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 dyes but no suitable, completely noninvasive analytical method is known. Here we report resonance Raman studies of a series of pure dyes, of plant materials and extracts, and of dyed ancient and modern paper samples. Resonance Raman spectroscopy is used to enhance the spectra of the dyes over the signals from the paper matrixes 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 dyes, native bark from the 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 documents 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 an urgent need of conservation due to unsuitable previous treatments and repairs. This document, dated of May 11, 868 A.D., is one of 14,000 manuscripts recovered by Sir Marc Auel 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. The conservation of the Diamond Sutra has focused on the problem of analyzing for the presence of these dyes in the paper, since they must not be removed or destroyed during any conservation procedure. 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 spectrometry and by HPLC. 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 dyes, 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 spectroscopy as a nondestructive in situ technique for analysis of irreplaceable documents and artifacts have now been well-established.1–32 In cases where the main interest is in a region of the sample

(20) Best, S. P.; Clark, R. J. H.; Withnall, R. Endavour 1992, 16, 66–73.
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) 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. 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 matrix 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. This SSRS (subtracted shifted 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 convolution (un differenced) 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. 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 spectrograph, 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 δ cm⁻¹ (21 cm⁻¹) and record a second spectrum, and then move to the third position (approximately 2δ cm⁻¹, 43 cm⁻¹, from the original position) for the final data acquisition of the cycle. To minimize the effect of changes in

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 a cylindrical lens to line-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 disk while the bark was fixed to a similar disk 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 to the sample holder allowing 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 to the sample holder allowing experiments to be carried out with no glass or quartz window between the laser beam and the sample.

Raman data were transferred to GRAMS 386 software for processing.40 The data were fitted to double Lorentzian functions of the type

$$I = \frac{H\sigma^2}{\sigma^2 + (v - v_c)^2} - F \frac{H\sigma^2}{\sigma^2 + (v - v_c - \delta)^2}$$

where $I$ is the signal intensity at $v - v_c$, $\delta$ 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.36

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 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 $\approx 90\%$ 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

![Figure 1. Resonance Raman spectra of a sample of bark from the 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-url)
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.39 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
We have found that berberine is changed to an as yet unidentified red and this can be achieved without a detailed vibrational analysis. To determine whether the pigments are present in paper samples and this can be achieved without a detailed vibrational analysis. The bark of Phellodendron amurense is bright yellow and is known to contain both berberine and palmatine, along with much lower quantities of a related pigment (jatrorrhizine). 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 Phellodendron 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\(^{-1}\), 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.

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 6:4 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 SSRS 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

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.

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.

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 6:4 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 SSRS 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.

AC9909812