Cumulative Examination in Analytical Chemistry

March 4, 2000

S.D. Brown and B. Munson, Examiners

This examination concerns Raman spectroscopy of difficult samples and the signal processing methods needed to extract signal in the presence of background and noise. The two papers


and


provide the background for the examination. Copies of these papers are attached for your use. You are presumed to have read these papers carefully prior to attempting the exam.

Please note ONLY the number of your exam on your blue book. DO NOT put your name on the exam or on the blue book! Enter your name and exam number on the 3x5 note card provided. Place the note card with your name and exam number in the white envelope provided. Return only your blue book to the manila envelope when you have completed the exam.

Each question is scored from 10 points.

1. Both papers concern methods to avoid, reduce or eliminate fluorescence of the sample subjected to Raman spectroscopy. Why is fluorescence a problem when we seek to measure conventional and electronic Raman spectra? Explain using energy level diagrams and a discussion of the instrumentation used to make a Raman measurement.

2. The paper by Maquelin, et al. on measurement of Raman spectra of bacteria measured on growth media uses "NIR confocal Raman spectroscopy." What is the confocal technique, and why do they need the technique to measure bacteria on growth media? What are the benefits obtained from NIR and from confocal microscopy? What would happen if we did conventional (visible) Raman spectrometry on the samples studied here? Would FT-Raman measurements improve the quality of spectra? Why or why not?

3. Even with the effort put into the instrumental approach used to measure spectra of the bacteria, there still seems to a problem with the background signal. Where does the residual background signal come from in this experiment? Why does this background signal have to be removed?
4. The authors report a “vector algebra, double correction method” Use a diagram to show what are they correcting for here and how the correction works.

5. In the end, they get a series of Raman spectra of bacteria. They show that S. aureus, S. epidermis, E. coli and E. faecium differ in their Raman spectra, but that S. aureus ATCC and S. aureus UHR spectra are very similar. What effects are being used to distinguish the bacteria examined here? Should these organisms have different Raman spectra? Why or why not?

6. The second paper (by Bell, et al.) uses Raman spectrometry in a different way. They use UV Raman to examine dyes from the bark of P. amurense on ancient and modern paper. What is the benefit from the UV excitation here? Why would excitation from the more normal lines in the visible or the NIR fail to produce the Raman spectra of the dyes, yet UV excitation succeed? Use an energy level diagram in answering this question.

7. The Raman measurements here also appear on a fluorescence background. In these studies, the fluorescence is compensated by twice shifting by δ the position of the grating of the Raman spectrometer. The Raman lines shift while the fluorescence does not. Why do the Raman lines shift with the grating change? Why not use a change in laser excitation to shift the Raman lines relative to the fluorescence background here?

8. Removal of background is often done by fairly conventional data preprocessing involving a background subtraction step or the fitting of the background with a spline function. Why might this shifted spectrum approach be preferable to the subtraction of a “background paper” spectrum? Why is the spectral shift method preferable to fitting the fluorescence background and removing it?

9. The authors then curve-fit the subtracted Raman spectrum with a “double Lorentzian peak.” They fit both the 18 and the 28 spectra separately. What is the point of this step in the data reconstruction and analysis? Why is it useful to fit both spectra rather than a single subtracted spectrum?

10. The aim of this work is to identify the pigments in artifact paper, yet the spectra of the paper artifacts seem different from those taken on modern paper dyed with the pigments studied, and they also differ from the spectra of the P. amurense bark. How do the authors rationalize the difference seen here? How might this rationalization be tested?
Raman Spectroscopic Method for Identification of Clinically Relevant Microorganisms Growing on Solid Culture Medium

Kees Maquelin,† Lin-Ping Choo-Smith,† Tamara van Vreeswijk,† Hubert Ph. Endtz,† Brian Smith,† Robert Bennett,† Hajo A. Bruining,† and Gerwin J. Puppels*†

Laboratory for Intensive Care Research and Optical Spectroscopy, Department of General Surgery 10M, Erasmus University Rotterdam and University Hospital Rotterdam "Dijkzigt", Dr. Molewaterplein 40, 3015 GD Rotterdam, The Netherlands, Institute of Medical Microbiology and Infectious Diseases, University Hospital Rotterdam "Dijkzigt", Dr. Molewaterplein 40, 3015 GD Rotterdam, The Netherlands, and Spectroscopy Products Division, Renishaw plc, Old Town, Wotton-under-Edge, Gloucestershire GL12 7DW, United Kingdom

Routine clinical microbiological identification of pathogenic microorganisms is largely based on nutritional and biochemical tests. In the case of severely ill patients, the unavoidable delay associated with such identification procedures can be fatal. We present a novel identification method based on confocal Raman microspectroscopy. With this approach it is possible to obtain Raman spectra directly from microbial colonies on the solid culture medium, which have developed after only 6 h of culturing for the most commonly encountered organisms. Due to the limited thickness of microcolonies, some of the underlying culture medium is sampled together with the bacteria. Spectra measured at different depths in a microcolony contain different amounts of the medium signal. A mathematical routine, involving vector algebra, is described for the nonsubjective correction of spectra for variable signal contributions of the medium. To illustrate the possibilities of our approach for the identification of microorganisms, Raman spectra were collected from 6-h microcolonies of five bacterial strains on solid culture medium. The classification results show that confocal Raman microspectroscopy has great potential as a powerful new tool in clinical diagnostic microbiology.

Routine microbiological identification of clinical samples is largely based on nutritional and biochemical characteristics of microorganisms. Following receipt of patient material, microbes are usually first cultured for 16–24 h on solid culture medium. A biomass of 10^6–10^9 cells is then used for biochemical assays, and after or in parallel with the identification assays, an antibiotic susceptibility profile of the microorganism is generated. Consequently, it is usually not until 2–3 days later that the clinician is presented with the full results of this labor-intensive procedure. In critical life-threatening infections, such a delay can be fatal for the patient. Therefore, common clinical practice is to start broad-spectrum empiric antimicrobial therapy based on experience with similar cases before the test results from the microbiology laboratory are known. Early identification of a causative microorganism will enable the clinician to modify and target the initial therapy to the microorganism that causes the specific infection. This practice of streamlining and drug targeting is important to optimize the efficiency of antimicrobial therapy, to reduce the development of drug resistance and other undesired side effects on the microflora. Furthermore, possible toxic effects of broad-spectrum empiric therapy, which sometimes includes a combination of two different antimicrobial agents, can be reduced.

Molecular biological techniques are now being evaluated and used as methods for the identification of microorganisms and the detection of specific antibiotic resistance genes. Although these techniques are potentially rapid, they are relatively expensive and require highly skilled personnel. Problems with false positive reactions due to DNA contamination and false negative reactions due to inhibitors introduced while preparing or collecting a sample are complicating factors in DNA amplification-based molecular diagnostics. At present, molecular diagnostics are usually second lines of investigation and are seldom the sole basis for microbial identification. Therefore, there is a need for new techniques that can rapidly identify pathogenic microorganisms and provide information on drug susceptibility.

---

* To whom correspondence should be addressed. Tel.: +31 10/4635890. Email: Puppels@HLKD.AZNL.
† General Surgery, Erasmus University Rotterdam and University Hospital Rotterdam.
‡ Medical Microbiology and Infectious Diseases, University Hospital Rotterdam.
§ Renishaw plc.

12 Analytical Chemistry, Vol. 72, No. 1, January 1, 2000

10.1021/ac091011h CCC: $19.00 © 2009 American Chemical Society Published on Web 11/10/2009
An alternative approach in microbial characterization is the use of spectroscopic methods. Pyrolysis mass spectrometry has been evaluated as a method for bacterial characterization; however, thus far the high instrument costs have hindered the widespread use of this method. The use of Fourier transform infrared (FT-IR) spectroscopy for microbial identification and characterization has been gaining acceptance since Naumann and co-workers published their pioneering work in this field.\textsuperscript{16-18} Manfait et al.\textsuperscript{20-21} have used FT-IR spectroscopy to identify drug resistance in bacteria, indicating the high information content of this technique.

The application of Raman spectroscopy in microbiology has also been explored previously. Studies have been reported in which FT-Raman\textsuperscript{22,23} or ultraviolet (UV) resonance Raman\textsuperscript{24-26} spectroscopy was used to study suspensions and dried films of microorganisms as well as hydrated microbial smears taken from a solid culture medium.\textsuperscript{27} Using UV resonance Raman spectroscopy to study bacterial suspensions and bacterial cell constituents, Nelsen and Sperry were able to identify microorganisms on the basis of their Raman spectra.\textsuperscript{28} They also reported measurements with only very small numbers of cells (1–50).\textsuperscript{29,30} However, the application of UV resonance Raman spectroscopy requires the cells to be suspended in liquid medium in order to avoid damage due to heating\textsuperscript{31} and photochemical effects as a result of the strong absorption of UV radiation by nucleic acids and proteins. For the clinical application of Raman spectroscopy targeted at rapid identification of pathogenic microorganisms, this is not a practical solution.

\textsuperscript{19} Zerouni, W.; Manfait, M.; Cholay, C. \textit{Pathol. Biol.} 1995, 43 (4), 300–305.

\textbf{Figure 1.} (A) Diagrammatic representation of the measurement volume \(v\) during sampling of a bacterial microcolony \(b\) on a solid culture medium \(m\). (B) Vector diagram of the background subtraction routine depicting the orthogonal vectors of the medium \(M\) and bacteria \((B)\) and the vector of the combined bacteria and medium signal \((C)\). See text for details.

For our studies, we have chosen to use near-infrared (NIR) multichannel confocal Raman microspectroscopy. The use of NIR laser light minimizes the excitation of sample autofluorescence, which tends to mask the much weaker Raman signal when using visible light excitation. Moreover, the use of a confocal signal detection scheme enables Raman spectroscopic measurements of very small sample volumes (even down to about 1 \(\mu\)m\(^3\)).\textsuperscript{32,33} Bacteria are therefore required to be cultured only until microcolonies are formed; microcolonies here are defined as colonies that develop in 6 h of growth after plating and have average colony diameters of 10–110 \(\mu\)m (depending on the type of microorganism). Raman spectra can be directly acquired from the microcolonies on solid culture media. When performing measurements on microorganisms still growing on the solid culture medium, there are minimal sample preparation steps prior to spectral acquisition. Since spectra with good signal-to-noise ratios can be obtained of microcolonies within a few minutes of signal collection time, this Raman spectroscopic approach offers the potential for rapid identification of microorganisms. However, a major obstacle of this approach is the presence of signal contributions from the underlying culture medium in the bacteria Raman spectrum due to the limited thickness of the microcolonies (Figure 1A). Because the medium signal contribution is neither negligible nor constant, it will interfere with strain identification. We have developed a new approach to deal with this problem. The method does not aim to subtract the exact amount of signal that is contributed by the culture medium. Instead, we subtract all signal contained in the combined microorganism and culture medium spectrum that is indistinguishable from the culture medium spectrum.

In this paper, the methodological aspects of obtaining and analyzing Raman spectra of microorganisms directly on solid culture medium are discussed. In so doing, confocal Raman microspectroscopy can be developed for the rapid, routine identification and characterization of microorganisms.

\textbf{MATERIALS AND METHODS}

\textbf{Sample Preparations.} The various bacterial strains used in the studies were derived either from the American Type Culture Collection (ATCC) or from the collection of the Department of Medical Microbiology and Infectious Diseases of the University of


\textsuperscript{A} Analytical Chemistry, Vol. 72, No. 1, January 1, 2000 13
Hospital Rotterdam (UHR- and BM-labeled strains). Strains were stored at −80 °C in a brain-heart infusion broth (Becton Dickinson, Franklin Lakes, Nj) containing 10% glycerol until use. Following an overnight passage (37 °C) on Mueller-Hinton (MH) medium (Merck, Darmstadt, Germany), the strains were recultured on MH medium for 6 h at 37 °C prior to Raman measurement of microcolonies.

For studies involving bacterial smears on CaF2 substrate, a second overnight (16 h) culturing step was performed on MH medium. From these overnight colonies, a biomass from several well-isolated colonies was picked up using an inoculating loop and smeared onto a CaF2 substrate. The samples were allowed to dry in air prior to the Raman measurements.

Ribonucleic acid (RNA) from baker’s yeast (Sigma, St. Louis, MO) was dissolved in water to a concentration of 80 ng/mL prior to measurement.

Raman Measurements. Raman spectra were acquired using a Renishaw System 1000 Raman microspectrometer (Renishaw plc, Gloucestershire, UK). The accompanying Leica DM-IM microscope was fitted with an 80× near-infrared objective (MR Plan 80×/0.75, Olympus). The spatial resolution of the setup was determined to be approximately 1.5 μm in the lateral direction and 7–8 μm along the optical axis. This depth resolution is to a larger degree, dictated by the entrance slit width of the spectrometer. The spectrometer was equipped with a 300 lines/mm grating. Raman signal was collected in the spectral interval from 250 to 2150 cm⁻¹, with a spectral resolution of 8 cm⁻¹. Raman measurements were performed using 830-nm excitation from a titanium–sapphire laser (model 9900, Spectra Physics, Mountain View, CA) pumped by an argon ion laser (series 2000, Spectra Physics), delivering 100 mW of laser power on the sample.

The constant background signal contribution originating from optical elements in the laser light delivery pathway was subtracted from all spectra. The reference spectrum of a tungsten band lamp of known temperature was used to correct for the wavelength-dependent signal detection efficiency of the Raman setup.34,35

Correction for Background Medium Signal Contribution. A nonsubjective method was developed to subtract Raman signal contributions of the culture medium from spectra obtained from bacterial microcolonies growing on the culture medium. This procedure involves the use of vector algebra. In mathematical terms, a Raman spectrum of bacteria on culture medium consisting of n data points can be thought of as a vector in an n-dimensional space. Similarly, the spectrum of the culture medium alone can be thought of as another vector in this n-dimensional space. The combined bacteria and medium vector can now be decomposed into a vector parallel to and a vector orthogonal to the medium vector (M) (Figure 1B). (Throughout the text, uppercase boldface type is used to denote vectors). The projection of the “combined signal vector” (C) of bacterial and medium signal on M (projM) gives the amount of signal in C that cannot be distinguished from M. Subsequent subtraction of this projection from C results in the desired non-medium-related bacteria spectrum (B) (i.e., the vector component of C orthogonal to M, see eq 1). When eq 1 is elaborated in terms of the dot product of C and M, we obtain eq 2:

\[
B = C - \text{proj}_M C \quad (1)
\]

\[
B = C - \frac{C \cdot M}{||M||^2} M \quad (2)
\]

A similar approach for spectral subtraction was described earlier by Berger et al.36 for the subtraction of a pure component spectrum from the spectrum of a mixture, containing that component, as a first step in a linear multivariate calibration algorithm.

In the case of measurements on fully hydrated microcolonies, a second variable that can interfere with microorganism identification is the water concentration in the measuring volume. Since fluctuations in ambient temperature and humidity levels can influence the water content of the culture medium, and hence the water signal contribution in the Raman spectra, it is necessary to eliminate this variable as well. The vector correction routine we have developed therefore involves a “double-correction” approach. This procedure starts with independently correcting both the combination spectrum (i.e., the spectrum of bacteria and medium) and the medium spectrum for water signal contributions by subtracting their respective projections on the water vector (eqs 3 and 4, projw is the vector projection on the water vector, W). The two resulting spectra are then used in the next vector correction step to actually correct the combination spectrum for the medium signal (eq 5):

\[
C' = C - \text{proj}_W C \quad (3)
\]

\[
M' = M - \text{proj}_W M \quad (4)
\]

\[
B' = C' - \text{proj}_M C' \quad (5)
\]

In the work presented here, this method was applied to first derivative spectra instead of the actual measured spectra. In most cases, Raman spectra of biological molecules, cells, or tissues contain a broad, relatively featureless background signal, usually ascribed to fluorescence. Its intensity and shape may vary somewhat (and sometimes quite considerably) from measurement to measurement and from strain to strain. When the method described above is applied to spectra that differ only in their Raman-to-fluorescence background signal ratio, the resulting “corrected” spectra would differ. Therefore, the fluorescence background signal contribution needs to be eliminated before application of the vector correction method. This is achieved by making use of first derivative spectra.

It is important to note that the vector correction procedure is not the same as subtraction of the exact amount of signal contributed by the medium. When a pure bacteria spectrum (i.e., without medium signal contribution) is decomposed into vectors parallel and perpendicular to the medium vector, the parallel


component would normally not be (exactly) zero. This implies that the result of the vector correction method described is not the exact "pure" bacteria spectrum but the component of the bacteria spectrum that is orthogonal to the medium vector. Therefore, this subtraction method may affect the possibilities of a precise biochemical interpretation of the spectra. However, since the components of the microorganism spectrum that are parallel to the medium and/or water vector cannot be distinguished from the actual signal contributions of water and medium (whose intensity varies from measurement to measurement), they cannot be considered as useful information when it comes to microorganism identification.

A schematic of the whole vector correction procedure is given in Figure 2. The code for the calculations in this scheme was developed under the Matlab software package (The Mathworks Inc., Natick, MA).

Validation of the Medium Subtraction Method. The application of the vector correction method to nonsubjectively subtract background signal contributions is expected to result in reproducible spectra regardless of the amount of medium and water signal initially present in the raw data. This method was tested in several ways.

First, a simulation was performed, in which the measurement of spectra at different depths within a microcolony on solid culture medium (and therefore with varying medium signal contributions) was mimicked. The resulting spectra were subsequently subjected to vector correction in order to remove water and culture medium signal contributions. Second, actual measurements at different depths in a microcolony growing on a solid culture medium served as an illustration of the practical situation. Finally, the effect of separately correcting for water signal contributions in addition to medium signal correction with the double-correction procedure was investigated.

(1) Simulation. To simulate measurements at different depths in a bacterial colony, a pure bacteria spectrum (i.e., without culture medium signal contributions) and a spectrum of the culture medium were added in different ratios. Raman spectra of dried bacterial smears on CaF$_2$ substrate served as the pure bacteria spectra, since in the spectral region of interest CaF$_2$ has no Raman features. Raman spectra were collected of a dried bacterial smear of *Escherichia coli ATCC 25922* on CaF$_2$ of Mueller–Hinton culture medium, and of water. Raman spectra of the bacterial smear and of the culture medium were obtained at random locations in the sample, comprising 60 min total signal collection time. The same signal collection time was used to obtain a spectrum of water. The Raman spectrum of the culture medium used for subtraction in the vector correction routine was also obtained with a high signal-to-noise level. The spectra were scaled to standard normal variance (SNV; i.e., zero mean and unit variance).\(^{37}\) The bacterium and medium spectra were then added in the ratios of 1:1, 1:5, and 1:10. First-derivative spectra were calculated for the combination spectra, the second MH spectrum, and the water spectrum, followed by cutting all the spectra to the region of interest, 400–1800 cm$^{-1}$ (schematic representation in Figure 2). These combination spectra were corrected for the water and medium signal contributions by an independently measured culture medium spectrum, as described above. Having a second independent spectrum parallels the scenario of correcting acquired microcolony spectra with a reference culture medium spectrum.

(2) Depth Measurements. To illustrate the validity of the vector correction method on spectra of bacterial samples growing on culture media, Raman spectra were acquired directly from a 6-h microcolony (*E. coli ATCC 25922*) with the laser focused at three depths within the microcolony. Spectra were acquired from the top (2 μm below surface), middle (4 μm below surface), and bottom (10 μm below the surface) of the microcolony, each in 5 min signal collection time. Water and medium Raman spectra were obtained in 60 min signal collection time as described above. A Raman spectrum of RNA was obtained in 30 min signal integration time. Data treatment was performed as described in Figure 2.

(3) "Double Correction". The effect of correcting for both water and medium signals was illustrated by culturing *E. coli ATCC 25922* on three media that differed only in their water content. Separately, 2.25, 3.8, and 5.6 g portions of dehydrated MH medium (Difco Laboratories, Detroit, MI) were dissolved in 100 mL of distilled water to prepare the solid culture media. Raman spectra of water and the three media were collected as described above. Spectra were treated as outlined in Figure 2, and for illustration purposes, the water correction step was initially omitted.

Identification/Classification of Bacterial Strains. Spectra were obtained of 6-h microcolonies of five bacterial strains (*Staphylococcus aureus ATCC 29213, S. aureus UHR 28624, Staphylococcus epidermidis UHR 29489, E. coli ATCC 25922, and Enterococcus faecalis BM 4147*). These measurements were performed in triplicate on separate days. Measurements were carried out on five microcolonies per bacterial strain, with five spectra obtained from various positions within each microcolony. For each measurement, Raman signal was collected for 30 s. Raman spectra of the Mueller–Hinton culture medium were obtained at random locations in the medium, comprising 60 min signal collection time. A water spectrum was also obtained in 60 min total signal collection time.

Vector corrections for water and medium signal contributions were performed as described above. Per strain, the five spectra collected from each microcolony were averaged. The complete

spectral range (400–1800 cm\(^{-1}\)) of the microcolony spectra was then used in the subsequent multivariate analysis. Data from all 3 days were combined and analyzed as one data set.

For multivariate analyses, the amount of data was first reduced using principal component analysis (PCA)\(^{(38)}\) performed using the Matlab PLS toolbox (Eigenvector Research Inc., Manson, WA). A total of \(n - 1\) PCA scores were calculated (\(n\) being the number of spectra in the analysis), typically accounting for 99–100% of the variation in the data set. These PCA scores were used in a cluster analysis (SPSS, Chicago, IL). Ward’s clustering algorithm method and squared Euclidean distance measure were used in generating the dendrogram of the hierarchical cluster analysis.

Linear discriminant analysis (LDA) was performed on principal component scores using SPSS. Two-thirds of the data was used as a training set and one-third as a test set. Selection of principal components that were included in the LDA model was based upon Wilk’s Lambda\(^{(39)}\) method and 95% \(F\)-test inclusion criterion,\(^{(38)}\) thus maximizing group separation. The strength of the model based on the training set was evaluated using the leave-one-out method.\(^{(39)}\)

**RESULTS AND DISCUSSION**

The Raman spectrum obtained directly from a bacterial microcolony on solid culture medium contains signal contributions from both bacteria and the culture medium. The relative signal contribution of the culture medium will vary, as it critically depends on the exact depth at which the laser light is focused as well as on the thickness of the microcolony. Since the medium signal contribution is not negligible and is variable, it would interfere with multivariate analysis intended for strain identification.

**Correction for Signal Contributions of the Culture Medium.** In Figure 3A, a spectrum of a dried smear of bacteria and a spectrum of medium from a culture plate are shown. The spectrum of the medium does not contain any clear marker bands that could be used to determine the exact intensity of the medium signal contribution. Therefore, it is not possible to accurately and nonsubjectively subtract the medium signal contribution on the basis of the intensity of such a band. For this reason we have developed the vector correction method described in the materials and methods section, which was evaluated in three experiments.

As explained in the Materials and Methods section, the result of the vector correction method is a spectrum that, when viewed as a vector, is perpendicular to the medium spectrum (i.e., not only medium signal is subtracted but also the component of the bacteria spectrum that is parallel to the medium spectrum). However, for the purpose of readability we will, from here on, refer to the action of the vector correction method as “subtraction of medium signal”.

**Simulated Experiment.** The two spectra of Figure 3A were used to simulate an experiment in which measurements are carried out at different depths within a microcolony on solid culture medium. To achieve this, the bacteria spectrum and a medium spectrum were co-added in various ratios (Figure 3B) (see Materials and Methods section for details). Spectral features belonging to the bacteria become less obvious when the fraction of the medium is increased in the combination spectrum. For example, there was an increasing sloping background contribution due to the greater water content of the medium. In the spectra with higher proportions of medium added in, the water contribution was also more noticeable by a marked broadening of the 1550–1700 cm\(^{-1}\) region. This spectral region overlaps with Raman features predominantly arising from COO\(^{-}\)amide backbone groups within proteins. In addition, various other spectral changes are observed around 1300 and 1000 cm\(^{-1}\) and between 800 and 900 cm\(^{-1}\).

Using a second medium spectrum, the vector correction method was used to subtract out the corresponding signal contributions. This routine resulted in bacteria spectra with high similarity, illustrating that, with the correction scheme, the background signal could be reproducibly subtracted (Figure 3C). There were no residual features of the medium following vector correction.

**Measurement on Microcolony Growing on Solid Culture Medium.** Having demonstrated that the vector correction procedure is able to nonsubjectively subtract medium signal contributions to yield similar bacteria spectra, Raman spectra were acquired at different depths within an actual bacterial microcolony (E. coli ATCC 25922). We observed that, at the various focusing depths, there are clear differences in the untreated spectra (Figure 4A). Following subtraction of water and medium signal contributions, the spectra were very similar (Figure 4B). However, it was noticed that there were still some spectral differences between spectra taken at different depths within the microcolony, as became clear when the difference was taken between the spectra acquired from the top and from the bottom of the microcolony (Figure 4C). Also

---

**Figure 3.** Validation of the vector correction routine. Simulation of measurements at different focusing depths within a bacterial microcolony. (A) Raman spectrum of E. coli ATCC 25922 bacterial smear on CeFe substrate (s) and Raman spectrum of Mueller-Hinton medium (m) to be used for the simulation. (B) Spectra obtained by adding the bacteria spectrum and a medium spectrum in the ratios (a) 1:1, (b) 1:5, and (c) 1:10. (C) Resulting bacteria spectra (a’, b’, c’) after vector correction for water and medium signal contributions. (au = arbitrary units.)
Table 1. Tentative Wavenumber Assignments of Some of the Raman Features in the Spectra Presented*  

<table>
<thead>
<tr>
<th>wavenumber (cm⁻¹)</th>
<th>tentative assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>533</td>
<td>δ(COC) glycosidic ring</td>
</tr>
<tr>
<td>668</td>
<td>ν(CS)</td>
</tr>
<tr>
<td>725</td>
<td>ρ(CH₃)</td>
</tr>
<tr>
<td>749</td>
<td>thymine (T) ring</td>
</tr>
<tr>
<td>780</td>
<td>uracil (U) ring</td>
</tr>
<tr>
<td>810</td>
<td>C–O–P–O–C in A-RNA backbone</td>
</tr>
<tr>
<td>856</td>
<td>ν(CC)</td>
</tr>
<tr>
<td>1004</td>
<td>ν(COC) 1,4-glycosidic link</td>
</tr>
<tr>
<td>1095</td>
<td>ν(CC) aromatic ring (Phe)</td>
</tr>
<tr>
<td>1220–1290</td>
<td>amide III</td>
</tr>
<tr>
<td>1334</td>
<td>δ(CD)</td>
</tr>
<tr>
<td>1452</td>
<td>δ(CH₂)</td>
</tr>
<tr>
<td>1573</td>
<td>δ(NH) and ν(CN), amide II</td>
</tr>
<tr>
<td>1630–1680</td>
<td>amide I</td>
</tr>
</tbody>
</table>

*δ, Deformation; ν, stretching; ρ, rocking; Phe, phenylalanine (refs 22 and 23).

Figure 4. Raman spectra of E. coli ATCC 25922 measured at various focusing depths: t (top) = 2 μm, m (middle) = 4 μm, and b (bottom) = 10 μm, below the surface of the microcolony. (A) Spectra obtained at the various depths before vector correction. (B) Spectra (t’, m’, b’) following vector correction for water and the underlying Mueller–Hinton culture medium signals. (C) Difference spectrum between the vector-corrected spectra acquired from the top and the bottom of the microcolony (t’–b’) and the Raman spectrum of a solution of ribonucleic acid (RNA) after a water spectrum has been subtracted. (au = arbitrary units.)

depicted in Figure 4C is a spectrum of a solution of RNA from which a water spectrum has been subtracted. Comparison of the difference spectrum with this RNA spectrum reveals many similarities, suggesting that different RNA levels account for the differences observed. However, not all of the features in the difference spectrum can be accounted for by RNA bands alone. Among others, the difference spectrum also shows peaks of thymine (749 cm⁻¹) and phenylalanine (1004 cm⁻¹) (Table 1), indicating the presence of more biological differences. Hence, these spectral differences observed are not artifacts of the correction routine. They arise from the naturally occurring biochemical heterogeneity of the bacterial colonies. One explanation could be that, even in microcolonies with a thickness of 6–8 μm, the various layers contain bacteria in different growth stages (older versus younger cells), and these differences are reflected in changes in the Raman spectrum. Manoharan et al. reported increased RNA levels in bacteria entering the logarithmic phase. We speculate that the cells in the higher layers of the microcolony are more actively dividing than cells in the deeper layers. Further investigations into microcolony heterogeneity are currently underway.

Correction for Varying Water Contribution. In our subtraction method, it is necessary to use a double-correction approach involving water and the culture medium. The validity of this approach was demonstrated in a study in which spectra were measured from bacteria grown on various culture media that differed only in water content. Figure 5A shows bacteria spectra after vector correction for the medium alone. Although the spectra look remarkably similar, difference spectra revealed residual variation in the spectra especially in the region around 1650 cm⁻¹. This broad band is characteristic of the Raman spectrum of water. When the extra correction step for the water signal was performed, the spectra were obtained, although the intensity of one peak at 780 cm⁻¹ was higher in the medium richer in nutrients (i.e., lower water content) (Figure 5B). This peak could be attributed to uracil (RNA) (Table 1), but one would expect to observe other RNA peaks in the difference spectrum as well. Further research should give additional information and is currently being performed.

Limitations in the comparison of organisms cultured on different media therefore do not originate from the correction method presented here but rather from intrinsic physiological cell differences. Minimization of these influences can be achieved by thorough standardization in the preparation of culture media and in the culture conditions. Commercially available, ready-to-use media have the advantage that major manufacturers use standard protocols and quality controls in their medium preparations. To what extent minor fluctuations from batch to batch can influence the identification process is not yet clear and needs further investigation. The presence of heterogeneity within microcolonies suggests that the Raman signal from several locations in the colony needs to be averaged in order to optimize reproducibility.

Identification of Bacteria. Now that complicating factors such as the variability in the water and medium signal contributions are eliminated, the Raman spectra can be treated by multivariate analysis for (nonsubjective) classification of bacterial strains. Representative spectra acquired from five bacterial strains are shown in Figure 6A. Closer inspection of the spectra reveals that there are, indeed, spectral differences characteristic of the various strains. For example, the Staphylococcus strains are characterized by markedly high intensities at 780 cm⁻¹. To the naked eye, the two S. aureus stains resemble one another very closely. The spectra of E. coli have a characteristically intense band at around
1004 cm\(^{-1}\) arising from the ring breathing vibration found in the amino acid phenylalanine. The biological relevance of the increased intensity of this peak in *E. coli* Raman spectra is presently unknown. Raman spectra arising from *S. epidermidis* and *E. faecium* have slightly poorer signal-to-noise levels when compared to those of the other strains. This is because the 6-h microcolonies of these strains are very thin. The colony thickness was approximately 1–3 μm for *S. epidermidis* and *E. faecium*, as opposed to between 6 and 8 μm for the strains of *S. aureus* and *E. coli*.

These differences lie in the intrinsic biological growth difference of the strains and are reflected in the resulting Raman spectra.

Hierarchical cluster analysis of the microcolony spectra of the 5 strains over 3 days resulted in the dendrogram shown in Figure 6B. We observed that there are two major clustering branches, consisting of the *Staphylococcus* strains in one group and *E. faecium* and *E. coli* in the other. Within this latter group, a clear division is observed between the *E. faecium* and *E. coli* strains. In the *Staphylococcus* branch, subclusters are formed of *S.
epidermidis and S. aureus. Therefore, it is possible to distinguish these bacterial genera and the Staphylococcus species on the basis of their Raman spectra.

Within the clusters containing spectra of one strain, spectra measured on the same day tended to cluster together, indicating the presence of some day-to-day variation. This variation, however, did not interfere with identification down to the species level of the Staphylococcus strains, suggesting that spectral differences between the various species are greater than any subtle day-to-day variation in the spectra of the strain.

When data from all 3 days were analyzed together, the two S. aureus strains could not be clearly separated with the unsupervised classification approach. PCA followed by the supervised linear discriminant analysis on only the S. aureus spectra resulted in a 100% correct classification of the training set and 83% of the test set. When hierarchical cluster analysis was performed on the PCA scores of each day separately, 100% separation of the two S. aureus strains was obtained. This observation suggests that when data from the 3 days were combined, the complete discrimination of the two S. aureus strains was hindered by day-to-day variations in the spectra. The effect of small deviations in wavenumber calibration of the spectra as well as fluctuations in the biochemical composition of the cells are currently being investigated as possible sources of this day-to-day variation. The issue of repeatable wavenumber calibration of multichannel Raman instruments and instrument-to-instrument calibration transfer is an area of active research. A thorough analysis of the problems involved as well as potential solutions was recently published by Mann and Vickers.11

Other studies using FT-IR spectroscopy combined with multivariate analysis have also reported the successful separation of different bacterial species, such as Eubacterium spp.,24 Lactobacillus spp.,33 Streptococcus spp.,33 and Enterococcus spp.41 Hence, although the various Raman spectra of microorganisms look similar upon first inspection, as with FT-IR spectra, there is a high information content in these spectra which can be used in multivariate analysis for the discrimination of microorganisms.

Overall, then, the classification results presented indicate that it is possible to identify bacteria from their Raman spectra acquired directly from microcolonies growing on solid culture medium.

CONCLUSION

We have presented a novel approach of the use of Raman microspectroscopy for the identification and characterization of clinically relevant microorganisms. An approach for measuring Raman spectra of 6-h microcolonies directly on the solid culture medium could, in principle, enable identification results to be obtained within the same day of receipt of patient material. In contrast, conventional microbiological approaches require significantly more time to arrive at the same result. With respect to other spectroscopic methods mentioned in the introduction, Raman spectroscopy also has the advantage of minimizing culturing time, sample handling, and the use of chemicals and disposables. Using online data processing routines, the Raman spectra can be analyzed and clustered rapidly. However, key features to the successful application of this technique are the careful correction for variable parameters, such as water signal contributions of the culture medium and the amount of culture medium signal collected with the Raman signal of bacteria. Averaging of signal over several positions in the microcolony is fundamental as well, to compensate for colony heterogeneity. The methods must be rigorously standardized, such that reproducible spectra of good signal-to-noise levels are obtained. Furthermore, the use of a standard growth medium with carefully controlled composition is essential for accurate classification of bacteria. The different signal contributions of different media can, in principle, be compensated for through the application of double-vector correction schemes. However, it is likely that differences in medium composition also have an influence on the overall biochemical composition of the bacteria. This is, therefore, a point of concern for all spectroscopic methods.

To our knowledge, this is the first report on spectroscopic measurements of bacteria directly on solid culture media. Good signal-to-noise levels could be obtained after an incubation period of 6 h, followed by 2.5 min of Raman signal collection. From the limited data set presented here, it appears that bacteria may be distinguished at the genus and species level on the basis of unsupervised analysis of their Raman spectra. Supervised methods, such as wavenumber region selection methods14 and minimization of day-to-day variation in instrumental and/or biological parameters, should enable discrimination down to the strain level. Therefore, we conclude that, when thoroughly standardized and optimized, confocal Raman microspectroscopy has potential as a powerful new tool in diagnostic microbiology.

ACKNOWLEDGMENT

We thank Nicole van den Braak from the Department of Medical Microbiology and Infectious Diseases (University Hospital Rotterdam) for assistance with obtaining the strains used for the studies. We also gratefully acknowledge financial support from the European Union Biomed II program, Project No. BMH4-97-2054.

Received for review September 1, 1999. Accepted October 20, 1999.

AC991011H
Identification of Dyes on Ancient Chinese Paper Samples Using the Subtracted Shifted Raman Spectroscopy Method

S. E. J. Bell,* E. S. O. Bourguignon, A. C. Dennis, J. A. Fields, J. J. McGarvey, and K. R. Seddon

School of Chemistry, The Queen’s University of Belfast, Belfast BT9 5AG, Northern Ireland

The Stein Collection in the British Library contains the Diamond Sutra, the world’s oldest, dated, printed document. The paper of the Diamond Sutra and other documents from the Stein collection is believed to be dyed yellow by a natural extract, called huangbo, from the bark of Phellodendron amurense, which contains three major yellow chromophores: berberine, palmatine, and jatrorrhizine. Conservation of these documents requires definite information on the chemical composition of the dyestuff but no suitable, completely noninvasive analytical method is known. Here we report resonance Raman studies of a series of pure dyestuff, of plant materials and extracts, and of dyed ancient and modern paper samples. Resonance Raman spectroscopy is used to enhance the spectra of the dyestuff over the background from the paper matrix in which they are held. The samples all give resonance Raman spectra which are dominated by intense fluorescence, but by using SSRS (subtracted shifted Raman spectroscopy) we have obtained reliable spectra of the pure dyestuff, native bark from Phellodendron amurense, modern paper dyed with huangbo extracted from this bark, and ancient paper samples. For both ancient paper samples whose pigment bands were detected, the relative intensities of the bands due to berberine and palmatine suggest that the ancient paper is richer in berberine than its modern counterpart. This is the first nondestructive in situ method for detection of these pigments in manuscripts, and as such has considerable potential benefit for the treatment of irreplaceable documents that are believed to be dyed with huangbo but on which conservation work cannot proceed without definite identification of the chemical compounds that they contain.

The Stein Collection in the British Library contains the Diamond Sutra, the world’s oldest, dated, printed document, which is in urgent need of conservation due to unsuitable previous treatments and repairs.1 This document, dated 868 A.D., is one of 14,000 manuscripts recovered by Sir Marc Aurel Stein from a cave library in the town of Dunhuang in Northwest China. The paper of the Diamond Sutra and other documents from the Stein collection is believed to be dyed yellow by a natural extract called huangbo from the bark of Phellodendron amurense, which contains three major yellow chromophores: berberine, palmatine, and jatrorrhizine.2 The conservation of the Diamond Sutra has focused on the problem of analyzing for the presence of these dyestuffs in the paper, since they must not be removed or destroyed during any conservation procedure.3 Analysis for the components of the dye in fragments of paper from the Stein Collection has already been carried out on paper samples both by liquid secondary ion mass spectrometry4 and by HPLC.5 However, these analyses require removal of a small sample (1 × 3 mm) from the document: a completely noninvasive analytical method is still needed.

Here we report resonance Raman studies of a series of pure dyestuffs, of plant materials and extracts, and of dyed paper samples (both fragments from the same source as the Diamond Sutra and modern analogues). The advantages of Raman spectrometry as a nondestructive in situ technique for analysis of irreplaceable documents and artifacts have now been well-established.6–12 In cases where the main interest is in a region of the sample


10.1021/ac9908812 CCC: $19.00 © 2000 American Chemical Society Published on Web 12/01/1999
containing predominantly a single species, then conventional, spontaneous Raman experiments will yield the required vibrational spectra directly. In situations where the analyte is colored, it is possible to enhance the signal due to the chromophore selectively, by choice of an appropriate excitation wavelength, so that the chromophore’s signal dominates the scattering from the sample (resonance Raman scattering). Both these approaches have been applied to a wide range of objects, particularly those containing inorganic pigments. However, one of the factors which limits the range of samples that can be probed by Raman techniques is the occurrence of strong luminescence, which can, in many cases, dominate the signals recorded from the sample. This luminescence can arise either from the analyte of interest (if it absorbs at the laser wavelength used for excitation, as in resonance Raman spectroscopy) or from adventitious impurities within unpurified samples.

The problem of sample luminescence has been recognized for decades, and many different strategies have been developed to circumvent the problem. It may be possible either to quench the luminescence (most commonly by using surface-enhanced Raman techniques)25 or to shift the excitation wavelength to one where the Raman signal lies in a wavelength range different from that of the luminescence or where the excitation source does not generate the luminescence.24,26 In the case of the ancient papers which are of interest here, it is necessary to use resonance Raman methods to enhance the spectra of the dyes over the signals from the paper matrices in which they are held. Unfortunately, the dyed paper samples all give resonance Raman spectra which are dominated by intense fluorescence and it is not possible to change the excitation wavelength to one where fluorescence is less pronounced, since this moves the excitation wavelength off-resonance. The alternative approach of using a surface-enhanced method is precluded because the enhancement requires intimate, and irreversible, contact between the paper (or the dyes they contain) and the enhancing metal surface.

Given that the Raman spectra of these samples must be obtained under conditions where the fluorescence is considerably stronger than the Raman signal, it is necessary to accumulate very high total detected photon levels using a multichannel (in this case CCD) detector to reduce the photon shot noise associated with the fluorescence background. Unfortunately, even after extensive signal accumulation apparently random noise is still detected at levels comparable to that of the Raman signals. This fixed pattern response, which is caused by random variations in sensitivity between different detector elements, is sufficiently large that it can obscure the Raman signal, and it must therefore be removed from the data.

We recently reported a method for analysis of highly luminescent samples which is eminently suitable for situations such as the one described above.26 This SERS (subtracted Raman spectroscopy) technique involves taking two or more Raman spectra of the same sample at slightly shifted spectrometer grating positions and then subtracting these spectra. The shift is chosen to be sufficiently small that the background fluorescence remains approximately constant while the Raman bands follow the shifted spectrometer grating positions. Subtraction of the two spectra gives a derivative-like spectrum from which the background has been almost eliminated. This method minimizes the apparently random noise on the spectra caused by random variations in sensitivity between different detector elements and gives signals in which the noise level is determined by the photon shot noise on the background. This photon noise can be reduced to acceptable levels by increasing accumulation times. Finally, curve-fitting the difference data gives peak parameters which can be used to reconstruct a conventional (undifferenced) representation of the spectrum. We have previously shown that, provided the degree of uncertainty in the data is correctly characterized, it is completely valid to come to conclusions about the spectra of the sample on the basis of the reconstructed data. The method is similar to that developed by Mathies et al. for removal of fluorescence from spectra of photosynthetic pigments, which uses pairs of spectra recorded with slightly shifted excitation wavelengths, but it has the advantage that it does not require a tunable laser excitation source.27-39 This means that even though a tunable, CW, moderate-power UV laser is not available, these experiments can be carried out with a simple, widely available Ar+ laser source.

EXPERIMENTAL SECTION

Raman spectra were recorded using 363.8 nm excitation (100 mW) from a Spectra-Physics 2020 Ar+ laser with a 180° backscattering geometry and a Kaiser holographic notch filter. Scattered photons were collected, dispersed by a Jobin-Yvon HR640 single-stage spectrophotograph, and detected with a Princeton Instruments (PI LN/1152UV) CCD detector. The spectrometer was calibrated using the standard Raman band positions of solid naphthalene.

For the difference spectroscopy, the spectrometer grating was manually moved from its initial calibrated position to the required shifted value by monitoring the position of a strong line from a medium-pressure Pen-Ray Hg lamp in real time. The normal acquisition protocol was to record a spectrum at the initial position, shift the spectrometer by 0.5 cm⁻¹ (21 cm⁻¹) and record a second spectrum, and then move to the third position (approximately 3 cm⁻¹, 43 cm⁻¹, from the original position) for the final data acquisition of the cycle. To minimize the effect of changes in

---

(25) Derbyshire, A.; Withall, R. J. Raman Spectrosc. 1999, 30, 185–188.
(36) Bell, S. E. J.; Bourguignon, R. S. O.; Dennis, A. Anal. Chem. 1998, 70, 1723–1729.
background luminescence level, this three-step acquisition cycle was normally repeated several times to average out, as far as possible, gradual changes in excitation laser power. Typical accumulation times were 1−2 h.

The average irradiance of the laser was reduced to a minimum by using an elliptical lens to fine-focus the beam onto the sample and by rotating the sample to decrease the time each point of the sample was exposed to the laser. The powders were placed in a grooved metal dish while the bark was fixed to a similar dish using thin metallic bands. The paper samples were placed between two flexible plastic sheets which were joined on two sides. An aperture was cut into the top plastic sheet to enable the paper sample to be exposed directly to the laser beam. In each case, the samples were placed horizontally on the spindle of a small electric motor and held in place magnetically. The horizontal configuration of the sample holder allowed experiments to be carried out with no glass or quartz window between the laser beam and the sample. With larger, intact documents which cannot be rotated, the laser beam can be directed onto the sample by a rotating mirror fixed at 45° to the surface. Tilting this mirror a few degrees off its rotation axis causes the beam to traverse a nearly circular elliptical path on the paper, which has an effect similar to that of rotating the sample itself.

Raman data were transferred to GRAMS 386 software for processing. The data were fitted to double Lorentzian functions of the type

\[
I = \frac{Ho^2}{\sigma^2 + (\nu - \nu_0)^2} - \frac{F \cdot Ho^2}{\sigma^2 + (\nu - \nu_e - \delta)^2}
\]

where \(I\) is the signal intensity at \(\nu - \nu_e\), \(\delta\) is the shift between the subtracted spectra, and \(H\) and \(\sigma\) are the band height and width, respectively. \(F\) is a scaling factor which allows optimization of the subtraction factor between the shifted spectra, to give the lowest residual luminescence background. Subsequent reconstruction of conventional representations of the fitted data was carried out as described previously.

RESULTS

Both berberine and palmatine show significant luminescence backgrounds when their resonance Raman spectra are recorded using 363.8 nm excitation; the problem of luminescence backgrounds is even more pronounced when attempts are made to run spectra of paper or bark samples. The latter, more challenging, spectra were all found to be broadly similar to each other so that, for brevity, a full set of representative spectra are shown for only one of the samples, bark from Phellodendron amurense. Figure 1 shows resonance Raman spectra of a bark sample, which were recorded at three slightly shifted spectrometer positions. It is clear that the spectra are dominated by luminescence from the sample, but small features are just visible at ca. 1640 and 1400 cm\(^{-1}\) when the spectra are expanded in this spectral region (see Figure 1 (insert)). Spectra a and b of Figure 2 show the results of subtracting the spectrum taken at the first spectrometer grating position from those taken at shifts of \(\delta\) and \(2\delta\), respectively. In the subtracted

![Figure 1. Resonance Raman spectra of a sample of bark from Phellodendron amurense, recorded at three slightly shifted spectrometer positions. Small features are just visible at ca. 1640 and 1400 cm\(^{-1}\) when the spectra are expanded in this spectral region (see insert).](image)

![spectra the signal-to-noise ratio is shot noise limited; i.e., the noise level is \(2^{1/2}\) (number of detected photons)^{1/2}. The factor of \(2^{1/2}\) arises because each of the spectra used in the subtraction has a shot noise of (number of detected photons)^{1/2}. In the examples shown here, this shot noise level is approximately 10 times less than the apparent noise on the original data which arises from the fixed pattern response. Since no trace of the fixed pattern can be discerned in the subtracted spectra we can give a lower limit of >90% of the fixed pattern being removed in the subtraction, the shot noise on the data prevents a more accurate estimate of the degree of fixed pattern removal.

The results of curve-fitting each of the subtracted spectra independently and then reconstructing conventional spectra using the parameters derived from the curve-fitting are shown in Figure 2c,d. The similarity between these independently reconstructed spectra gives an indication of the degree of confidence which can be placed in the data. The uncertainty is surprisingly low, considering the intense luminescence background on the spectra from which these data were derived. All the strong bands are present at the same positions and with similar relative intensities in both spectra; the level of uncertainty lies around the presence or absence of much smaller bands, such as those found at 1428 and 1423 cm\(^{-1}\). The latter bands occur in only one of the two reconstructed spectra, so it is not possible to determine with a reasonable degree of certainty whether they are actually present or are artifacts of the fitting procedure (caused by fitting random noise, for example). The final step in the reconstruction process is to derive the best guess spectrum, which is the one which fits both sets of subtracted data satisfactorily.

Figure 3 shows the best guess spectrum of the bark sample along with those of berberine, palmatine, and modern paper dyed with huangbo, all of which were obtained by exactly the same method. Similarly, Figure 4 shows reconstructed spectra of three fragments of ancient paper which were taken from the same source as the Diamond Sutra and are contemporary with it. The background luminescence level and degree of uncertainty in the fitted data for these samples are similar to those of the bark sample.

DISCUSSION

At first sight, the very high background luminescence on the resonance Raman spectra of all the samples studied appears to

\[\text{(40) GRAMS 386 is available from Galactic Industries Corp., 255 Main St., Salem, NH 03079.}\]
Figure 2. Subtracted shifted resonance Raman spectra obtained by subtracting spectra taken at the first spectrometer grating position from those taken at shifts of (a) δ and (b) 2δ. The result of curve-fitting the subtracted spectrum at a shift of δ (a) with double Lorentzians and reconstructing conventional spectra using the parameters derived from the curve-fitting is shown as (c). (d) Is the reconstructed Raman spectrum obtained from the 2δ shift spectrum (b).

Figure 3. Final reconstructed resonance Raman spectra of (a) berberine and (b) palmatine. (c) Is the scaled sum of the spectra of berberine and palmatine for comparison with (d), the final reconstructed spectrum of Phellodendron amurense bark. The spectrum of modern paper dyed with Huangbo pigment extracted from the bark of Phellodendron amurense is shown as (e).

Figure 4. Reconstructed resonance Raman spectra of three samples of ancient Chinese paper from Dunhuang. Catalog numbers of the samples are (a) UN006, (b) DU017, and (c) TM009. The background luminescence level and degree of uncertainty in the fitted data for these samples are similar to those of the bark samples (see Figures 2 and 3(c)).

mask the vibrational data almost completely. Removing the background signal by fitting a smooth polynomial function will not improve the data significantly since the spectra are dominated by the apparently random noise which arises from irregularity in the detector response (see insert to Figure 1). Since the irregularity is fixed, the most obvious method to remove it would be to use a broad-band continuum source to "flat-field" the detector response. Unfortunately, this type of flat-fielding, although attractive in principle, is extraordinarily difficult to carry out in practice. There are several reasons for this, but the most obvious is that it is very difficult to match the vertical distribution of light falling on the detector in the Raman experiments with that obtained using a continuum source. Since the response varies between detector elements along both the horizontal (wavelength) and vertical axes, mismatch in the vertical distribution results in mismatch of the flat-fielding correction. In addition, more subtle effects, such as changes in the flat-field response with changes in the effective aperture seen by each detector element, are also large enough to cause poor flat-fielding correction with continuum sources. (A reviewer has pointed out that these problems may be minimized if fiber optic coupled collection optics are used.)

However, simple subtraction of the spectra taken at slightly different spectrometer positions removes the fixed pattern response and gives spectra in which the Raman signals are immediately apparent, albeit in an unfamiliar form, with positive peaks matched by their negative images. Moreover, comparison of spectra reconstructed from two different subtracted spectra (Figure 2) shows that, even for the spectra with high background luminescence levels, the degree of uncertainty in the results is acceptably small so that the best fit data can be compared with some confidence.

The resonance Raman spectra of berberine and palmatine, shown in Figure 3, are similar, which reflects their similar chemical structures, but there are sufficient differences between them to allow one to be distinguished easily from the other. The most striking differences are the appearance of a strong band at
1400 cm$^{-1}$ in the spectrum of berberine, which is not present in that of palmatine, and the presence of a distinct peak at 1607 cm$^{-1}$ in the spectrum of palmatine. No attempt has been made to assign these bands, since the primary objective of this work is to determine whether the pigments are present in paper samples and this can be achieved without a detailed vibrational analysis.

The bark of *Peltodenon amurense* is bright yellow and is known to contain both berberine and palmatine, along with much lower quantities of a related pigment (gastrorhazine). The spectrum of the bark (Figure 3d) is dominated by the strong 1640 cm$^{-1}$ band which is common to both major pigments present and has a distinct shoulder at 1607 cm$^{-1}$, a band associated with palmatine. However, the spectrum also contains a band at 1400 cm$^{-1}$, which is associated with berberine, so that the spectrum contains signals from both berberine and palmatine. Some of the bands appear in similar positions in both spectra (for example, those at 1640 and 1520 cm$^{-1}$) and thus reinforce each other when the spectrum of a mixture is recorded. Conversely, those bands which appear in the spectrum of only one of the compounds will appear to lose intensity in the spectra of mixed samples. It is possible to model the expected spectrum of a berberine/palmatine mixture by making a scaled sum of the spectra of the individual components; Figure 3c shows the result of such an addition. This model spectrum is very similar to that of the bark sample, having both a shoulder at 1607 cm$^{-1}$ and a peak at 1400 cm$^{-1}$, demonstrating that both species are present in the native bark. It is possible to estimate (±20%) the relative proportions of the two pigments in the spectra of mixtures from the scaling factor used in making the scaled sum. However, for the purposes of conservation, the relative amount of each pigment present in a sample is not important because both pigments have very similar chemical properties and any treatment chosen would have to be compatible with both.

The spectrum of modern paper dyed with huangbo pigment extracted from the bark of *Peltodenon amurense* (Figure 3e) is strikingly similar to that of the bark itself, which is hardly surprising since the extraction process is known to produce a liquor rich in berberine and palmatine. However, there appears to be a loss of resolution in the bands adjacent to the strongest band at 1640 cm$^{-1}$; a shoulder appears at 1607 cm$^{-1}$, but the distinct palmatine peak has disappeared. Since this spectrum was run under conditions identical to those for the others, the loss of resolution is not due to the instrumentation but must be due to additional peaks in the same spectral region overlapping with those of the expected signal. In fact, these additional bands were observed much more clearly when the spectra of the ancient paper samples (Figure 4) were recorded. The spectra of two of these samples, Figure 4a,b (the samples are identified here simply by their catalog numbers), are broadly similar to those discussed above, but that of one of the ancient samples (Figure 4c), which had previously been classified as undyed by visual inspection at the British Museum, has a dominant broad feature centered at 1600 cm$^{-1}$ and no distinct berberine or palmatine bands. The origin of this broad feature is unknown: it does not correspond to the known spectrum of modern cellulose-based papers but is too broad to contain other indications of its source. It might arise from other components of plant-derived paper treatments. For example, it is known that waxes and polysaccharides are extracted along with berberine and palmatine from *Peltodenon amurense* in the preparation of huangbo.

In the spectrum of the ancient paper sample TM006, there are no strong distinct bands, apart from the broad feature, but in the spectra of the other ancient paper samples studied, there is clear evidence for the presence of berberine and related dyes. The spectrum of UN006, shown in Figure 4a, is very similar to that of berberine; it shows only a slight trace of broadening adjacent to the strong feature at 1640 cm$^{-1}$, which could be due to the presence of palmatine (characteristic band at 1607 cm$^{-1}$) and/or the unidentified broad feature which is centered at 1600 cm$^{-1}$. The ancient paper DU017 similarly contains several peaks characteristic of berberine and a shoulder in the position characteristic of an additional palmatine component along with a stronger underlying broad feature. The intensity of the broad feature is larger than that found for paper sample UN006 but is similar to that found for modern paper dyed with huangbo.

In the spectra of both ancient paper samples in which pigment bands were detected, the relative intensities of the bands at 1400 cm$^{-1}$ (berberine) and 1370 cm$^{-1}$ (palmatine) suggest that the ancient papers are richer in berberine than their modern counterpart (> 4:1 berberine:palmatine in the ancient papers vs 6:1 in the modern paper sample). This could be due to several factors: The exact source of the dye and method of dye preparation used for these particular ancient samples are not known (although a description of an ancient method for dye extraction from plant sources has been found; see ref 2) so that the original composition of the dye may have been different from that used for the modern sample. In addition, chemical changes in the dye may have occurred in the centuries since the paper was originally treated. However, irrespective of the relative proportions of the two dyes, it is clear that the yellow tint of the ancient paper samples arises from them and not from any one of the numerous other classes of dye that could give rise to the yellow coloration, such as carotenoids or inorganic pigments, whose Raman spectra would be completely dissimilar from those found here.

**CONCLUSION**

The SERS method can be used to detect the presence of berberine and related dyes in both modern and ancient papers. This is a particularly challenging analytical problem, given the low levels of compound which must be detected and the large luminescence background which the spectra contain. The spectra of the ancient papers show that the relative concentrations of berberine and palmatine are different from those of modern paper dyed with huangbo and the native bark from which the modern huangbo was prepared. This is the first nondestructive in situ method for detection of these pigments in manuscripts and as such has considerable potential benefit for the treatment of irreplaceable documents which are believed to be dyed with huangbo but documents on which conservation work cannot
proceed without definite identification of the chemical compounds that they contain.

ACKNOWLEDGMENT

We are indebted to Nadezhda Brovenko and Yuri Petrosyan of the Oriental Institute of St. Petersburg for providing samples from their collection. E.S.O.B. acknowledges the award of an Erasmus studentship.

Received for review August 24, 1999. Accepted October 12, 1999.

AC9809812