. We have developed a fitting routine that determines biochemical composition using basis spectra for the major types of biochemicals found in mammalian cells (protein, DNA, RNA, lipid and glycogen), which is shown to be robust and reproducible. We measured both infrared and Raman spectra of viable suspensions of pairs of nontumorigenic and tumorigenic rat fibroblast cell lines. To model *in vivo* conditions, we compared nonproliferating, nontumorigenic cells to proliferating, tumorigenic cells. Reproducible differences in biochemical composition were found for both nontumorigenic/tumorigenic cell models, using both spectroscopic techniques. These included an increased fraction of protein and nucleic acids in the tumorigenic cells, with a corresponding decrease in lipid and glycogen fractions. Measurements of each cell type in both the proliferating and nonproliferating states showed that proliferative status was the major determinant of differences in vibrational spectra, rather than tumorigenicity *per se*. The smallness of the spectral changes associated with tumorgenicity may be due to the subtle nature of the oncogenic change in this system (a single mutant oncogene).
[DOI: 10.1117/1.1928050]

Keywords: mid-infrared spectroscopy; Raman spectroscopy; cancer diagnostics; biological cells; biochemical analysis.

Paper SS04100 received Jun. 15, 2004; revised manuscript received Mar. 9, 2005; accepted for publication Mar. 14, 2005; published online Jun. 16, 2005.

1 Introduction

Vibrational spectroscopy has the potential to detect and quantify changes in biochemistry occurring during carcinogenesis, and several papers have reported differences in Raman or Fourier transform infrared (FTIR) spectra of cancerous and noncancerous cells/tissue. Primarily these papers have used multivariate classification methods such as neural networks, cluster analysis, or principle component analysis combined with linear discriminant analysis to distinguish cancerous and normal tissue.[1–11] Although such an approach can discriminate spectra from different tissues, these analysis methods cannot provide detailed information regarding biochemical changes. One of the major advantages of infrared and Raman spectroscopies is their ability to provide detailed biochemical information. Therefore, with multivariate analysis techniques, one of the advantages of Raman and infrared spectroscopies is not realized. When attempts are made to identify biochemical changes, they frequently rely on (tentative) peak assignments and/or result in only qualitative statements.[5–11] As noted by other authors, as well as ourselves, there are discrepancies in the literature regarding peak assignments and most peaks have contributions from more than one biochemical component.[4,5,12] Consequently, biochemical changes determined by peak assignments are often merely speculative. Our goal in this work is to use a more robust biochemical analysis of vibrational spectra to gain a fundamental understanding of the biochemical changes accompanying carcinogenesis.

Most cancers originate in the epithelium, which is comprised primarily of cells with very little interstitial structural material. Consequently, one model for carcinogenesis would be to measure cancerous and noncancerous epithelial cells. Ideally this model would use cancerous cells that were derived directly from the noncancerous cells, mimicking an *in vivo* setting. Such models are not readily available for epithelial cells. There are, however, such models for fibroblast cells, and we have performed Raman and infrared spectroscopy on cells from two such models. An additional advantage of using

Address all correspondence to Judith Mourant, Los Alamos National Lab., Bioscience Division, MS E535, Los Alamos, New Mexico 87545. Tel: 505-665-1190; Fax: 505-665-4637. E-mail: jmourant@lanl.gov

1083-3668/2005/$22.00 © 2005 SPIE

cells rather than tissue is that the biochemical changes in the cells can be isolated.

In this work we have used M1, MR1, Rat1, and Rat1-T1 fibroblast cells. M1 and Rat1 cells are immortal but not tumorigenic and were derived from rat embryo fibroblast cells: M1 by transfection of a mutant *myc* oncogene and Rat1 by an unknown spontaneous event. MR1 and Rat1-T1 cells are tumorigenic and were derived from M1 and Rat1 cells, respectively, by transfection of a mutant *ras* oncogene.[13] Vibrational spectra were obtained of all four cell lines from both exponentially growing cell cultures and from cell cultures which had reached a plateau in growth. The reason for measuring cells in different proliferative states is that cell proliferation plays an important role in cancer initiation and progression. It is well known that transformed cells have a higher proliferative index than the normal tissue from which they originated.[14] Therefore, as one model of cancer, we compare exponentially growing MR1 cell cultures with M1 cell cultures that have reached the plateau phase of growth. In the other model, exponentially growing Rat1-T1 cell cultures are compared with Rat1 cell cultures that have reached a plateau of growth.

2 Materials and Methods

2.1 *Preparation of Cells*

Monolayer cultures were routinely maintained and subcultured for up to 20 passages (cumulative population doublings 120) as described in detail elsewhere.[13,15] Briefly, cells were cultured as monolayers in standard tissue culture flasks using Dulbecco's modified eagle's medium (DMEM, Invitrogen) containing 4.5 g/L D-glucose, 5% (v/v) fetal calf serum (Hyclone), 100 IU/mL penicillin, and 100 μg/mL streptomycin (Invitrogen) referred to hereafter as complete medium. Cell suspensions were obtained from monolayer cultures by treatment for 10 min with 0.25% trypsin in a phosphate buffer (*p*H 7.4) containing 1 mM EDTA and 25 mM HEPES, followed by the addition of cold complete medium. Cell suspensions were passed twice through an 18 gauge needle, centrifuged into a pellet (1500 rpm for 10 min), and the medium was removed. The cell pellet was resuspended in phosphate buffered saline (PBS), centrifuged again to remove residual medium, and PBS was added to obtain the final concentration of 1×10^8 to 2.5×10^8 cells/mL used for infrared and Raman measurements of M1 and MR1 cells. For Raman spectroscopy of Rat1 and Rat-T1 cells, cell suspensions are prepared as described except that after the second time supernatent is removed and PBS is added, the cells are centrifuged into a black, delron plastic cylinder 1 cm in diameter$\times$1 cm in depth to obtain a cell pellet of $2-5\times10^8$ cells/mL. Growth curve experiments showed that monolayers of MR1 and Rat1-T1 cells reached their growth plateau at $\sim6\times10^5$ cells/cm^2, and $3-5\times10^5$ cells/cm^2, respectively, while M1 and Rat1 cells reached their growth plateau at $1-2\times10^5$ cells/cm^2. Based on these data, exponentially growing cell suspensions were obtained from monolayer cultures harvested at a cell density of less than 1/3 of confluent cultures, while plateau-phase suspensions were obtained from monolayer cultures harvested after 2–3 days at confluence. The proliferative status of each of these suspensions was confirmed by flow cytometric DNA content analysis as described below.

2.2 *Cell Counting and Cell Volume Analysis*

An aliquot of each cell suspension was counted using an electronic particle counter equipped with a pulse-height analyzer (Coulter Electronics) as described previously.[16] Briefly, a cell volume distribution was obtained and gates were set to count only intact cells, excluding acellular debris. Three counts were taken for each sample and averaged to determine the concentration of cells in the suspension. After counting, a cell volume distribution containing $>10^4$ cells was saved and processed on a computer to obtain the mean volume of the cells in the suspension. Absolute volumes were determined through calibration of the particle counter using five different sizes of polystyrene microspheres (Duke Scientific). Cell volume distributions measured before and after spectroscopy were compared to monitor for changes in the cell size distribution. No significant changes were found.

2.3 *Cell Cycle Analysis and Proliferative Status*

Determination of the cell cycle distribution was performed using flow cytometric DNA content analysis as described in detail previously.[16] Briefly, an aliquot of 10^6 cells was fixed in 70% ethanol and refrigerated. Fixed samples were prepared for analysis by centrifuging the cells to a pellet (1500 rpm for 10 min), decanting the ethanol and resuspending the cells in 1 mL of a DNA staining solution containing 50 μg/mL propidium iodide (Sigma) and 100 units/mL RNase (Sigma) in PBS containing calcium and magnesium (InVitrogen). Cells remained in the staining solution overnight at 4 °C. DNA content analysis was performed on a FACS Calibur (Becton-Dickenson) flow cytometer using 488 nm excitation and fluorescence collection with the propidium iodide filter set. DNA content histograms containing greater than 10^4 cells were collected with coefficients of variation on the G_1-phase peak of $<5\%$. These histograms were analyzed for cell cycle distribution with the MODFIT LT program (Verity Software House) using the debris and aggregate elimination options.

2.4 *Collection of FTIR Data*

The infrared spectra of cells in PBS were obtained in transmission mode. The sample chamber consists of two rectangular barium fluoride windows held in apposition with an oval-shaped 50-μm-thick ring of Teflon between them. The upper window has a hole at the top for inserting the sample and a hole at the bottom for allowing air to escape while loading the sample. A rubber gasket was placed over the upper window to seal the holes and to prevent the sample from leaking. Fibroblast cells were loaded into the 50-μm-thick sample space between the two $BaFl_2$ windows by using a syringe to push the cell suspension into the sample space. The infrared beam spot was $\sim$1 cm in diameter and the spectra were obtained at 2 cm^{-1} resolution with 200 scans per spectra. Collection time for a single spectrum on our Fourier transform infrared spectrometer equipped with a DTGS detector (Mattson Cygnus-100) was 7 min. The cells were not on ice during the measurement, however, we have found that this amount of time at room temperature does not affect cell viability.[17] A typical measurement protocol consisted of taking two spectra of phosphate buffered saline (PBS), two spectra of cells in PBS, two spectra of PBS, two spectra of cells in PBS, and two spectra of PBS. All measurements were performed in the

Table 1 Number of samples measured. Cell cultures harvested in the exponential and plateau phase are denoted by an e or a p after the cell name, respectively.

	Rat1e	Rat1p	Rat1-T1e	Rat1-T1p	M1e	M1p	MR1e	MR1p
Raman	4	6	6	4	7	8	6	7
FTIR	3	11	11	4	9	7	6	7

same sample chamber, which meant reloading the sample chamber for each set of two measurements. The absorbance of the cells is calculated according to Eq. (1) where $I_{\text{cells\_ave}}$ is the average of the intensity measured when cells were in the sample chamber and $I_{\text{PBS\_ave}}$ is the average of the PBS spectra. The number of samples measured are shown in Table 1.

$$\text{Absorbance} = -\log(I_{\text{cells\_ave}}/I_{\text{PBS\_ave}}). \quad (1)$$

2.5 *Description of the Raman Instrumentation*

Excitation is provided by a 785 nm diode laser coupled via a fiber-optic to a probe head containing a holographic grating to eliminate Raman signals from the fiber (Kaiser Optical Systems). The excitation light was focused onto the sample using an aspheric lens, f0.7 (Edmund Industrial Optics). Backscattered light is collected through the same lens and passed through the probe head, where the excitation light was eliminated with two notch filters. A fiber-optic cable connects the probe head to a spectrograph (HoloSpec f1.8, Kaiser Optical Systems) which has an additional notch filter. Dispersed Raman signal is detected using a deep depletion, backilluminated liquid nitrogen cooled charged coupled device (Princeton Instruments) attached to the spectrograph. Two removable holographic gratings were used in the spectrograph. The first covered the low wavenumber spectral range 100–1900 cm^{-1}, while the second covered the high wavenumber spectral range from 1780 to 3250 cm^{-1}. The spectral resolution, defined as twice the spectral bandpass (full width at half maximum), was ~6.2 and 5.0 cm^{-1}, for the low and the high wavenumber regions, respectively.

2.6 *Collection of Raman Data*

All data were collected with a laser power of 185–200 mW. This laser power is not expected to cause any significant increase in sample temperature due to the extremely low absorption coefficent of cells at 785 nm. Assuming water is the primary absorber at 785 nm, the absorption coefficient of our samples is ~0.025 cm^{-1}. While tissues containing large amounts of blood, such as liver, show significant heating, tissues containing small amounts of blood undergo little or no heating.[18] A measurement of a standard tungsten source (Optronic Laboratories Inc.) was used to determine and correct for the instrumental response. The number of cell samples measured is given in Table 1.

During collection of data on M1 and MR1 cell suspensions, the spatial cosmic ray correction of the WinSpec/32 program (Princeton Instruments) was used at 100%. This correction had a slight smoothing effect on the data, equivalent to about a five-point boxcar smooth. For a given experiment, collection times were normally 300 s per spectrum. The M1 and MR1 cell suspensions were contained in 5-mm-diameter suprasil nuclear magnetic resonance (NMR) tubes (Wilmad/Lab Glass) and the bottom of the tubes were placed in an ice water bath during data collection. For each experiment, two spectra were collected of the empty NMR tube, two spectra of the NMR tube containing the buffer, and ten spectra of the NMR tube containing the cell culture. The spectrum of the tungsten source was smoothed using a 15-point boxcar function and divided by the manufacturer provided spectral irradiance of the tungsten lamp to give the final spectrum used to correct for instrument response.

The data collection methods for the Rat1 and Rat1-T1 cells were slightly different than for the M1 and MR1 measurements. The cells were measured in an open-faced, black, delron plastic sample chamber to eliminate the need to subtract out the spectrum of the NMR tubes. Second, the cells were measured at a higher concentration to increase the signal-to-noise ratio. Third, a tungsten lamp spectrum was recorded every time the grating was changed to facilitate a flat-field correction of the data. Finally, the temporal cosmic ray filter was used rather than the spatial cosmic ray filter to avoid the smoothing effect of the spatial filter and for better cosmic ray elimination. For measurement, the Rat1 or Rat1-T1 cells were spun into the 1 cm in diameter by 1 cm in depth sample chamber which was then placed in an ice water bath. The focus of the laser was lowered onto the top of the cell pellet until the signal intensity was maximized. Spectra were acquired for 20 min each in the low and high wavenumber regions. A spectrum of the empty sample chamber, and a spectrum of PBS in the empty sample chamber, were also obtained on most days when cells were measured. The intensity of the empty sample chamber spectrum was found to consist primarily of Raman scattering and fluorescence from the measurement optics with practically no contribution from the sample chamber itself.

The volume of sample measured was estimated to be 0.008 mm^3. This estimate was based on a measurement of the beam profile as well as measurements of different depths of cells which were examined for Raman scattering from the sample chamber under the cells. The number of cells in the 0.008 mm^3 volume is estimated to be 2000 using the measured average volume of a cell of 2000 μm^3 and assuming 50% of the volume is cells and 50% of the volume is interstitial space.

2.6.1 *Determination of biochemical component spectra*

Choice of biochemical components DNA: Calf thymus DNA (Sigma-Aldrich) dissolved in TE buffer (10 mM TRIS-HCl, 1 mM EDTA, *p*H 8) was measured. This source of DNA was chosen, because it is highly polymerized like the DNA in cells. Raman measurements were made at concentrations of

10.9, 15.6, and 17.5 mg/mL, while all of the infrared measurements were made at a concentration of 10 mg/mL.

RNA: Calf liver RNA (Sigma-Aldrich) was measured in TE buffer at concentrations of 40 and 42.7 mg/mL for the Raman measurements and 10 and 40 mg/mL for the IR measurements.

Lipid: Extract from liver cells was used (Avanti Polar Lipids). The composition is given by the manufacturer to be 5%–7% cholesterol, 42% phosphatidylcholine (PC), 22%–26% phosphatidylethanolamine (PE), 8%–9% phosphatidylinositol (PI), and 18%–22% other. This composition compares reasonably well with the known composition of lipids in fibroblasts: 9%–12% cholesterol,[19] 42%–51% PC, 14%–18% PE, ~8% PI.[20] For measurement, the lipid was dissolved in a buffer containing 10 mM TRIS-HCl, 100 mM NaCl, 20 mM sodium azide at a concentration of 40 mg/mL.

Protein: Protein was isolated using nondenaturing conditions from previously frozen cells. For the initial extraction, the cells were resuspended with tissue protein extraction reagent (TPER, Pierce) at 1 mL of TPER for each 0.05 g of cells. The protease inhibitors (stock solution concentrations in parenthesis), aprotinin (1 mg/mL), phenylmethylsufonylfluoride (10 mg/mL), leupeptin (1 mg/mL), and dithiothreitol (1 mol/L) were added at 1 μL/mL TPER. The suspension was sonicated at 4 °C for five 10 s periods with 20 s between each sonication, using 7 W of output power from a sonic dismembrator (Fisher Scientific). The sample was centrifuged at 6000 rpm for 25 min at 4 °C to remove cellular debris. RNase (1 mg/mL) and DNase (1 mg/mL) (both from bovine pancreas, Sigma-Aldrich) were added at 10 μL/mL TPER to the supernatant. The supernatant was allowed to sit at room temperature for 90–120 min after which it was run on a size exclusion column (Biogel A 0.5 m gel, Bio-Rad) at 4 °C equilibated with 100 mM NaCl. Fractions were collected using an automatic fraction collector. UV detection at 280 nm during collection showed three broad peaks. A Bradford assay[21] showed the middle peak to be protein. IR spectroscopy suggested that the first and third peaks contained mostly lipid-like components with no protein. The fractions from the middle peak were concentrated by ultrafiltration first using a 200 mL stirred cell and then a 10 mL stirred cell (Millipore). Further concentration was done in an ultracentrifuge using Microcon centrifugal filter devices (Millipore). During all concentration steps, 10 000 molecular weight cutoff filters were used. Final protein concentration was determined using a Bradford assay. IR and Raman spectra were collected immediately after concentration, as we have found that the spectra change if the protein is allowed to sit for more than 12 h at 4 °C. For the fits of Raman data, an average spectrum from four isolations of protein was used. Two were of Rat1 cells harvested in the plateau phase of growth (~19 mg/mL) and two were from Rat1-T1 cells harvested in the exponential phase of growth (~26 mg/mL). No difference was found in the protein spectra of the two cell types to within the signal to noise of our measurements. For the fits of infrared data on M1 and MR1 cells, a spectrum of protein extracted from MR1 cells in the plateau phase of growth was used (7 mg/mL).

Glycogen: Glycogen from bovine liver (Sigma-Aldrich) was dissolved in either *p*H 7 phosphate buffer or PBS. No dependence of the spectra on buffer was found. Concentrations of 15 and 20 mg/mL were used for Raman spectroscopy and concentrations of 10, 15, and 20 mg/mL were used for the infrared measurements.

Data acquisition Multiple measurements of separate samples were made on both the Raman spectrometer and the FTIR. The purpose of making multiple measurements was to assure that the data were repeatable. For the Raman measurements, reproducibility of the spectral intensities over a time span of weeks insured that the configuration used for the measurements was stable and facilitated the extraction of absolute concentrations of biochemical components from the analysis of the cell data. All Raman data were collected at a laser power of 185–200 mW using the temporal cosmic ray filter. The amount of time data were collected varied depending on the signal to noise being obtained. For example, more data were collected on protein than on DNA because the fluorescence background is higher for protein.

Analysis of component spectra Regions of the low wavenumber Raman spectra were simultaneously fit to a fifth order polynomial plus a spectrum of the empty sample cell plus a spectrum of the buffer. An example of a fit is shown in Fig. 1. The regions of the spectrum chosen for the fit were regions in which, based on the literature, no Raman bands were expected from the biochemical being measured. The purpose of the fifth order polynomial was to model the fluorescence. After fitting, the contributions of the polynomial, the buffer, and the empty cell spectra were subtracted from each component spectrum. High wavenumber Raman spectra were also fit to a fifth order polynomial, a spectrum of the empty sample cell and a spectrum of the buffer, and these contributions were subtracted from each component spectrum. The regions chosen for the fit were regions where there did not appear to be Raman bands due to the biochemical of interest. Before fitting the Raman M1/MR1 data, the component spectra were smoothed with a 5-pt boxcar to simulate the spatial cosmic ray filter used in acquiring the cell data. The analysis of the FTIR data was analagous to the analysis of the cell data given by Eq. (1).

To quantitatively analyze overlap between the component spectra, normalized dot products were calculated. Let A and B be component spectra and *i* be the index for wavenumber. The normalized dot products were then calculated as

$$\sum A(i)*B(i) \Big/ \left[\sum A^2(i)\right] \Big/ \left[\sum B^2(i)\right],$$

where the summations are over the range of wavenumbers used in the fitting procedures described below.

2.6.2 *Analyzing the cell spectra*

Raman and FTIR spectra were each fit to a combination of the biochemical components and baseline terms by minimizing chi-squared. All of the data were fit using the Levenberg–Marquardt method[22] as implemented in Igor Pro (Wavemetrics) to minimize chi-squared. For the Raman data, fits were done with several different starting conditions to assure that the global minimum rather than a local minima was found. Additionally, all of the FTIR data on Rat1 and Rat1-T1 cells were also fit using a linear least squares routine. Results from this method were identical (to three signficant figures) to

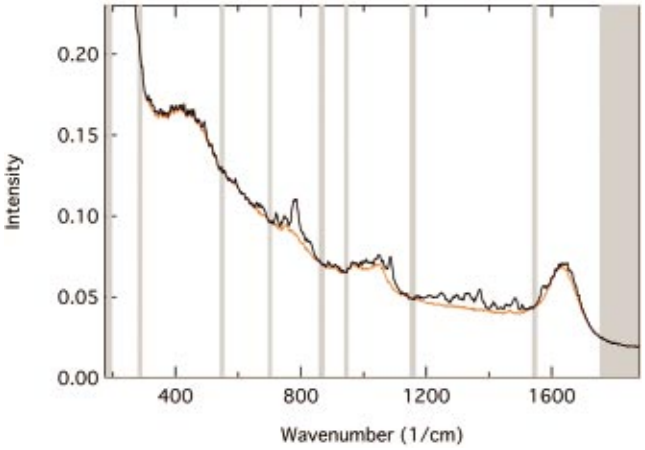

Fig. 1 Example of a fit used for determining the contributions of the buffer (PBS), the optics and empty sample cell, and fluorescence to a spectrum of DNA. The black line is the spectrum of DNA, the red line is a fit which is a linear combination of (1) a spectrum taken with the sample cell empty, (2) a spectrum of phosphate buffered saline (PBS) and (3) a fifth order polynomial (representing fluorescence). The fit was performed only at the wavelengths that are shaded gray.

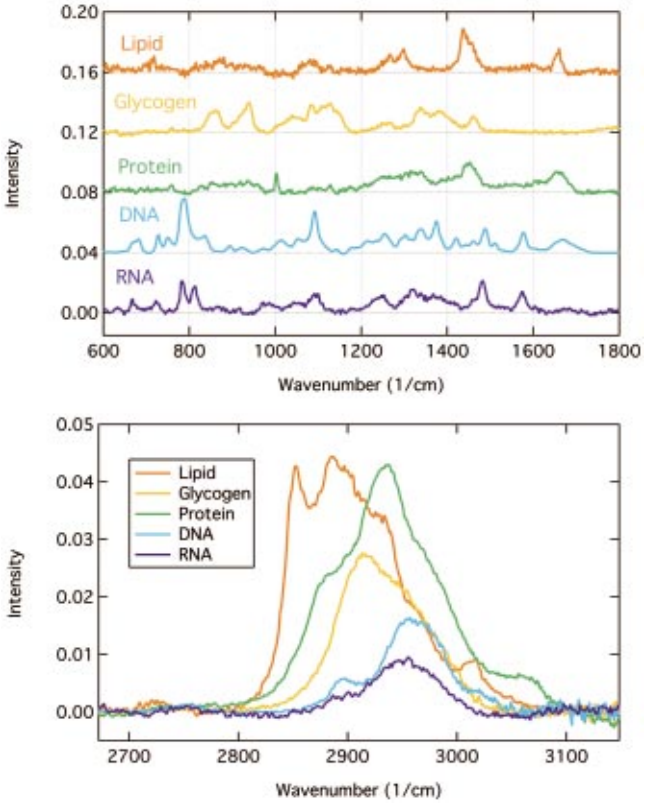

Fig. 2 Raman spectra of biochemical components of cells. The low wavenumber spectra (top) have been offset by multiples of 0.04 for clarity. All spectra were scaled to a concentration of 20 mg/mL.

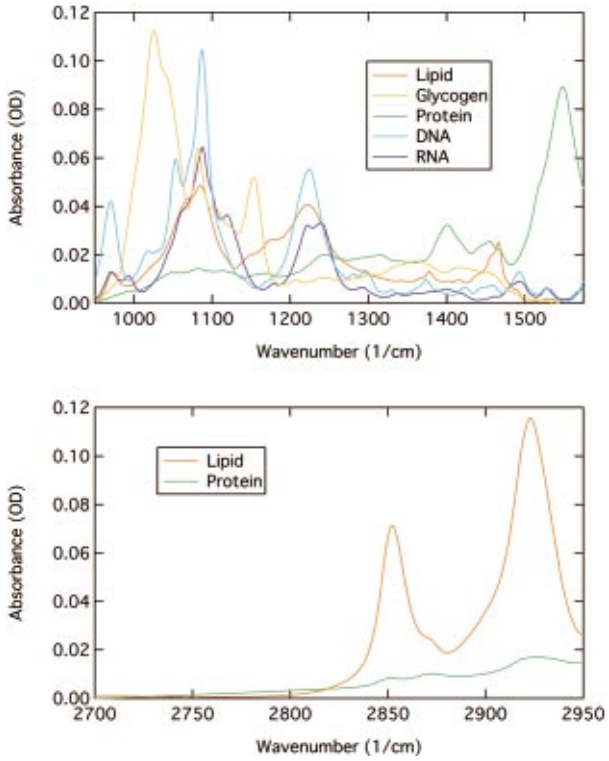

Fig. 3 Infared component spectra. (The protein spectrum shown was used for analysis of Rat1 and Rat1-T1 data. A different protein spectrum was used for M1 and MR1 cells.) All spectra were scaled to a concentration of 10 mg/mL. In the high wavenumber region only lipid and protein spectra are shown, because no other components had significant absorbance.

those obtained using the Levenberg–Marquardt method to minimize chi-squared, demonstrating that the Levenberg–Marquardt method was finding the global minimum rather than a local minimum. The concentrations of the components were constrained to be greater than zero when doing the fits.

Raman spectra of cells, corrected for system response, were fit to a linear combination of a fifth order polynomial, the five biochemical basis spectra, a spectrum of PBS, and a spectrum of an empty sample cell. The fifth order polynomial was originally meant to represent the fluorescence, although it also compensated for errors in estimating the baseline when analyzing the component spectra. The spectra of PBS and an empty sample cell were averages of multiple measurements made on separate days. It was assumed that the intensity of the signal might change slightly between the low and high wavenumber regions and that the fluorescence amplitudes might be different. These assumptions were made because the sample was realigned after the grating was changed and because during the measurement with the first grating there was some fluorescence bleaching. There were 15 parameters in our fits of Rat1 and Rat1-T1 data; six for the polynomial, five for the basis spectra, one for PBS, one for the empty sample cell, one for the ratio of fluorescence amplitudes in the low and high wavenumber regions, and one for the ratio of the signal amplitude between the low and high wavenumber regions. For the analysis of M1 and MR1 data one change was made in order to correct for some minor baseline artifacts. Separate fourth order polynomials were used for the low and high wavenumber regions instead of the single fifth order polynomial and ratio of fluorescence amplitudes used for the Rat1, Rat1-T1 data. The spectral regions used in the fits were 450–1775 cm^{-1}, and 2600–3125 cm^{-1}, for the Rat1/Rat1-T1 data. For the M1/MR1 data, the lower wavenumber region was changed to 450–1750 cm^{-1} because the signal to noise past 1750 cm^{-1} was poor. The errors of the cell measurements used in computing chi-squared were assumed to be proportional to the square root of the intensity.

Two spectral regions (1011–1575 and 2800–2950 cm^{-1}) were chosen for fitting the FTIR data in order to avoid regions where the high absorption peaks of water greatly decrease the signal to noise. In these spectral regions, the absorbance of PBS is below 2 OD's (for our 50 μm pathlength) except from 1572 to 1575 cm^{-1} where it reaches a maximum of 2.06. Cell absorption spectra were normalized to 10^8 cells/mL. The baseline terms include a spectrum of PBS and a linear polynomial for each spectral region. The errors used for the chi-squared fits were calculated from the standard deviation of multiple cell spectra after normalizing the areas under each curve.

3 Results

Before presenting results of the Raman and IR measurements on tumorigenic and nontumorigenic cells, results are first presented on component spectra, on the fitting method, and on an analysis of the growth stage of the samples. The results of the analysis of the component spectra motivates some of the data presentation methods and provides background information for the fitting method. The section on data fitting provides examples and some evaluation of the method. Finally, the analysis of the growth stage of the cell samples used for vibrational spectroscopy demonstrates that exponential phase cell cultures and plateau phase cell cultures contained very different distributions of cells in the stages of the cell cycle and quantitates the variability of the samples.

Table 2 Similarity of Raman component spectra as quantified by calculation of dot products.

	Protein	Lipid	RNA	DNA	Glycogen
Protein	1.0	0.84	0.56	0.58	0.78
Lipid		1.0	0.37	0.40	0.67
RNA			1.0	0.87	0.58
DNA				1.0	0.61
Glycogen					1.0

3.1 *Component Spectra*

Figures 2 and 3 show the component spectra used in analyzing the Raman and IR data, respectively. These data show significant overlap between component spectra. A simple way to quantify this overlap is to treat the component spectra as vectors in wavenumber space and calculate the normalized dot product between the component spectra. If there is no overlap of bands in the component spectra, the dot product will be 0. If the component spectra are identical the dot product will be one. In other words, the dot product increases from 0 to 1 as the similarity of the spectra increases. Table 2 shows results for the Raman spectra. The greatest overlap is between RNA and DNA, with the next greatest overlap being between protein and lipid. The overlap between lipid and protein is less important because these components are present in large quantities. The results for the IR component spectra are shown in Table 3. Again, the greatest overlap is between the RNA and DNA spectra. However, there is less similarity in the IR lipid and protein spectra than in the corresponding Raman spectra. Given the large overlap between the spectra of RNA and DNA and the fact that, separately, RNA and DNA are relatively minor components, it is difficult to accurately determine their concentrations separately. Consequently, the results of fits of component spectra to cell data are reported in terms of total nucleic acids (i.e., RNA plus DNA).

Table 3 Similarity of infrared component spectra as quantified by calculation of dot products.

	Protein	Lipid	RNA	DNA	Glycogen
Protein	1.0	0.57	0.36	0.35	0.31
Lipid		1.0	0.47	0.48	0.37
RNA			1.0	0.95	0.66
DNA				1.0	0.69
Glycogen					1.0

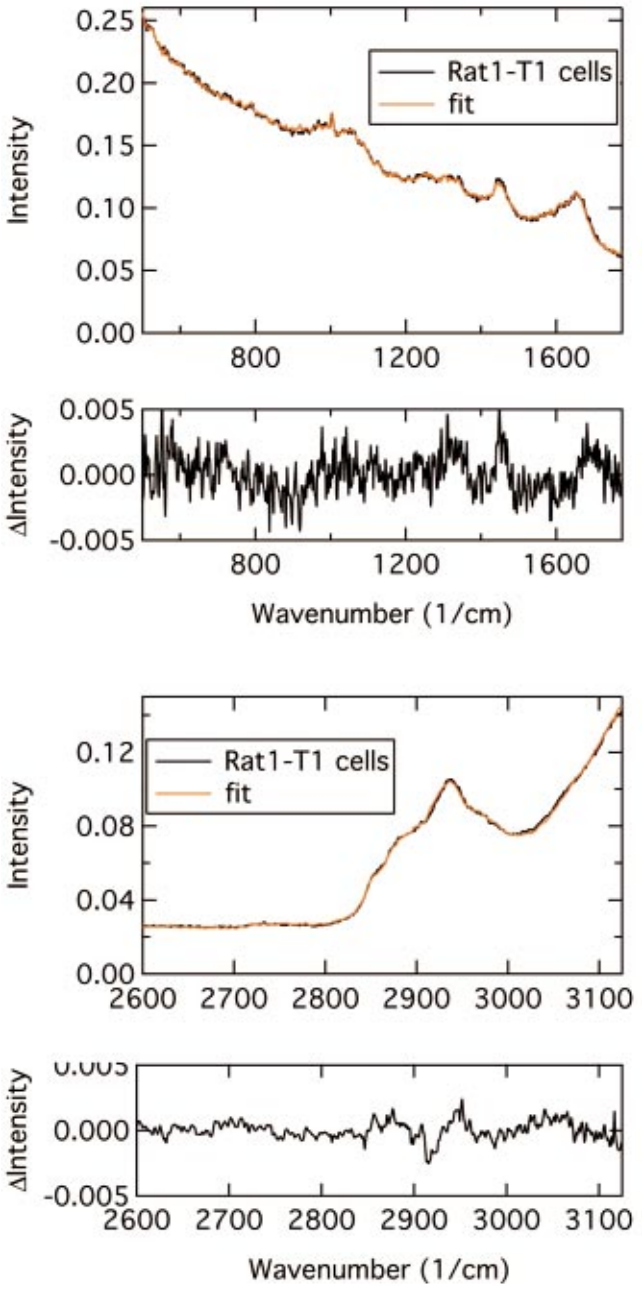

Fig. 4 Top two panels: Example of Raman data for exponentially growing Rat1-T1 cells and the corresponding fit (very top). The average residual for all fits to Rat1-T1e cells is shown below the fit. Bottom two panels: Same data, fit and residual in the high wavenumber region.

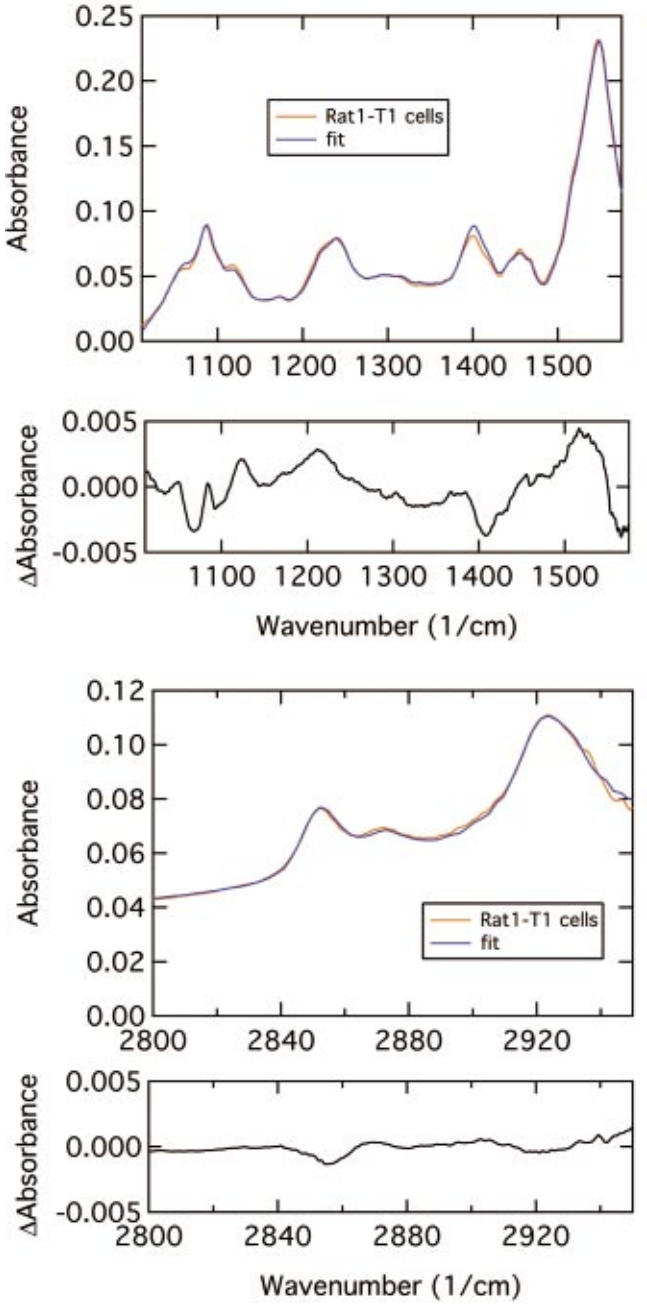

Fig. 5 Top two panels: Infrared data for exponentially growing Rat1-T1 cells and the corresponding fit (very top) and the average residual for the fits to Rat1-T1e cells. Bottom two panels: Same data, fit and residual in the high wavenumber region.

Table 4 Percent of cells in G1, S, and G2 for cell suspensions measured by FTIR and Raman spectroscopy. The mean percent is followed by the standard deviation in parentheses.

	Rat1-T1			
	Raman		FTIR	
	Exponential	Plateau	Exponential	Plateau
G1	50(2.6)	84.4(5.3)	48.5(3.9)	84.4(5.4)
S	37.7(2.6)	8.4(3.9)	34.4(7.4)	8.4(8.9)
G2	12.3(1.4)	7.3(2.0)	16.7(8.9)	7.3(1.9)
	MR1			
	Raman		FTIR	
	Exponential	Plateau	Exponential	Plateau
G1	49.2(10.0)	83.0(4.8)	48.8(8.1)	85.5(3.6)
S	35.9(7.5)	9.9(5.4)	37.3(5.0)	10.1(3.3)
G2	15.7(5.3)	7.1(3.0)	13.8(3.9)	4.3(2.5)
	Rat1			
	Raman		FTIR	
	Exponential	Plateau	Exponential	Plateau
G1	50.7(2.7)	86.2(4.1)	52.1(0.2)	85.6(3.3)
S	35.0(4.3)	10.7(4.4)	33.1(2.7)	9.8(3.8)
G2	14.3(2.5)	3.8(1.7)	14.8(2.8)	5.3(1.9)
	M1			
	Raman		FTIR	
	Exponential	Plateau	Exponential	Plateau
G1	45.6(1.8)	80.0(6.6)	44(3.7)	81.9(8.6)
S	40.8(3.0)	11.0(6.7)	42.1(2.3)	11.0(7.0)
G2	14.2(2.2)	8.9(2.8)	13.8(3.3)	7.0(3.7)

3.2 *Data Fits*

Figure 4 shows a typical Raman spectrum of Rat1-T1 cells harvested in the exponential phase of growth as well as the fit to this spectrum and the average residual of fits to the Rat1-T1e data. The general decreasing slope of the data in the low wavenumber region is due to fluorescence. In the high wavenumber region there is less fluorescence; the rising trend at the highest wavenumbers is due to Raman scattering from water. The fit shows a small peak at 916 cm^{-1} not present in the data. This peak is due to a contribution from the sample chamber to the protein spectrum. The residuals shown below the data plots are the average of the residuals of the fits to the six individual measurements of Rat1-T1e cells. The nonran-

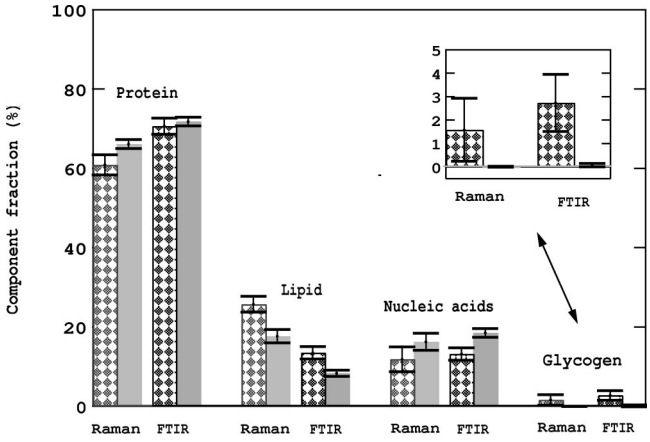

Fig. 6 Percentage of various biochemicals in the nontumorigenic and tumorigenic models. Plateau phase Rat1 cells (gray checkerboard) and exponential phase Rat1-T1 cells (solid gray). Data are means and standard deviations for several independent measurements (see Table 1).

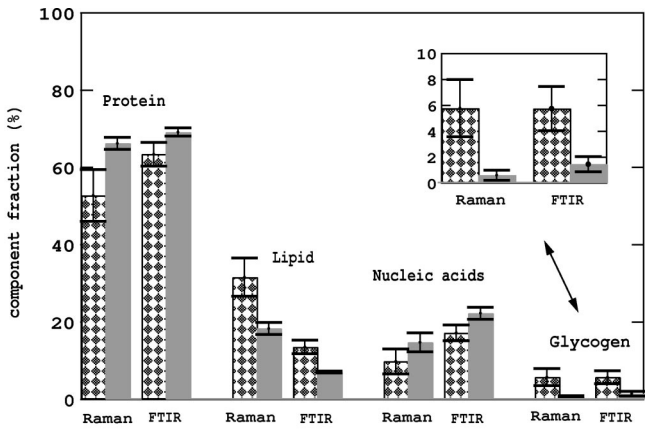

Fig. 7 Percentage of various biochemicals in the nontumorigenic and tumorigenic models. Plateau phase M1 cells (gray checkerboard) and exponential phase MR1 cells (solid gray). Data are means and standard deviations for several independent measurements (see Table 1).

Table 5 *t* Test results for the significance of differences in biochemical composition of the tumorigenic and nontumorigenic models.

	Rat1-T1e versus Rat1p		MR1e versus M1p	
	(Raman)	(Infrared)	(Raman)	(Infrared)
Protein	>99.5%	n.s.	>99.9%	>99.5%
Lipid	>99.9%	>99.9%	>99.9%	>99.9%
Nucleic acids	>95.0%	>99.9%	>97.5%	>99.9%
Glycogen	>95.0%	>99.9%	>99.9%	>99.9%

domness of the amplitudes of this residual indicates that there is a systematic error, albeit a small one since the amplitude is small. The average residuals for the Rat1p, Rat1-T1p, and Rat1e data all show the same general features as the residual shown in Fig. 4 such as the peaks near 1325, 1450, and 1700 cm^{-1} and the dips at 900 and 2920 cm^{-1}. The average residuals of the fits to M1 and MR1 data also show some nonrandom behavior (data not shown). In the low wavenumber region, the residuals show the same peaks at 1325, 1450, and 1700 cm^{-1} along with an additional broad peak near 1000 cm^{-1}, while the dip at 900 cm^{-1} is not present. In the high wavenumber region, there are no distinct peaks like the 2920 cm^{-1} one for the Rat1/Rat1-T1 data. One explanation for the systematic errors in the fits is that there are systematic errors in the component spectra. With the exception of the broad peak near 1000 cm^{-1}, the peaks and dips in the residual spectra are at wavenumbers where Raman scattering from protein is strong or on the edge of the amide I peak. Possibly the spectra of protein used for the fits does not correspond exactly to the spectra of the protein *in vivo*, due to the fact that it is difficult to extract membrane bound proteins and that the environment for the *in vitro* measurements is different than *in vivo* conditions.

Figure 5 shows a typical infrared spectrum of Rat1-T1 cells harvested in the exponential phase of growth as well as the fit to these data and the average residual of fits to the Rat1-T1e data. The average residuals for the Rat1p, Rat1-T1p, and Rat1e data all show the same general features as the residual shown in Fig. 5 such as the peaks near 1530, 1220, and the dips near 1100, 1070, and 2860 cm^{-1}, except that the Rat1e residual does not have a dip near 1400 cm^{-1}. The M1 and MR1 residuals have similar amplitudes and show some of the same features, but there are some differences. There are dips at 1070 and 1100 cm^{-1} and peaks near 1220 and 1530 cm^{-1} as was true for the Rat1 and Rat1-T1 data. There are also peaks near 1460 and 2850 cm^{-1}. Many of these features, such as those at 1400, 1460 and 1530 cm^{-1} can be attributed to protein, however, the dips near 1070 and 1100 cm^{-1} are probably due to nucleic acids or lipids and the peak at 2850 cm^{-1} is most likely lipid. There are probably small discrepancies between the *in vitro* and *in vivo* infrared spectra of most of the components.

3.3 *Cell Cycle Analysis*

Cell cycle analysis of cell cultures in the exponential phase of growth should yield a significant (i.e., ~50%) fraction of cells in the S and G2 phases of the cell cycle. Cell cycle analysis of cell cultures which have reached a growth plateau should have the majority of cells in G1. Therefore, cell cycle analysis was performed to assure that the cell cultures were in the expected phase of growth when harvested. Table 4 shows the mean and standard deviations for the percent of cells in G1, S, and G2 for the measured cell suspensions. Cell suspensions in the exponential phase of growth have about 50% of the cells in G1, 35% in S, and 15% in G2. Cell suspensions in the plateau phase of growth have about 85% of the cells in G1, 10% in S, and 5% in G2.

3.4 *Spectra of Tumorigenic and Nontumorigenic Models*

As models of cancerous versus normal cellular tissue, Rat1-T1 cells harvested in the exponential phase of growth (Rat1-T1e cells) were compared to Rat1 cells harvested in the plateau phase of growth (Rat1p cells) and MR1 cells harvested in the exponential phase of growth (MR1e cells) were compared to M1 cells harvested in the plateau phase of growth (M1p cells). Results of the biochemical analysis obtained by fitting the Raman and FTIR data as described in the methods section are shown in Figs. 6 and 7. The percentage of protein is on average higher in the tumorigenic model, with this difference being greater than the error bars for all the data except the FTIR measurements of Rat1p and Rat1-T1e cells. The lipid content is clearly greater for the nontumorigenic models as there is no overlap of error bars. The nucleic acid content is on average greater in the tumorigenic model, with overlapping error bars for the Raman data, but not for the FTIR data. The differences in glycogen concentration are also significant with no overlapping error bars. The nontumorigenic models have more glycogen than the tumorigenic models. In addition to presenting means and standard deviations in Figs. 6 and 7, we have also performed *t* tests to assess the significance of the differences in the mean values of the biochemical percentages. As shown in Table 5, the confidence levels for the differences in mean values of the concentrations of protein, lipid, nucleic acids, and glycogen are all greater than 95.0% except for the difference in protein concentration for Rat1-T1e and Rat1p cells as measured by FTIR spectroscopy.

3.5 *Source of the Differences Between the Tumorigenic and Nontumorigenic Models*

Figures 6 and 7 demonstrate that there are differences in the biochemical composition of the exponentially growing tumorigenic cells and the plateau phase nontumorigenic cells. These results raise the question of whether the differences are due to intrinsic changes in the cells due to their tumorigenicity or due to their proliferative phase. Comparisons of the same cell line harvested in the exponential and plateau phase show results similar to those seen in Figs. 6 and 7. Figures 8, 9, 10 and 11 compare the results of analyses of exponential and plateau cell spectra for Rat1-T1, MR1, Rat1, and M1 cells, respectively. For the tumorigenic cells, the Raman data show a greater percent of protein in the exponential cells, but the

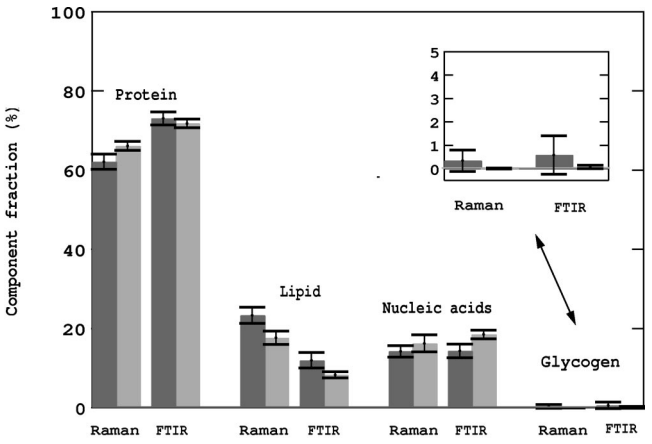

Fig. 8 Comparison of the percentage of various biochemicals in plateau and exponential phase Rat-T1 cell cultures. Plateau phase Rat1-T1 cells (dark gray) and exponential phase Rat1-T1 cells (light gray). Data are means and standard deviations for several independent measurements (see Table 1).

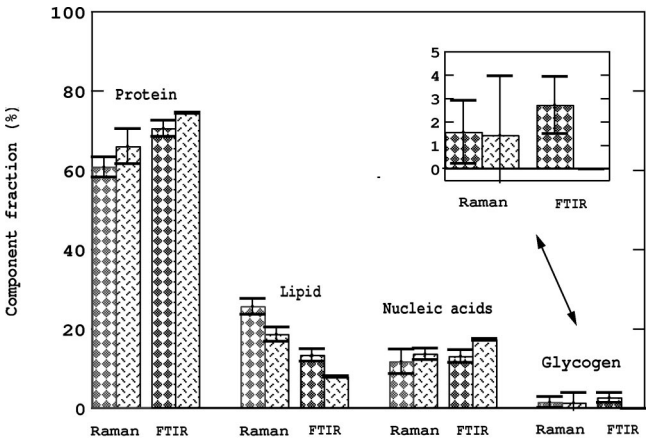

Fig. 10 Comparison of the percentage of various biochemicals in plateau and exponential phase Rat1 cell cultures. Plateau phase Rat1 cells (gray checkerboard) and exponential phase Rat1 cells (gray slashes). Data are means and standard deviations for several independent measurements (see Table 1).

FTIR data do not. For nontumorigenic cells, the protein percentage also appears to be higher in the exponential phase cells, although error bars overlap for the Raman data on Rat1 cells. Lipid percentage is clearly greater for the plateau phase cells, with no overlap of the error bars for either tumorigenic or nontumorigenic cells. The percentage of nucleic acids appears to be greater for the exponential phase cells, although all of the Raman data and the FTIR data on Rat1 cells have overlapping error bars. Glycogen content appears to be greater for plateau phase cells, although in half the cases the error bars overlap. The results of *t* tests of the significance of the differences in the means of the percentages of biochemical components are shown Tables 6 and 7. Most of the significant differences between the biochemical composition of the tumorigenic and nontumorigenic models are also found in Tables 6 and 7.

Comparison of tumorigenic Rat1-T1 cells and nontumorigenic Rat1 cells harvested in the same growth phase are shown in Figs. 12 and 13 with Fig. 12 showing results for exponential cell cultures and Fig. 13 showing results for plateau phase cultures. The differences seen in the means for the plateau phase cultures (Fig. 13), greater percentages of protein and nucleic acids in the tumorigenic cells, and greater lipid and glycogen percentages in the nontumorigenic cells, were also observed in Fig. 6. However, for the exponential phase cultures (Fig. 12), only the changes in nucleic acids appear. The results of *t* tests assessing the significance of differences in the means are given in Tables 8 and 9. Very few of the differences in mean concentrations are significant, particularly for the exponential phase cells. Examination of the data, however, indicates two biochemical component ratios that might differentiate the tumorigenic and nontumorigenic cells. The ratio of lipid to nucleic acid concentration was found to be greater for the nontumorigenic cells for seven of the eight data sets. (The exception was the FTIR data on M1e and MR1e cells.) *t* Tests on the difference in the mean values of this ratio

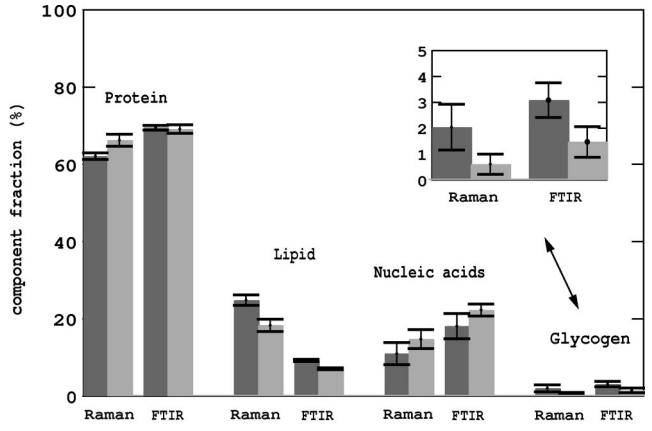

Fig. 9 Comparison of the percentage of various biochemicals in plateau and exponential phase MR1 cell cultures. Plateau phase MR1 cells (dark gray) and exponential phase MR1 cells (light gray). Data are means and standard deviations for several independent measurements (see Table 1).

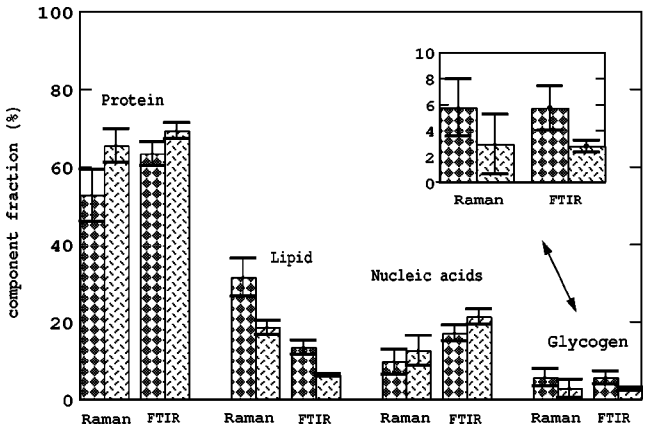

Fig. 11 Comparison of the percentage of various biochemicals in plateau and exponential phase M1 cell cultures. Plateau phase M1 cells (gray checkerboard) and exponential phase M1 cells (gray slashes). Data are means and standard deviations for several independent measurements (see Table 1).

Table 6 *t* Test results for exponential and plateau phase tumorigenic cells. If the confidence level was less than 90%, it was considered not significant (n.s.).

	Rat1-T1e versus Rat1-T1p		MR1e versus MR1p	
	(Raman)	(Infrared)	(Raman)	(Infrared)
Protein	>99.0%	n.s.	>99.9%	n.s.
Lipid	>99.5%	>99.9%	>99.9%	>99.9%
Nucleic acids	n.s.	>99.9%	>95.0%	>97.5%
Glycogen	n.s.	n.s.	>99.0%	>99.5%

were of greater than 90% significance for Raman data on Rat1e versus Rat1-T1e cells, on M1p versus MR1p cells, on M1e versus MR1e cells, and for FTIR data on M1p versus MR1p cells. The ratio of lipid to protein was found to be consistently larger for the plateau phase nontumorigenic cells than for the plateau phase tumorigenic cells. *t* Tests were of greater than 90% significance for Raman and FTIR data on M1p versus MR1p cells and for FTIR data on Rat1p versus Rat1-T1p. Therefore, we can conclude that there are some small differences between tumorigenic and nontumorigenic cells, particularly when they are compared in the plateau phase of growth. In particular, both the lipid to nucleic acid ratio and the lipid to protein ratio can differentiate M1 and MR1 cells in the plateau phase of growth by either Raman or FTIR spectroscopy.

Based on the results presented in the previous paragraphs, it is concluded that the differences in biochemical composition seen between the tumorigenic and nontumorigenic models are primarily due to the difference in proliferative status.

3.6 *Comparison of FTIR and Raman Results*

The previous two subsections of the results section demonstrate that Raman and FTIR spectroscopy give very similar results for the changes in biochemical composition. However, examination of the Raman and infrared results in Figs. 6–13 demonstrates that the lipid concentration is always greater as measured by Raman scattering compared to FTIR

Table 7 *t* Test results for exponential and plateau phase nontumorigenic cells. If the confidence level was less than 90%, it was considered not significant (n.s.).

	Rat1e versus Rat1p		M1e versus M1p	
	(Raman)	(Infrared)	(Raman)	(Infrared)
Protein	>90.0%	>99.0%	>99.5%	>99.9%
Lipid	>99.9%	>99.9%	>99.9%	>99.9%
Nucleic acids	n.s.	>99.5%	n.s.	>99.5%
Glycogen	n.s.	>99.5%	>95.0%	>99.9%

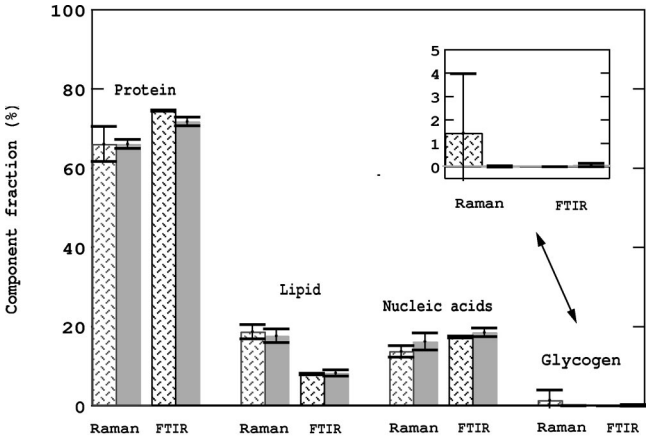

Fig. 12 Comparison of the percentage of various biochemicals in nontumorigenic and tumorigenic exponential phase cell cultures. Exponential phase Rat1 cells (gray slashes) and exponential phase Rat1-T1 cells (solid gray). Data are means and standard deviations for several independent measurements (see Table 1).

spectroscopy. The protein percentage, however, is always lower when measured by Raman scattering rather than FTIR spectroscopy. Similarly there is a trend that the nucleic acid and glycogen percentages are lower for the Raman results than for the FTIR results. The reason for these minor differences lies in the fact that the FTIR absorbance does not depend on scattering properties of the sample due to the very short measurement pathlength, but the measured Raman intensity does. To demonstrate that the Raman results depend on scattering properties, measurements were made of glycogen with and without the addition of polystyrene spheres to increase scattering. The addition of the polystyrene spheres reduced the amplitude of the Raman scattering from glycogen by about a factor of 2 for a reduced scattering coefficient of 1.0 cm^{-1} (glycogen concentration held constant at 20 mg/mL, data not shown). With this information we can explain the minor discrepancies between the FTIR and Raman results. Of

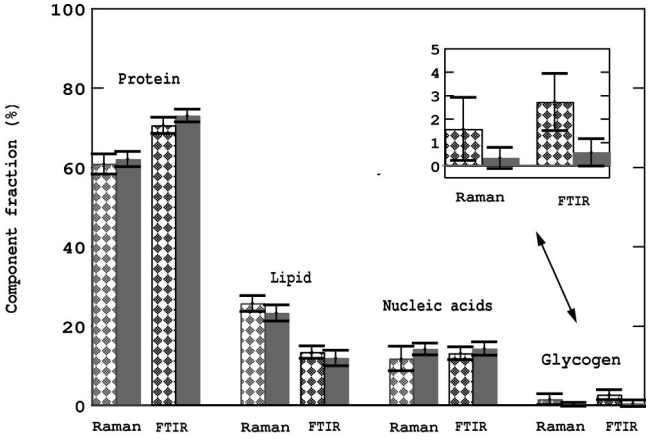

Fig. 13 Comparison of the percentage of various biochemicals in nontumorigenic and tumorigenic plateau phase cell cultures. Plateau phase Rat1 cells (gray checkerboard) and plateau phase Rat1-T1 cells (solid gray). Data are means and standard deviations for several independent measurements (see Table 1).

Table 8 *t* Test results for tumorigenic and nontumorigenic cells harvested in the exponential phase of growth. If the confidence level was less than 90%, it was considered not significant (n.s.).***The Rat1-T1e cells were found to have a lower percentage of protein than the Rat1e cells, unlikely all other cases where the tumorigenic cells had more protein.

	Rat1-T1e versus Rat1e		MR1e versus M1e	
	(Raman)	(Infrared)	(Raman)	(Infrared)
Protein	n.s.	***	n.s.	n.s.
Lipid	n.s.	n.s.	n.s.	>99.5%
Nucleic acids	n.s.	n.s.	n.s.	n.s.
Glycogen	n.s.	n.s.	>95.0%	>99.9%

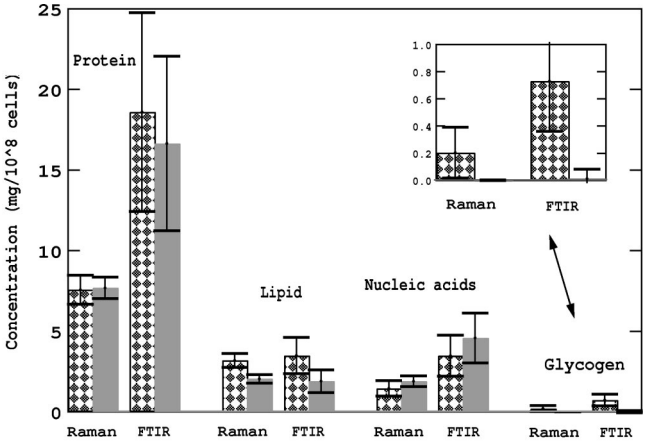

Fig. 14 Concentration of biochemicals in Rat1p (checkerboard) and Rat1-T1e (light gray) cells. The Raman data have been multiplied by a factor of 2.

the biochemical components we measured, only the lipid sample was turbid (i.e., scattering). The cells were also turbid. Therefore, the percentages of protein, nucleic acids, and glycogen were probably underestimated by the Raman spectroscopy measurements.

3.7 *Absolute Concentrations of Biochemical Components*

The results presented so far have all shown relative amounts of biochemical components. FTIR, and, in some cases, Raman spectroscopy can also provide estimates of absolute amounts of the different components. In order to use Raman spectroscopy to determine absolute concentrations of biochemicals, the intensity of the Raman spectra must be linearly proportional to concentration. All of the Raman measurements of Rat1 and Rat1-T1 cells and biochemical components were made using the same measurement configuration. Measurements of pure biochemicals were interspersed with the measurements of cells over a period of months. During this time, the measurements of biochemicals were found to be proportional to concentration. Figure 14 shows the concentration of biochemicals in plateau phase Rat1 cell cultures and exponential phase Rat1-T1 cell cultures. As expected, based on the discussion in the previous paragraph, the lipid results are similar for the two methods, however, the Raman results underestimate the amount of protein, nucleic acids, and glycogen. Despite these differences in results for the two methods, some conclusions can be made regarding cell composition. The amount of protein in 10^8 cells is, to within error bars, the same for Rat1p and Rat1-T1e cells. The amount of lipid and the amount of glycogen, however, is greater in the Rat1p cells. There is also a trend that there is more nucleic acid in a Rat1-T1e cell than a Rat1p cell. Some of these differences can be partially accounted for by cell size; a Rat1-T1e cell is on average 1841 μm^3 in volume, while a Rat1p cell is on average 2005 μm^3 in volume. Therefore, some of the differences in lipid and glycogen amounts are due to cell size. The trend that nucleic acid content is greater in the Rat1-T1e cells cannot be accounted for by a difference in cell size.

Table 9 *t* Test results for tumorigenic and nontumorigenic cells harvested in the plateau phase of growth. If the confidence level was less than 90%, it was considered not significant (n.s.).

	Rat1-T1e versus Rat1p		MR1e versus M1p	
	(Raman)	(Infrared)	(Raman)	(Infrared)
Protein	n.s.	>90.0%	>99.5%	>99.9%
Lipid	n.s.	n.s.	>99.0%	>99.9%
Nucleic acids	n.s.	n.s.	n.s.	n.s.
Glycogen	n.s.	>99.0%	>99.5%	>99.5%

4 Discussion

Our data show that both FTIR and Raman spectra measured on intact, viable cell samples can be well fit to a linear combination of basis spectra comprising the major classes of biochemical components (protein, nucleic acids, lipids and glycogen). To the best of our knowledge, the general idea of fitting infrared spectra to a linear combination of biochemical basis spectra has not been previously implemented by other research groups. However, the general idea of fitting Raman spectra to a linear combination of biochemical basis spectra has been previously implemented to analyze *in situ* spectra of human coronary arteries, breast tissue, and brain tissue. The coronary artery tissue were fit to a linear combination of 12 basis spectra including structural proteins, cholesterols and lipids, and mineral components.[23] This biochemical model described reasonably accurately a range of coronary artery morphologies. In the study of breast tissue, spectra were fit to a combination of collagen, fat, cholesterol/necrosis, calcium hydroxyapatite, calcium oxalate, β carotene, water, a cell nucleus, and a cell cytoplasm spectrum.[24] The Raman study of brain tissue used only four biochemical components, protein (bovine serum albumin), cholesterol, lipids, and water, to model spectra of white matter, gray matter, an astrocytoma, and a meningeoma.[25] The residuals from the fits indicated that while most of the intensity was accounted for when only these

components were used, there was some intensity probably due to other biochemical components. Both coronary artery, and to a large extent breast tissue, can undergo fairly significant changes in morphology, e.g., a change in the amount of fatty tissue in the breast, and, therefore, many of the biologically relevant changes can be modeled without a detailed biochemical analysis of the epithelial cells where most cancers begin. The work described in this paper aims to understand the more subtle biochemical changes at the cellular level accompanying carcinogenesis and facilitate biochemical analysis of largely cellular tissues such as epithelium *in vivo*.

The primary differences in relative biochemical composition between plateau phase nontumorigenic cells and exponential phase tumorigenic cells are greater protein and nucleic acid content in the tumorigenic model and greater lipid and glycogen content in the nontumorigenic model. These changes were the same for both the Rat1/Rat1-T1 system and the M1/MR1 system. The changes were found to primarily result from the change in proliferative status with only minor contributions due to the change in intrinsic tumorigenicity. The fact that cell spectra depend on cell proliferative status has been previously reported.[12,26,27] Also, consistent with a dependence on proliferative status, spectra of mammalian cells have been found to depend on cell cycle.[28,29] In Ref. 29, the authors note that the observed spectral differences seen for the leukemia cells in different stages of the cell cycle are similar to those observed between normal and abnormal (cervical) exfoliated cells. In addition, they state that it appears that many of the differences between normal and abnormal cells noted previously are due to differences in proliferative status. However, no data measuring the potential contribution of proliferative status to the differences previously noted for normal and abnormal exfoliated cells were presented.

The result that there is very little difference in the spectra of tumorigenic and nontumorigenic cells may appear somewhat surprising in light of many papers reporting differences in cancerous and noncancerous cells and tissues.[1–11] The tumorigenic and nontumorigenic cells used in our model only differ by a single gene mutation. *In vivo* there are generally many genetic changes and therefore the intrinsic differences between normal and cancerous cells are larger. In addition to the intrinsic change in tumorigenicity of the cells, there may be other changes ocurring in tissue; for example, a thickening of the epithelium, changes in the stroma underlying the epithelium, and differences in the proliferative status of the cells in the epithelium.

There have been at least four papers published examining tumorigenesis of fibroblast cells with somewhat conflicting results.[30–33] Three of these papers report on fibroblast cells and their malignant counterparts transformed via infection of a MuSV virus.[31–33] Significant spectral differences between mouse primary fibroblasts and their transformed malignant counterparts as well as between primary human fibroblasts and their transformed malignant counterparts were reported.[31,32] The contribution of proliferative status to the spectral differences was considered and it was concluded that the spectral differences appeared to not be due to differences in replication rates of the cells.[32] In a subsequent publication[33] by several of the same authors using the same spectroscopic techniques, primary rabbit fibroblasts were compared to their MuSV transformed malignant counterparts. In this paper, the cells were reported to be synchronized in the G1 growth phase. A cluster analysis of 111 high-signal-to-noise spectra (out of 140 total spectra) yielded complete separation of the spectra of the primary fibroblasts and the transformed malignant fibroblasts. Heterogeneity of the spectra of the primary fibroblasts was found to be larger than that of their transformed malignant counterparts. It was stated that growth stages might be responsible for this higher heterogeneity, although no evidence for this was provided and the cells were reported to be synchronized in G1. In conclusion, the papers on fibroblast cells and their MuSV transformed counterparts give the consistent result that there are differences between the spectra of the primary and transformed cells. However, conflicting statements are made about whether any of the spectral differences are due to growth stage. Finally, infrared spectra of human skin fibroblasts and giant sarcoma cells from the same patient have been compared.[29] Only insignificant differences were found and the spectra of the cell lines display only subtle differences as seen in Fig. 2 of Ref. 29. The authors state that their data supports the hypothesis that IR (micro)spectroscopy monitors the level of cell activity rather than signatures specific to cancer, although no cell cycle or growth stage data are reported in the paper.

Some of the biochemical changes reported in this paper are expected based on cell biochemistry. For example, proliferating cells are expected to have more nucleic acids. Second, the greater ratio of glycogen to protein in cell cultures in the plateau growth phase (Figs. 10 and 11) is consistent with a report of glycogen/protein ratio increasing by three to fourfold in cultured human malignant epithelial cells in the stationary (i.e., plateau) phase[34] and with a negative correlation between the ratio of glycogen to protein and proliferative index previously reported based on FTIR spectroscopy of human colorectal cancer tissue.[35] To further verify the change in glycogen concentration, we have performed independent biochemical analyses of glycogen in all four cell lines. We found that for all four cell lines, there was little or no glycogen in the exponentially growing cultures and that plateau phase Rat1 cells and plateau phase M1 cells contained the most glycogen (data not shown). These biochemical results are consistent with our spectroscopy results.

A further indication that the spectroscopy results are credible is that the results obtained with infrared and Raman spectroscopy are similar despite the fact that the sensitivity of infrared and Raman spectroscopy to biochemical components differs and that the intrinsic errors of the two methods are different. Infrared and Raman spectra of a given component differ because IR and Raman spectroscopy are sensitive to different vibrations. For example, in IR spectroscopy the phosphate groups are the predominate absorbing groups of nucleic acids while the absorption of the nucleic acid bases is relatively weak. In contrast, the intensities of base vibrations and phosphate groups are similar in Raman spectra. Due to these differences in IR and Raman spectroscopy, the overlap of spectral components is different for the two methods and consequently the ability of each method to quantitate the individual biochemical components varies.

Qualitatively, sensitivity to a biochemical component is greatest when that component has one or more strong bands in spectral regions that do not overlap with other components. In the low-frequency infrared region, protein and glycogen have

strong absorption peaks where other components do not absorb. In the high-frequency IR region, lipid has some strong absorption bands with the only interfering component being some weak protein absorption. These facts are consistent with the results in Table 3 in that the dot products of protein, lipid, and glycogen with each other or with the nucleic acids are relatively small. In contrast, there is strong overlap between the spectra of RNA and DNA and their dot product is near 1. Therefore, infrared spectroscopy is expected to be particularly sensitive to lipid, glycogen and protein. In the low-wavenumber Raman spectra, relative peak intensities of the components are similar and there is strong overlap of spectral bands. Exceptions include the very sharp phenylalanine band at $\sim$1000 cm^{-1} and some of the glycogen bands between 800 and 1000 cm^{-1}. For the high-wavenumber region, all components contribute, with lipid being the most distinguishable with a fairly sharp band near 2850 cm^{-1} that only overlaps with the shoulder of a protein band. Dot products between the components are generally greater for the Raman spectra than the IR spectra and in particular the dot product between protein and lipid is greater. Consequently, Raman spectroscopy may not be as good as infrared spectroscopy for quantitating lipid and protein concentration.

A weakness of both IR and Raman spectroscopy is the difficulty in separately quantifying RNA and DNA using either method since the spectra of RNA and DNA are very similar (dot product$>$0.87). Furthermore, both are minor biochemical components and small inaccuracies in spectra of the major biochemical components, where they overlap with DNA and RNA, could lead to erroneous results. Therefore, the potential for inaccuracies in the quantification of RNA and DNA is great. We found that while nucleic acid content was robust to details of the fitting method and that results were similar for infrared and Raman spectroscopy, results for RNA and DNA as individual components were sensitive to small changes in the fitting range and were quite different for Raman and infrared spectroscopy. IR spectroscopy indicated that there was more RNA than DNA in the cells, while Raman spectroscopy results showed more DNA in the cells. Nonetheless, there were trends in the data that made sense, e.g., in nearly all cases analysis showed that the proliferative cells had more DNA than the nonproliferative cells.

The method of fitting vibrational spectra of cellular tissue with biological components (basis spectra) can potentially be applied to *in vivo* measurements, although there will be challenges both in the areas of data acquisition and model refinement. The acquisition of Raman spectra *in vivo* has previously been demonstrated and recent developments in optical probe design have demonstrated the acquisition of good quality *in situ* Raman spectra in 1 s.[36] In the case of infrared spectroscopy, *in vivo* spectra based on evanescent wave spectroscopy have been reported.[37,38] Other methods of acquiring infrared spectra *in vivo* may also be possible, although the tissue depth measured will always be limited to the superficial epithelium due to water absorption. The modifications needed for the data analysis model will depend on tissue type. The lipid, glycogen, DNA and RNA basis spectra used in this paper should still be valid. It is unclear whether the protein basis spectra will be valid for epithelial tissue. In the limited study of related cells reported here, we found that there was no significant difference in the protein spectra of different cells in different growth stages. Spectra of protein isolated from cancerous and noncancerous epithelial cells will need to be measured to determine whether the protein spectra depend on the exact type and carcinogenicity of the cells. When the tissue being measured has a thin epithelium compared to the depth probed by Raman or infrared spectroscopy, spectra of structural tissue components and possibly hemoglobin will need to be added to the model.

The experimental models used in this paper, nonproliferating, nontumorigenic cells and proliferating, tumorigenic cells, were meant to mimic normal and cancerous tissue, respectively. Our results indicate that Raman and IR spectroscopy can differentiate these conditions primarily because of the change in proliferative status. This affect of proliferation needs to be taken into account in any use of vibrational spectroscopy for cancer diagnosis. There are diagnostically challenging *in vivo* situations, such as hyperplastic and dysplastic polyps in which proliferative status is increased in both conditions. The results presented here indicate that when there is only a small genetic change such as a *ras* mutation, differentiation will be difficult. However, in many cancers, there are multiple mutations and IR and Raman spectroscopy may have diagnostic potential in these situations. We are currently extending this approach to epithelial cell models.

Acknowledgments

We would like to thank Lynne Dominique for a measurement of cells on the FTIR, and Anabel Guerra for help with cell cycle analysis and determination of proliferative status. We are also grateful to Steve Doorn and Laura Smilowitz for loaning their lasers while ours was being repaired. The flow cytometry resource at Los Alamos (NIH Grant No. RR01315) provided us with the resources to measure the DNA content of our cells. This work was supported by NIH Grant No. CA89255.

References

1. C. Kendall, N. Stone, N. Shepherd, K. Geboes, B. Warren, R. Bennett, and H. Barr, ''Raman spectroscopy, a potential tool for the objective identification and classification of neoplasia in Barrett's oesophagus,'' *J. Pathol.* **200**, 602–609 (2003).
2. T. C. Bakker Schut, M. J. H. Witjes, H. J. C. M. Sterenborg, O. C. Speelman, J. L. N. Roodenburg, E. T. Marple, H. A. Bruining, and G. J. Puppels, ''In vivo detection of dysplastic tissue by Raman spectroscopy,'' *Anal. Chem.* **72**, 6010–6018 (2000).
3. A. Nijssen, T. C. Bakker Schut, F. Heule, P. J. Caspers, D. P. Hayes, M. H. A. Neumann, and G. J. Puppels, ''Discriminating basal cell carcinoma from its surrounding tissue by Raman spectroscopy,'' *J. Invest. Dermatol.* **119**, 64–69 (2002).
4. D. M. Haaland, H. D. T. Jones, and E. V. Thomas, ''Multivariate classification of the infrared spectra of cell and tissue samples,'' *Appl. Spectrosc.* **51**, 340–345 (1997).
5. N. Stone, C. Kendall, J. Smith, P. Crow, and H. Barr, ''Raman spectroscopy for identification of epithelial cancers,'' *Faraday Discuss.* **126**, 141–157 (2004).
6. S. Sigurdsson, P. A. Philipsen, L. K. Hansen, J. Larsen, M. Gniadecka, and H. C. Wulf, ''Detection of skin cancer by classification of Raman spectra,'' *IEEE Trans. Biomed. Eng.* **51**, 1784–1793 (2004).
7. K. Venkatakrishna, J. Kurien, K. M. Pai, M. Valiathan, N. N. Kumar, C. M. Krishna, G. Ullas, and V. B. Kartha, ''Optical pathology of oral tissue: A Raman spectroscopy diagnostic method,'' *Curr. Sci.* **80**, 665–669 (2001).
8. G. Steiner, A. Shaw, L.-P. Choo-Smith, M. H. Abuid, G. Schackert, S. Sobottka, W. Steller, R. Salzer, and H. H. Mantsch, ''Distinguishing and grading human gliomas by IR spectroscopy,'' *Biopolymers* **72**, 464–471 (2003).

9. M. A. Cohenford and B. Rigas, ''Cytologically normal cells from neoplastic cervical samples display extensive structural abnormalities on IR spectroscopy: Implications for tumor biology,'' *PNAS* **95**, 15327–15332 (1998).
10. M. Huleihel, A. Salman, V. Erukhimovitch, J. Ramesh, Z. Hammody, and S. Mordechai, ''Novel spectral method for the study of viral carcinogenesis in vitro,'' *J. Biochem. Biophys. Methods* **50**, 111–121 (2002).
11. A. Mahadevan-Jansen, M. F. Mitchell, N. Ramanujam, A. Malpica, S. Thomsen, U. Utzinger, and R. Richards-Kortum, ''Near-infrared Raman spectroscopy for the *in vitro* detection of cervical precancers,'' *Photochem. Photobiol.* **68**, 123–132 (1998).
12. J. R. Mourant, Y. R. Yamada, S. Carpenter, L. R. Dominique, and J. P. Freyer, ''FTIR spectroscopy demonstrates biochemical differences in mammalian cell cultures at different growth stages,'' *Biophys. J.* **85**, 1938–1947 (2003).
13. L. A. Kunz-Schughart, A. Simm, and W. Mueller-Klieser, ''Oncogene-associated transformation of rodent fibroblasts is accompanied by large morphological and metabolic alterations,'' *Oncol. Rep.* **2**, 651–661 (1995).
14. I. F. Tannock and R. P. Hill, *The Basic Science of Oncology*, 3rd ed., Chapt. 7, pp. 134–165, McGraw–Hill, New York (1998).
15. J. R. Mourant, M. Canpolat, C. Brocker, O. Esponda-Ramos, T. M. Johnson, A. Matanock, K. Stetter, and J. P. Freyer, ''Light scattering from cells: the contribution of the nucleus and the effects of proliferative status,'' *J. Biomed. Opt.* **5**, 131–137 (2000).
16. J. P. Freyer, ''Mitochondrial function of proliferating and quiescent cells isolated from multicellular tumor spheroids,'' *J. Cell Physiol.* **176**, 138–149 (1998).
17. K. M. Omberg, J. C. Osborn, S. L. Zhang, J. P. Freyer, J. R. Mourant, and J. R. Schoonover, ''Raman spectroscopy and factor analysis of tumorigenic and non-tumorigenic cells,'' *Appl. Spectrosc.* **56**, 813–819 (2002).
18. G. J. Puppels, T. C. Bakker Schut, P. J. Caspers, R. Wolfhuis, M. vaan Aken, A. van der Laarse, H. A. Bruining, H. P. J. Buschman, M. G. Shim, and B. C. Wilson, ''In vivo Raman spectroscopy,'' in *Handbook of Raman Spectroscopy*, I. R. Lewis and H. G. M. Edwards, Eds., pp. 549–574, Marcel Dekker, New York (2002).
19. K. Harzer and B. Kustermann-Kuhn, ''Quantified increases of cholesterol, total lipid and globotriaosylceramide in filipin-positive Niemann-Pick type C fibroblasts,'' *Clin. Chim. Acta* **305**, 65–73 (2001).
20. W. Zhang, B. Asztalos, P. S. Roheim, and L. Wong, ''Characterization of phospholipids in pre-α HDL: selective phospholipid efflux with apolipoprotein A-I,'' *J. Lipid Res.* **39**, 1601–1607 (1998).
21. C. M. Stoscheck, ''Quantitation of protein,'' in *Methods of Enzymology*, Chap. 6, Academic, New York (1990).
22. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C, the Art of Scientific Computing*, Chapt. 15, Modeling of Data, Cambridge University Press, Cambridge (1997).
23. H. P. Buschman, G. Deinum, J. T. Motz, M. Fitzmaurice, J. R. Kramer, A. van der Laarse, A. V. Bruschke, and M. S. Feld, ''Raman microspectroscopy of human coronary atherosclerosis: Biochemical assessment of cellular and extracellular morphologic structures in situ,'' *Cardiovasc. Pathol.* **10**, 69–82 (2001).
24. K. E. Shafer-Peltier, A. S. Haka, M. Fitzmaurice, J. Crowe, J. Myles, R. A. Dasari, and M. S. Feld, ''Raman microspectroscopic model of human breast tissue: implications for breast cancer diagnosis in vivo,'' *J. Raman Spectrosc.* **33**, 552–563 (2002).
25. C. Krafft, S. Miljanic, S. B. Sobottka, G. Schackert, and R. Salzer, ''Near infrared Raman spectroscopy to study the composition of human brain tissue and tumors,'' *Proc. SPIE* **5141**, 230–236 (2003).
26. A. Pacifico, L. A. Chiriboga, P. Lasch, and M. Diem, ''Infrared spectroscopy of cultured cells. II. Spectra of exponentially growing, serum-derived and confluent cells,'' *Vib. Spectrosc.* **32**, 107–115 (2003).
27. K. W. Short, S. Carpenter, J. P. Freyer, and J. R. Mourant, ''Raman spectroscopy detects biochemical changes due to cell proliferation in mammalian cell cultures,'' *Biophys. J* **88**, 4274–4288 (2005).
28. H. N. Holman, M. C. Martin, E. A. Blakely, A. Bjornstad, and W. R. McKinney, ''IR spectroscopic characteristics of cell cycle and cell death probed by synchrotron based Fourier transform IR spectromicroscopy,'' *Biopolymers (Biospectroscopy)* **57**, 329–335 (2000).
29. S. Boydston-White, T. Gopen, S. Houser, J. Bargonetti, and M. Diem, ''Infrared spectroscopy of human tissue. V. Infrared spectroscopic studies of myeloid leukemia (ML-1) cells at different phases of the cell cycle,'' *Biospectroscopy* **5**, 219–227 (1999).
30. P. Lasch, A. Pacifico, and M. Diem, ''Spatially resolved IR microspectroscopy of single cells,'' *Biopolymers* **67**, 335–338 (2002).
31. M. Huleihel, A. Salman, V. Erukhimovitch, J. Ramesh, Z. Hammody, and S. Mordechai, ''Novel spectral method for the study of viral carcinogenesis in vitro,'' *J. Biochem. Biophys. Methods* **50**, 111–121 (2002).
32. V. Erukhimovitch, M. Talyshinsky, Y. Souprun, and M. Huleihel, ''Spectroscopic characterization of human and mouse primary cells, cell lines and malignant cells,'' *Photochem. Photobiol.* **76**, 446–451 (2002).
33. A. Salman, J. Ramesh, V. Erukhimovitch, M. Talyshinsky, S. Mordechai, and M. Huleihel, ''FTIR microspectroscopy of malignant fibroblasts transformed by mouse sarcoma virus,'' *J. Biochem. Biophys. Methods* **5**, 141–153 (2003).
34. M. Rousset, E. Dussauix, G. Chevalier, and A. Zwelbaum, ''Growth-related glycogen levels of human intestine carcinoma cell lines grown in vitro and in vivo in nude mice,'' *J. NCI* **65**, 885–889 (1980).
35. S. Takahashi, A. Satomi, K. Yano, H. Kawase, T. Tanimizu, Y. Tuji, S. Murakami, and R. Hirayama, ''Estimation of glycogen levels in human colorectal cancer tissue: relationship with cell cycle and tumor growth,'' *J. Gastroenterol.* **34**, 474–480 (1999).
36. J. T. Motz, M. Hunter, L. H. Galindo, J. A. Gardecki, J. R. Kramer, R. R. Dasari, and M. S. Feld, ''Optical fiber probe for biomedical Raman spectroscopy,'' *Appl. Opt.* **43**, 542–554 (2004).
37. L. Brancaleon, M. P. Bamberg, T. Sakamaki, and N. Kollias, ''Attenuated total reflection-Fourier transform infrared spectroscopy as a possible method to investigate biophysical parameters of stratum corneum in vivo,'' *J. Invest. Dermatol.* **116**, 380–386 (2001).
38. A. Brooks, N. I. Afanasyeva, V. Makhine, R. F. Bruch, S. F. Kolyakov, S. Artjushenko, and L. N. Butvina, ''New method for investigations of normal human skin surfaces in vivo using fiberoptic evanescent wave Fourier transform infrared spectroscopy (FEW-FTIR),'' *Surf. Interface Anal.* **27**, 221–229 (1999).