Statistical comparison of Fourier transform infrared spectra

Richard G. Spencer
National Institutes of Health
National Institute on Aging
Nuclear Magnetic Resonance Unit
Baltimore, Maryland, 21224

Ericka F. Calton
Nancy Pleshko Camacho
Hospital for Special Surgery
Musculoskeletal Imaging and Spectroscopy Laboratory
New York, New York 10021

Abstract. Spectroscopic assessment of whether a biological sample has changed as a result of processing or degradation is generally carried out by qualitative comparison of spectra, without statistical analysis, resulting in a subjective evaluation of sample stability. Here, we present a formalism for quantitative statistical comparison of signal-averaged Fourier transform infrared spectra, commonly used to assess molecular properties of biological samples. Expressions are derived permitting the comparison of 1. single beam spectra; 2. transmittance spectra obtained by calculating the ratio of single beam spectra of a sample and background; and 3. absorbance spectra derived from transmittance spectra. An application of these results to the degradation of cartilage is presented. Two absorbance spectra of a cartilage sample taken in succession are found to be statistically identical with respect to the ratio of the amplitude of the amide I band to the amplitude of the amide II band. However, a spectrum of the same sample acquired after a 24-h degradation period, while similar to the spectrum of the fresh sample, is found to have an altered ratio of these spectral band amplitudes, consistent with degradation of the cartilage matrix.
 [DOI: 10.1117/1.2393231]

Keywords: infrared spectroscopy; signal-to-noise; statistical comparison; signal averaging; sample stability.

Paper 06030R received Feb. 13, 2006; revised manuscript received Apr. 29, 2006; accepted for publication Jul. 18, 2006; published online Dec. 8, 2006.

1 Introduction

It is axiomatic that comparisons cannot be made between groups composed of a single sample. Even with highly precise measurements, without an estimate of population variance there is no way for a statistical comparison to be made.

However, certain questions pertain inherently to a single sample. In the analysis of biological material, for example, one often needs to determine whether a particular sample has degraded over time, suggesting that it may not be suitable for further studies. Similarly, tissue is often removed from an experimental animal to render subsequent study more convenient, and an attempt is made to design the removal procedure in order not to alter the sample. Note that these questions are very different from whether a particular process tends to degrade samples in general. Even if it is established, through studying a sufficiently large number of samples, that a process is or is not likely to have a certain effect on a statistical basis, the question remains as to whether a given sample has responded in that typical fashion. A related question is determination of criteria for declaring a sample to have changed.

Fourier transform infrared spectroscopy (FTIR) is a sensitive technique for characterizing biological materials. In practice, to determine whether a change has occurred in a sample using FTIR, one acquires spectra under the relevant conditions or at the relevant times and compares them in an *ad-hoc* fashion. Due to a finite signal-to-noise ratio (SNR), it will never be the case that the spectra will be identical. However, if the differences between spectral band amplitudes are "small," and the SNR is "good," the sample is declared unchanged. Otherwise, if the differences are intuitively thought to be too large to be explained by the SNR, it is decided that the sample has changed. We demonstrate that this common procedure can be put into an objective statistical framework when signal averaging is used, which is the typical case.

Our formalism is based on the fact that the final, signal-averaged spectrum is composed of the average of many scans, each of which is assumed to be itself composed of random noise superimposed on a reproducible signal. Therefore, the SNR of the final signal-averaged spectrum provides information about each of these two components, which is what one evaluates intuitively on comparing spectra.

Note that this analysis differs from an attempt to compare groups based on measurements of $N=1$ samples per group; it is measurement uncertainty in the evaluation of a single sample, rather than variability within a population, which is of concern here.

Our approach and results are similar to those of Spencer,[1] but with extensions necessary for application to FTIR. For clarity, we define certain terms used in FTIR. A *single-beam spectrum* is the photon signal amplitude, as a function of

Address all correspondence to Nancy Pleshko Camacho, Research Division, Hospital for Special Surgery, 535 E. 70th St., New York, NY 10021; Tel: 212-606-1435; Fax: 212-472-5331; Email: camachon@hss.edu

1083-3668/2006/11(6)/064023/7/$22.00 © 2006 SPIE

wavenumber, transmitted through a sample. This can be obtained either from a single scan, in the case of very favorable SNR, or, more commonly as the result of signal-averaging multiple scans. Normalization by the number of scans may be performed, but has no effect on relative intensities within the spectrum or on SNR. A *transmittance spectrum* is the wavenumber-by-wavenumber ratio of the single-beam spectra for the sample of interest and the relevant background (e.g., air). Such a spectrum is most often formed from a signal-averaged sample spectrum divided by a signal-averaged background spectrum. An *absorbance spectrum* is a mathematical transformation of a transmittance spectrum:

$$A(\nu) = -\log_{10} T(\nu), \tag{1}$$

where $A(\nu)$ and $T(\nu)$ denote the amplitudes of the absorbance and transmittance spectra at wavenumber ν, respectively.

The calculation previously presented for the case of nuclear magnetic resonance (NMR) spectroscopy[1] is formally identical to the one required for single-beam IR spectra. In the following sections, we 1. recapitulate this calculation in terms of FTIR spectra, and then present the appropriate calculations for 2. transmittance spectra and 3. absorbance spectra.

2 Theory

2.1 *Comparison of Single-Beam Spectra*

By assumption, each scan is composed of reproducible signal of magnitude $B_{NS=1}(\nu)$, where NS, the number of signal-averaged scans in a spectrum, equals one for an individual scan, combined with random noise of standard deviation $\sigma_{B_{NS=1}}$, assumed to be independent of the wavenumber. The SNR of the single scan is given by

$$\mathrm{SNR}_{NS=1}(\nu) = \frac{B_{NS=1}(\nu)}{\sigma_{B_{NS=1}}}, \tag{2}$$

with, in many interesting cases, $SNR_{NS=1}(\nu) < 1$. SNR is labeled by wavenumber to emphasize that SNR is taken with respect to a given spectral band.

Similarly, for a signal-averaged spectrum composed of N scans with non-normalized signal magnitude $B_{NS=N}(\nu)$ and noise $\sigma_{B_{NS=N}}$, the SNR at wavenumber ν is

$$\mathrm{SNR}_{NS=N}(\nu) = \frac{B_{NS=N}(\nu)}{\sigma_{B_{NS=N}}}. \tag{3}$$

Assuming a stable spectrometer and a sample that is unchanging over the time scale of the measurement, $B_{NS=N}(\nu)$ increases in proportion to N and is given by:

$$B_{NS=N}(\nu) = N \cdot B_{NS=1}(\nu), \tag{4}$$

while the noise amplitude increases in proportion to $\sqrt{N}$:

$$\sigma_{B_{NS=N}} = \sqrt{N} \cdot \sigma_{B_{NS=1}}, \tag{5}$$

so that

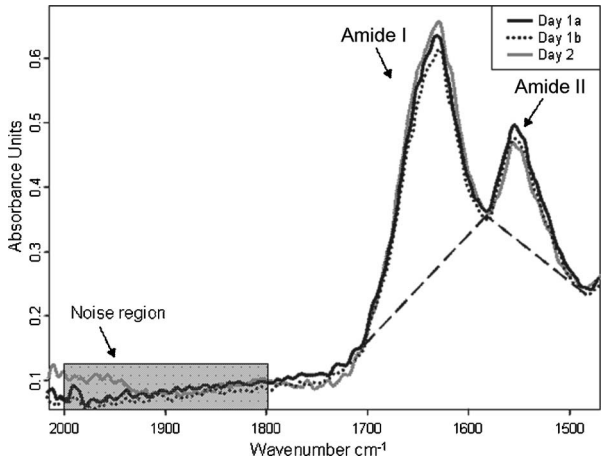

Fig. 1 Infrared fiber optic probe (IFOP) absorbance spectra (ten co-added scans) of cartilage. Two spectra were obtained successively on day 1 (day 1a and day 1b), and another was obtained on day 2 (day 2) after a degradation period. All spectra were acquired from the same position on the sample. Noise amplitude was determined from the wavenumber region 2000 to 1800 cm^{-1}, and the amide I and amide II absorbances were baseline corrected prior to calculation of band height (baseline denoted by dashed line). Statistical comparison showed that there was no difference in the ratio of the amplitude of the amide I spectral band to the amplitude of the amide II spectral band, an indicator of sample stability, between the day 1 spectra, while for the day 2 spectrum, this ratio was significantly different compared to the ratio for both of the day 1 spectra.

$$\mathrm{SNR}_{NS=N}(\nu) = \sqrt{N} \cdot \mathrm{SNR}_{NS=1}(\nu). \tag{6}$$

Note that Eqs. (4) and (5) are both written with the convention that the data in the spectrometer data averager is not renormalized by the number of co-added scans. If renormalization is performed, then the right-hand side of these equations would be divided by N. In this case, the apparent signal amplitude is constant, and the apparent noise decreases in proportion to $\sqrt{N}$. We employ the convention in which renormalization is not performed. Of course, in either case, without or with renormalization, Eq. (6) holds.

It can be difficult or impossible to directly measure $\sigma_{B_{NS=1}}$, since doing so would require measurement of the SNR of a single scan to set the overall measurement scale. Therefore, we wish to define $\sigma_{B_{NS=1}}$ in terms of a readily observed, signal-averaged spectrum. Indeed, from Eqs. (3) and (5),

$$\sigma_{B_{NS=1}} = \frac{B_{NS=N}(\nu)}{\sqrt{N} \cdot \mathrm{SNR}_{NS=N}(\nu)}, \tag{7}$$

corresponding to Eq. (8) of Ref. 1.

Thus, the mean and standard deviation of the ensemble of single-scan spectra comprising an N-scan signal-averaged spectrum can be expressed in terms of quantities observed from that signal-averaged spectrum:

$$B_{NS=1}(\nu) \pm \sigma_{B_{NS=1}} = \frac{B_{NS=N}(\nu)}{N} \pm \frac{B_{NS=N}(\nu)}{\sqrt{N} \cdot \mathrm{SNR}_{NS=N}(\nu)}, \tag{8}$$

analogous to Eq. (10) of Ref. 1.

This expression can be used for statistical comparison of two spectra. However, to avoid difficulties associated with

establishing an absolute amplitude scale, it is common to compare the ratio of the amplitudes of two bands, centered at ν_1 and ν_2, for example, in one spectrum, with their ratio in a different spectrum. The measurement uncertainty of this ratio, expressed as a standard deviation, follows from the formula for propagation of errors applied to a ratio[2]:

$$\sigma_{a/b} = \frac{a}{b}\left[\left(\frac{\sigma_a}{a}\right)^2 + \left(\frac{\sigma_b}{b}\right)^2\right]^{\frac{1}{2}}. \tag{9}$$

We will use propagation of errors, especially in the form of Eq. (9), frequently in what follows. One finds, in terms of quantities observable from the signal-averaged spectrum, that:

$$\frac{B_{NS=1}(\nu_2)}{B_{NS=1}(\nu_1)} \pm \sigma_{\left[\frac{B_{NS=1}(\nu_2)}{B_{NS=1}(\nu_1)}\right]} = \frac{B_{NS=N}(\nu_2)}{B_{NS=N}(\nu_1)} \pm \sqrt{N}\frac{B_{NS=N}(\nu_2)}{B_{NS=N}(\nu_1)}\left\{\left[\frac{1}{\mathrm{SNR}_{NS=N}(\nu_2)}\right]^2 + \left[\frac{1}{\mathrm{SNR}_{NS=N}(\nu_1)}\right]^2\right\}^{\frac{1}{2}}. \tag{10}$$

This is equivalent to Eq. (11) of Ref. 1, and an example of its application is contained therein. Note again that $\sigma_{\left[\frac{B_{NS=N}(\nu_2)}{B_{NS=N}(\nu_1)}\right]}$ and $\sigma_{\left[\frac{B_{NS=1}(\nu_2)}{B_{NS=1}(\nu_1)}\right]}$ are not spectral SNR values, but rather measurement uncertainties derived from such values.

2.2 *Comparison of Transmittance Fourier Transform Infrared Spectra*

We now derive the standard deviation of the noise for a single-scan transmittance spectrum in terms of quantities that can be measured from an N-scan transmittance spectrum, and then use this result to obtain the standard deviation of the ratio of two band amplitudes within single-scan transmittance spectra. With superscripts s and b labeling sample and background, respectively, the magnitude of the single-scan sample and background spectra are denoted by $B^s_{NS=1}(\nu)$ and $B^b_{NS=1}(\nu)$. We define

$$T_{NS=1}(\nu) = \frac{B^s_{NS=1}(\nu)}{B^b_{NS=1}(\nu)}, \tag{11}$$

and similarly for $NS=N$. We wish to obtain the standard deviation of $T_{NS=1}(\nu)$, represented as the noise amplitude in the $N=1$ transmittance spectrum, in terms of the observable SNR of a signal-averaged transmittance spectrum. From Eqs. (4), (5), (9), and (11),

$$\sigma_{T_{NS=1}} = \sqrt{N}\left\{\left[\frac{1}{B^b_{NS=N}(\nu)}\right]^2 \sigma^2_{B^s_{NS=N}} + \left[\frac{B^s_{NS=N}(\nu)}{[B^b_{NS=N}(\nu)]^2}\right]^2 \sigma^2_{B^b_{NS=N}}\right\}^{\frac{1}{2}}, \tag{12}$$

or

$$\sigma_{T_{NS=1}} = \sqrt{N}\left\{\frac{B^s_{NS=N}(\nu)}{B^b_{NS=N}(\nu)}\right\} \times \left\{\left[\frac{1}{\mathrm{SNR}^s_{NS=N}(\nu)}\right]^2 + \left[\frac{1}{\mathrm{SNR}^b_{NS=N}(\nu)}\right]^2\right\}^{\frac{1}{2}}. \tag{13}$$

Note that a direct application of Eq. (9) gives

$$\sigma_{T_{NS=N}} = \left\{\left[\frac{1}{B^b_{NS=N}(\nu)}\right]^2 \sigma^2_{B^s_{NS=N}} + \left[\frac{B^s_{NS=N}(\nu)}{[B^b_{NS=N}(\nu)]^2}\right]^2 \sigma^2_{B^b_{NS=N}}\right\}^{\frac{1}{2}}, \tag{14}$$

so that

$$\sigma_{T_{NS=N}} = \left\{\frac{B^s_{NS=N}(\nu)}{B^b_{NS=N}(\nu)}\right\} \times \left\{\left[\frac{1}{\mathrm{SNR}^s_{NS=N}(\nu)}\right]^2 + \left[\frac{1}{\mathrm{SNR}^b_{NS=N}(\nu)}\right]^2\right\}^{\frac{1}{2}}. \tag{15}$$

Comparing Eqs. (12) and (14) or Eqs. (13) and (15), one finds

$$\sigma_{T_{NS=N}} = \sigma_{\left[\frac{T_{NS=1}}{\sqrt{N}}\right]}. \tag{16}$$

Thus, the noise standard deviation (the *denominator* of the SNR) for a transmittance spectrum actually *decreases*, in proportion to $1/\sqrt{N}$, with signal averaging. This is rather counterintuitive, since noise amplitude in a simple additive random process *increases* with $\sqrt{N}$. Indeed, this result is in contrast to the case for a single-beam spectrum, for which the noise standard deviation *increases* in proportion to $\sqrt{N}$ [Eq. (5)]. The reason for this difference is that while Eq. (15) represents the noise in a calculated transmittance spectrum, it equally represents the uncertainty in measurement of $T_{NS=N}(\nu)$, which has an amplitude independent of NS:

$$T_{NS=N}(\nu) = \frac{B^s_{NS=N}(\nu)}{B^b_{NS=N}(\nu)} = \frac{N \cdot B^s_{NS=1}(\nu)}{N \cdot B^b_{NS=1}(\nu)} = \frac{B^s_{NS=1}(\nu)}{B^b_{NS=1}(\nu)} = T_{NS=1}(\nu). \tag{17}$$

To compare band amplitudes between two transmittance spectra, use can be made of Eqs. (11) and (13) in the same way as Eq. (8). However, it is again more useful to reformulate Eq. (13) in terms of spectral band amplitude ratios.

We therefore wish to calculate the standard deviation of the ratio of the amplitude of a spectral band centered at ν_2 to the amplitude of a spectral band centered at ν_1 in a transmittance spectrum. As before, the goal is to calculate the standard

deviation of the underlying single-scan quantities in terms of observations made from the signal-averaged spectra. We denote the standard deviation of $T_{NS=1}(\nu_2)/T_{NS=1}(\nu_1)$ by $\sigma_{\left[\frac{B_{NS=1}(\nu_2)}{B_{NS=1}(\nu_1)}\right]}$, and similarly for $NS=N$.

From Eqs. (9), (16), and (17),

$$\sigma_{\left[\frac{T_{NS=1}(\nu_2)}{T_{NS=1}(\nu_1)}\right]} = \left\{\frac{T_{NS=1}(\nu_2)}{T_{NS=1}(\nu_1)}\right\}\left\{\left[\frac{1}{T_{NS=1}(\nu_2)}\right]^2\sigma^2_{T_{NS=1}} + \left[\frac{1}{T_{NS=1}(\nu_1)}\right]^2\sigma^2_{T_{NS=1}}\right\}^{\frac{1}{2}}$$
$$= \sqrt{N}\left\{\frac{T_{NS=N}(\nu_2)}{T_{NS=N}(\nu_1)}\right\}\left\{\left[\frac{1}{T_{NS=N}(\nu_2)}\right]^2\sigma^2_{T_{NS=N}} + \left[\frac{1}{T_{NS=N}(\nu_1)}\right]^2\sigma^2_{T_{NS=N}}\right\}^{\frac{1}{2}}$$
$$= \sqrt{N}\left\{\frac{T_{NS=N}(\nu_2)}{T_{NS=N}(\nu_1)}\right\}\left\{\left[\frac{1}{\mathrm{SNR}^T_{NS=N}(\nu_2)}\right]^2 + \left[\frac{1}{\mathrm{SNR}^T_{NS=N}(\nu_1)}\right]^2\right\}^{\frac{1}{2}}. \quad (18)$$

where $\mathrm{SNR}^T_{NS=N}(\nu_1) = T_{NS=N}(\nu_1)/\sigma_{T_{NS=N}}$, and similarly for ν_2 and $NS=1$. Equation (18) is the desired result, permitting the comparison of band ratios between two signal-averaged transmittance spectra. The form of Eq. (18) is identical to the error term in Eq. (10) [or Eq. (11) of Ref. 1] in spite of the fact that Eq. (18) is for transmittance spectra, while Eq. (10) is for single-beam spectra. As in the discussion of the single-beam spectrum, $\sigma_{\left[\frac{T_{NS=1}(\nu_2)}{T_{NS=1}(\nu_1)}\right]}$ is the measurement uncertainty of $T_{NS=1}(\nu_2)/T_{NS=1}(\nu_1)$, rather than an SNR.

2.3 *Comparison of Absorbance Fourier Transform Infrared Spectra*

We now wish to derive the expression for noise amplitude in an absorbance spectrum. From Eq. (1), we have for spectra comprised of any number of scans:

$$|\sigma_{A(\nu)}| = \frac{1}{\ln(10)T(\nu)}\sigma_{T(\nu)}. \quad (19)$$

We denote the single-scan absorbance spectrum by $A_{NS=1}(\nu)$:

$$A_{NS=1}(\nu) = -\log_{10}T_{NS=1}(\nu), \quad (20)$$

and similarly for $NS=N$. Using Eqs. (17), (20), and

$$A_{NS=1}(\nu) = A_{NS=N}(\nu), \quad (21)$$

we have

$$\sigma_{A_{NS=1}} = \frac{1}{\ln(10)}\left\{\left[\frac{\sigma_{B^s_{NS=1}}}{B^s_{NS=1}(\nu)}\right]^2 + \left[\frac{\sigma_{B^b_{NS=1}}}{B^b_{NS=1}(\nu)}\right]^2\right\}^{\frac{1}{2}}, \quad (22)$$

and similarly for $NS=N$. Equations (4) and (5) then lead to

$$\sigma_{A_{NS=N}} = \frac{1}{\sqrt{N}}\sigma_{A_{NS=1}}, \quad (23)$$

as in the case of transmittance spectra.

Defining

$$\mathrm{SNR}^A_{NS=N}(\nu) = \frac{A_{NS=N}(\nu)}{\sigma_{A_{NS=N}}}, \quad (24)$$

and using Eq. (23), we have:

$$\sigma_{A_{NS=1}} = \sqrt{N}\frac{A_{NS=N}(\nu)}{\mathrm{SNR}^A_{NS=N}(\nu)}. \quad (25)$$

The mean and standard deviation of the underlying single-scan quantities in terms of observable signal-averaged quantities are given by Eqs. (21) and (25). However, as for the case of transmittance spectra, it is frequently absorbance band ratios that are compared in practice. From Eqs. (9), (21), and (23), we find

$$\sigma_{\left[\frac{A_{NS=1}(\nu_2)}{A_{NS=1}(\nu_1)}\right]} = \sqrt{N}\frac{A_{NS=N}(\nu_2)}{A_{NS=N}(\nu_1)} \times\left\{\left[\frac{1}{\mathrm{SNR}_{A_{NS=N}}(\nu_2)}\right]^2 + \left[\frac{1}{\mathrm{SNR}_{A_{NS=N}}(\nu_1)}\right]^2\right\}^{\frac{1}{2}}, \quad (26)$$

giving the standard deviation of the ratio of spectral band amplitudes in a single-scan absorbance spectrum in terms of quantities that are observable from the signal-averaged spectrum. An example of the application of Eq. (26) follows.

3 Experimental Application to *Ex Vivo* Cartilage

3.1 *Methods*

3.1.1 *Cartilage samples*

Fresh samples of immature bovine patellar cartilage were stored in saline at −80 °C for 6 weeks prior to data acquisition. At the time of experimentation, samples were thawed and then maintained in saline at 4 °C throughout the protocol.

3.1.2 *Fourier transform infrared spectral data acquisition*

Spectra were recorded using a mid-IR fiber optic probe (IFOP) (Remspec Corporation, Sturbridge, Massachusetts) attached to a Bruker Optics (Billerica, Massachusetts) spectrometer with an external liquid nitrogen cooled mercury cadmium telluride (MCT) detector. The IFOP was coupled to the detector via a flexible 1-m-long fiber optic cable composed of chalcogenide, a mid-infrared transmitting glass. A ZnS attenuated total reflectance (ATR) crystal with a 1-mm-diam flat tip was attached to the end of the fiber optic bundle. ATR data were acquired from 4000 to 900 cm^{-1} with a spectral resolution of four wavenumbers (cm^{-1}) and processed using a Blackman-Harris three-term apodization function and a zero-filling factor of 2. Data were processed using OPUS software (Bruker Optics).

Single-beam sample spectra were collected with the ZnS crystal contacting the cartilage surface after a sixty second period to permit tissue relaxation,[3] while single-beam background spectra were obtained with the crystal in contact with air only. Absorbance spectra were derived from a sample single-beam spectrum and a background single-beam spectrum as described above. Immediately after collection of the first absorbance spectrum, designated day 1a, a second spectrum, day 1b, was collected. A third spectrum, day 2, was obtained from the same position on the sample 24 h later. Each sample and background spectrum was the result of signal averaging NS=10 scans.

3.1.3 *Processing of spectra*

Spectral band positions were identified using the second-derivative function in the OPUS software, and absorbance amplitudes at these wavenumbers were calculated after applying linear baseline correction to the spectra. Absorbance values were measured at $\nu_1=1550\ cm^{-1}$ and at $\nu_2=1630\ cm^{-1}$, corresponding to amide II and amide I vibrations, respectively, which arise primarily from collagen, the main protein component of cartilage.[3] Changes in the amide I and/or amide II absorbance contours can arise from changes in collagen secondary structure or hydration,[4] temperature,[5] enzyme-induced degradation,[6] or from changes in the biochemical composition of the sample.[7] The ratio of these spectral band amplitudes are therefore reflective of sample stability, and were calculated for all spectra. Root mean square (rms) noise was calculated as deviation from a parabolic fit to the baseline over the wavenumber range 1800 to 2000 cm^{-1}. Data are presented as mean ± standard deviation.

3.1.4 *Statistical analysis*

One way analysis of variance (ANOVA) followed by *post hoc* pairwise t-tests using the Bonferroni correction was performed. Statistical significance was defined as $p<0.05$.

4 Results

The spectra used for analysis are shown in Fig. 1, while measured spectral absorbances, obtained after the standard procedure of baseline correction, and SNR values are shown in Table 1. Equations (21) and (26) with $N=10$ were used to determine the required $NS=1$ values:

Table 1 Values of band amplitudes and SNR for cartilage spectra. Results are derived from signal-averaged (ten co-added scans) ATR absorbance spectra of cartilage obtained as described in Fig. 1.

	$A_{NS=10}(\nu_1)$	$A_{NS=10}(\nu_2)$	$SNR^A_{NS=10}(\nu_1)$	$SNR^A_{NS=10}(\nu_2)$
Day 1a	0.1590	0.3447	20.02	43.39
Day 1b	0.1653	0.3577	23.42	50.67
Day 2	0.1262	0.3539	14.22	39.89

$$\text{day 1a: } \frac{A_{NS=1}(\nu_2)}{A_{NS=1}(\nu_1)} \pm \sigma_{\left[\frac{A_{NS=1}(\nu_2)}{A_{NS=1}(\nu_1)}\right]} = 2.17 \pm 0.37,$$

$$\text{day 1b: } \frac{A_{NS=1}(\nu_2)}{A_{NS=1}(\nu_1)} \pm \sigma_{\left[\frac{A_{NS=1}(\nu_2)}{A_{NS=1}(\nu_1)}\right]} = 2.16 \pm 0.32,$$

$$\text{day 2: } \frac{A_{NS=1}(\nu_2)}{A_{NS=1}(\nu_1)} \pm \sigma_{\left[\frac{A_{NS=1}(\nu_2)}{A_{NS=1}(\nu_1)}\right]} = 2.80 \pm 0.66. \quad (27)$$

The one-way ANOVA analysis showed a statistically significant difference among the three ratios. *Post hoc* t-tests showed that the amide I to amide II absorbance ratio calculated from the day 1a spectrum was not significantly different from the ratio calculated from the day 1b spectrum, while the ratio calculated from the day 2 spectrum was significantly different compared to both the day 1a and day 1b values ($p=0.017$ and $p=0.018$, respectively). Therefore, while as expected, there was no statistically significant difference in the amide I to amide II ratio between the spectra collected in immediate succession, this ratio did change after one day of storage, indicative of sample degradation. These results are impossible to discern from simple visual inspection of the spectra shown in Fig. 1.

5 Discussion

The purpose of this study is to present a practical method for statistical comparison of two or more FTIR spectra. The concept behind this is that a large difference in band amplitudes (or amplitude ratios) between two spectra, combined with high SNR, should result in a high probability that the band amplitudes are actually different. On the other hand, if the band amplitude differences are small, and the SNR is limited, the differences may be fortuitous. Our goal is to translate this simple concept into formulae permitting comparison of spectral bands in single-beam, transmittance, and absorbance spectra.

Certain assumptions are inherent in our analysis. We assume that the sample is unchanging and that the spectrometer gain and other characteristics do not vary over the timescale of the measurement. We also assume that the measured noise amplitude is similar to the noise amplitude at the location of the spectral bands under consideration. While imperfect, these are common working assumptions in FTIR spectroscopy. We also assume the basic premises of signal averaging, that is, that signal grows in proportion to the number of scans while

noise grows only as the square root of the number of scans. Again, the validity of these assumptions may be subject to instrumental limitations,[8] and requires that detector noise is the dominant noise source.[9]

We reiterate the distinction between our analysis and what would inevitably be an ill-fated attempt to derive a statistical comparison between groups comprised of single samples, based on their spectra. In general, to determine whether two groups of samples differ by spectral criteria, it is indeed necessary to have a sample size greater than one for each group. This permits statistical assessment of the relevant difference in mean values in light of the standard deviation of the sample variability. However, it is often necessary to compare individual signal-averaged spectra taken from the same sample, for example to determine whether a particular sample has degraded or been affected by treatment. The statistical problem in this case is quite different from a comparison of groups. For each scan of a given signal-averaged spectrum of the sample, the underlying sample component is invariant, under realistic assumptions, while scan-to-scan variability is the result of random noise. If it were possible to observe spectral band amplitudes in these single scans, statistical comparison of two sets of such scans could be performed.[10,11] However, the reason signal averaging is performed is that it is generally impossible to measure band amplitudes in individual scans; such a comparison of sets of individual scans is therefore not possible. Nevertheless, as shown here, a single signal-averaged spectrum contains information about the ensemble of scans of which it is comprised. It is the ability to compare these ensembles, each of which is co-added to create a single spectrum, which permits the comparison of two signal-averaged spectra.

Our development has been motivated by our work in cartilage biology. Such studies are often performed using multiple experimental modalities, necessitating the transport of samples among different laboratories.[12] To determine whether significant changes in particular molecular characteristics of a sample occur, spectra of the sample taken at the outset of the series of studies can be compared to spectra obtained after the studies have been completed. Stability of the sample is confirmed by finding variations in spectral band amplitudes to be sufficiently small that they can be accounted for by finite SNR. The approach described in this work can be used to make this determination more formally.

In addition to the previous application, we routinely perform FTIR analysis of *in-situ* cartilage samples using the IFOP as described above, followed by removal of the sample for subsequent analyses.[13] We record a spectrum before and after removal to determine whether the sample has been affected. Again, an intuitive assessment of whether these spectra are different can now be replaced by the analysis presented here.

Other settings in which this analysis may be particularly useful are those in which SNR is limited. This is seen in the analysis of samples that are available only in limited amounts[14] or when performing highly time-resolved analysis.[14,15] In these cases, the significance of apparent differences in spectral band amplitudes between spectra may be especially difficult to ascertain without a method for formal comparison.

Voigtman[10] and Williams[11] have presented general analyses, not restricted to spectroscopy, permitting the comparison of SNR between sets of measurements. These calculations take as their starting point the known probability density functions (PDF) of the mean and standard deviation of a Gaussian-distributed population. The PDF of the SNR itself, defined as the ratio of these two quantities, is then derived. Tables are generated permitting the desired comparison to be made based on repeated measurements of the SNR. These results permit the appropriate use of SNR as a figure of merit for data collection systems in general, and could, for example, be applied to evaluate the performance of FTIR spectrometers. Our goals are different from the ones in these papers. We define criteria for comparing spectral band amplitudes and ratios between signal-averaged spectra by determining the statistical characteristics of the ensemble of individual scans comprising the spectra; our application is to the usual case in which the individual scans cannot themselves be analyzed quantitatively. We further specify our results for particular realizations of FTIR spectroscopy.

Finally, we note that our results can also be used in the context of a type of power analysis calculation. The analysis permits the determination of the minimum number of scans required to ensure that two spectra will not appear to be statistically different, to a specified degree of statistical certainty, when the underlying band amplitudes are in fact the same.

Acknowledgments

This research was supported in part by the Intramural Research Program of the National Institutes of Health (NIH), National Institute on Aging (RGS), and by NIH R01 EB00744 (NPC).

APPENDIX:

We present some simple related results regarding signal-averaged transmittance spectra. Using Eqs. (16) and (17), the SNR of an N-scan transmittance spectrum is

$$\mathrm{SNR}^{T}_{NS=N}(\nu) \equiv T_{NS=N}(\nu)/\sigma_{T_{NS=N}} = \frac{T_{NS=1}(\nu)}{\sigma_{T_{NS=1}}/\sqrt{N}}, \qquad (28)$$

or

$$\mathrm{SNR}^{T}_{NS=N}(\nu) = \sqrt{N} \cdot \mathrm{SNR}^{T}_{NS=1}(\nu). \qquad (29)$$

This is the well-known result that the SNR, obtained by dividing an N-scan spectrum of the sample by an N-scan spectrum of the background, increases with the square root of N. It is therefore the version of Eq. (6) which is applicable to transmittance spectra.

Further, using Eqs. (23) and (24), we obtain the analogous result for absorbance spectra:

$$\mathrm{SNR}^{A}_{NS=N}(\nu) = \frac{A_{NS=1}(\nu)}{\sigma_{A_{NS=1}}/\sqrt{N}} = \sqrt{N} \cdot \mathrm{SNR}^{A}_{NS=1}(\nu). \qquad (30)$$

Thus, in this context as well, one recovers a result of the same form as Eq. (29). Surprisingly, it is difficult to find Eqs. (29) and (30) derived in the literature.

We also note the SNR properties with respect to signal averaging of another, nonequivalent, method of obtaining transmittance spectra, that is the addition of single-scan transmittance spectra. This in fact also leads to an increase in SNR proportional to the square root of the number of scans averaged. Let p denote the pth realization of a single-scan transmittance spectrum $T_{NS=1}(\nu)$. We wish to derive the noise of the N-scan transmittance spectrum

$$\sum_{p=1}^{N} [T_{NS=1}(\nu)]_p,$$

which we denote by

$$\sigma_{\left[\sum_{p=1}^{N} (T_{NS=1})_p\right]},$$

in terms of the noise of the single-scan ratio spectrum $\sigma_{T_{NS=1}}$. Applying propagation of errors to the sum, we have:

$$\sigma_{\left[\sum_{p=1}^{N} (T_{NS=1})_p\right]} = \sqrt{N} \cdot \sigma_{T_{NS=1}}, \tag{31}$$

with the position of the $\sqrt{N}$ opposite that in Eq. (16). However, the invariant underlying signal component of the spectrum is

$$\sum_{p=1}^{N} [T_{NS=1}(\nu)]_p = N \cdot T_{NS=1}(\nu). \tag{32}$$

Denote the SNR of the co-added single-scan transmittance spectra by $SNR^{T}_{(NS=1)_N}(\nu)$ and the SNR of a single sample scan divided by a single background scan by $SNR^{T}_{(NS=1)}(\nu)$. The relationship between these is, from Eqs. (31) and (32),

$$\mathrm{SNR}^{T}_{(NS=1)_N}(\nu) = \frac{N \cdot T_{NS=1}(\nu)}{\left[\sqrt{N}\sigma_{T_{NS=1}}\right]} = \sqrt{N} \cdot \mathrm{SNR}^{T}_{(NS=1)}(\nu), \tag{33}$$

which is of the same form as Eqs. (6), (29), and (30). In this sense, amplitudes within transmittance spectra resulting from forming the ratio of a single-scan sample spectrum to a single-scan background spectrum, can be regarded as identically independently distributed random variables in the same way as amplitudes within a single-scan single-beam spectrum.

References

1. R. G. Spencer, "Comparing individual signal-averaged spectra," *Magn. Reson. Med.* **40**, 327–329 (1998).
2. P. R. Bevington and D. K. Robinson, *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill, New York (1992).
3. P. A. West, M. P. Bostrom, P. A. Torzilli, and N. P. Camacho, "Fourier transform infrared spectral analysis of degenerative cartilage: an infrared fiber optic probe and imaging study," *Appl. Spectrosc.* **58**, 376–381 (2004).
4. H. Susi, J. S. Ard, and R. J. Carroll, "The infrared spectrum and water binding of collagen as a function of relative humidity," *Biopolymers* **10**, 1597–1604 (1971).
5. A. George and A. Veis, "FTIRS in H2O demonstrates that collagen monomers undergo a conformational transition prior to thermal self-assembly *in vitro*," *Biochemistry* **30**, 2372–2377 (1991).
6. P. A. West, P. A. Torzilli, C. Chen, P. Lin, and N. P. Camacho, "Fourier transform infrared imaging spectroscopy analysis of collagenase-induced cartilage degradation," *J. Biomed. Opt.* **10**, 14015 (2005).
7. N. P. Camacho, P. West, P. A. Torzilli, and R. Mendelsohn, "FTIR microscopic imaging of collagen and proteoglycan in bovine cartilage," *Biopolymers* **62**, 1–8 (2001).
8. B. T. Bowie and P. R. Griffiths, "Measurement of the sensitivity and photometric accuracy of FT-IR spectrometers," *Appl. Spectrosc.* **54**, 1192–1202 (2000).
9. P. R. Griffiths and J. A. De Haseth, "Signal-to-noise ratio," in *Fourier Transform Infrared Spectrometry*, pp. 248–253, Wiley Interscience, New York (1986).
10. E. Voigtman, "Comparison of signal-to-noise ratios," *Anal. Chem.* **69**, 226–234 (1997).
11. R. R. Williams, "Fundamental limitations on the use and comparison of signal-to-noise ratios," *Anal. Chem.* **63**, 1638–1643 (1991).
12. M. Kim, X. Bi, W. E. Horton, R. G. Spencer, and N. P. Camacho, "Fourier transform infrared imaging spectroscopic analysis of tissue engineered cartilage: histologic and biochemical correlations," *J. Biomed. Opt.* **10**, 031105 (2005).
13. G. Li, M. Thomson, E. Dicarlo, X. Yang, B. Nestor, M. P. Bostrom, and N. P. Camacho, "A chemometric analysis for evaluation of early-stage cartilage degradation by infrared fiber-optic probe spectroscopy," *Appl. Spectrosc.* **59**, 1527–1533 (2005).
14. K. Gerwert, "Molecular reaction mechanisms of proteins monitored by time-resolved FTIR-spectroscopy," *Biol. Chem.* **380**, 931–935 (1999).
15. E. Pinakoulaki, T. Soulimane, and C. Varotsis, "Fourier transform infrared (FTIR) and step-scan time-resolved FTIR spectroscopies reveal a unique active site in cytochrome caa3 oxidase from Thermus thermophilus," *J. Biol. Chem.* **277**, 32867–32874 (2002).