Phys. Med. Biol. **45** (2000) R1–R59. Printed in the UK PII: S0031-9155(00)91279-X

TOPICAL REVIEW

Prospects for *in vivo* Raman spectroscopy

E B Hanlon†, R Manoharan†‡, T-W Koo†, K E Shafer†, J T Motz†, M Fitzmaurice§, J R Kramer‖, I Itzkan†, R R Dasari† and M S Feld†

† Laser Biomedical Research Center, George R Harrison Spectroscopy Laboratory, Massachusetts Institute of Technology, 77 Massachusetts Avenue 6-014, Cambridge, MA 02139, USA
‡ Rosemount Analytical, Inc., 1201 North Main Street, PO Box 901, Orrville, OH 44667, USA
§ University Hospitals of Cleveland and Case Western Reserve University, Department of Pathology, 2085 Adelbert Road, Cleveland, OH 44106, USA
‖ The Cleveland Clinic Foundation, Department of Cardiology, 9500 Euclid Avenue F-25, Cleveland, OH 44195, USA

E-mail: ebhanlon@mit.edu, ramasamy.manoharan@frco.com, tkoo@mit.edu, keshafer@mit.edu, jtmotz@mit.edu, mxf39@po.cwru.edu, kramerj@cesmtp.ccf.org, itzkan@mit.edu, rrdasari@mit.edu and msfeld@mit.edu

Received 10 May 1999

Abstract. Raman spectroscopy is a potentially important clinical tool for real-time diagnosis of disease and *in situ* evaluation of living tissue. The purpose of this article is to review the biological and physical basis of Raman spectroscopy of tissue, to assess the current status of the field and to explore future directions. The principles of Raman spectroscopy and the molecular level information it provides are explained. An overview of the evolution of Raman spectroscopic techniques in biology and medicine, from early investigations using visible laser excitation to present-day technology based on near-infrared laser excitation and charge-coupled device array detection, is presented. State-of-the-art Raman spectrometer systems for research laboratory and clinical settings are described. Modern methods of multivariate spectral analysis for extracting diagnostic, chemical and morphological information are reviewed. Several in-depth applications are presented to illustrate the methods of collecting, processing and analysing data, as well as the range of medical applications under study. Finally, the issues to be addressed in implementing Raman spectroscopy in various clinical applications, as well as some long-term directions for future study, are discussed.

(Some figures in this article appear in colour in the electronic version; see www.iop.org)

Contents

1. Introduction

The physician's ability to diagnose disease is enhanced by the timely availability of objective, quantitative diagnostic information. Successive advances in biomedical technology have been driven by the need to provide such information.

Novel biomedical applications of optical spectroscopy, such as fluorescence, reflectance and Raman scattering, can provide information about the composition of tissue at the molecular level. Of these techniques, Raman spectroscopy can provide the most detailed information about the chemical composition of the tissue under study. Since the progression of disease is accompanied by chemical change, Raman spectroscopy can provide the physician with valuable information for diagnosing disease. And since light can be delivered and collected rapidly via optical fibres, which can be incorporated into catheters, endoscopes, cannulas and needles, as necessary, Raman spectroscopy can be performed *in vivo* in real time.

Diagnostic applications of Raman spectroscopy currently under investigation are wide-ranging. For example, Raman spectroscopy may be used to monitor blood analytes non-invasively. It may be used to perform minimally invasive, real-time, tissue diagnosis *in vivo*, in cases where biopsy cannot be performed readily, such as coronary artery disease and Alzheimer's disease, or where a high incidence of false positive screening tests leads to unnecessary biopsy procedures, as in breast cancer. The purpose of this article is to discuss the potential of Raman spectroscopy to provide real-time, objective, quantitative diagnostic information *in vivo*.

Chandrasekhara Venkata Raman (figure 1), discoverer of the Raman effect, was born on 7 November 1888, in Tiruvanaikkaval, Madras, India, and died on 21 November 1970. Although he was a brilliant student, higher education was considered below his caste, and so in 1907 he went to Calcutta to join the Indian Financial Civil Service as Assistant Accountant General. A self-made scientist with intense drive, Raman worked early mornings and evenings in the underutilized facilities of the Indian Association for the Cultivation of Science, in Calcutta, to study problems in acoustics. Conducting research independently for nearly 10 years, he established his reputation and was appointed professor at the University of Calcutta in 1917. On his first voyage outside India, to Oxford in 1921, he performed some experiments and

Figure 1. Professor Sir C V Raman (by permission of Professor N Kumar, Raman Research Institute).

published a note in *Nature* entitled, 'The colour of the sea', in which he showed that the colour of the ocean is independent of sky reflection or absorption, and is instead due to scattering. So began his scientific research into the scattering of light, the work that led to his discovery of the Raman effect in 1928, for which he was awarded the Nobel prize in record time, just two years later. He was the first Asian, and the only Indian, to receive this award. Among scientists, Raman is best known for his strikingly ingenious yet simple experimental design, the lucidity of his ideas, and the depth of his observations.

The Raman effect is a fundamental process in which energy is exchanged between light and matter (see, for example, Diem 1993). When light impinges on a substance it can be scattered or absorbed. Most of the scattered light will have the same frequency as that of the incident light. However, a small fraction of the incident light can go into setting molecules in the material into vibration. The energy for this must come from the incident light. Since light energy is proportional to frequency, the frequency change of this scattered light must equal the vibrational frequency of the scattering molecules. This process of energy exchange between scattering molecules and incident light is known as the Raman effect.

From the energy level point of view, the process of Raman scattering can be viewed as the transition of a molecule from its ground state to an excited vibrational state, accompanied by the simultaneous absorption of an incident photon and emission of a Raman scattered photon. This is illustrated in figure 2.

The Raman scattered light can be collected by a spectrometer and displayed as a 'spectrum', in which its intensity is displayed as a function of its frequency change. Since each molecular species has its own unique set of molecular vibrations, the Raman spectrum of a particular species will consist of a series of peaks or 'bands', each shifted by one of the characteristic vibrational frequencies of that molecule.

As an example, the Raman spectrum of cholesterol, a representative biological molecule, is shown in figure 3. The spectrum is a plot of the scattered light intensity versus its change in frequency, relative to that of the incident light. Raman frequency shifts are conventionally measured in wavenumbers (cm^{-1}), a convenient unit for relating the change in vibrational

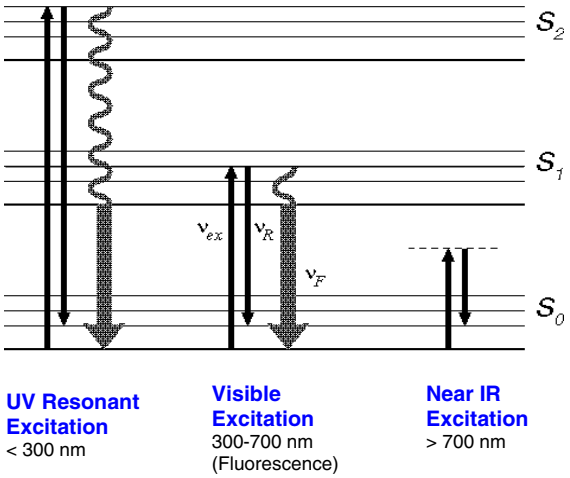


Figure 2. Raman effect: UV, visible and near-IR excitation strategies. S_0, ground electronic state; S_1 and S_2, excited electronic states. The horizontal lines indicate vibrational energy levels. The diagram shows how a molecule in the ground state, S_0, can make a transition from the lowest vibrational level to the first excited vibrational level by means of Raman scattering. Black up-arrows indicate the frequency of the laser excitation light; black down-arrows indicate frequency of the Raman scattered light. The difference in length between black up- and down-arrows indicates molecular vibration frequency. Grey arrows indicate frequency of the fluorescence light. Notice that Raman scattering at different excitation wavelengths, UV, visible and near-IR, produces the same change in vibrational energy; therefore the excitation wavelength can be chosen to avoid spectral interference by fluorescence. For visible excitation, the fluorescence light frequency and Raman scattered light frequency are similar. As discussed in section 2.1, this leads to intense fluorescence background in visible excitation Raman spectra. Near-IR light has too low a frequency to excite fluorescence, while for UV excitation, the fluorescence light frequency is much lower than the Raman scattered light frequency. Hence, fluorescence background in the Raman spectrum can be reduced by using UV or near-IR excitation.

energy of the scattering molecule to the change in frequency of the scattered light. One cm^{-1} equals 30 000 MHz; this is typically 10 000 times smaller than the frequency of the light itself. Each band of scattered light in the cholesterol Raman spectrum is characteristic of molecular vibrational motions, which, taken all together, are unique for cholesterol. For example, the characteristic Raman peak at 1440 cm^{-1} is due to the CH_2 and CH_3 deformation vibrations, and the peak at 1670 cm^{-1} is due to C=C stretching vibrations. If a sample of biological tissue contains cholesterol, these peaks will be present in its Raman spectrum. In other words, the molecular structure and composition of a material under study is encoded as a set of frequency shifts in the Raman scattered light. Thus, the Raman spectrum can provide a 'fingerprint' of a substance from which the molecular composition can be determined.

In Raman spectroscopy the incident light is often referred to as 'excitation' light, because it excites the molecules into vibrational motion. Ultraviolet, visible and infrared light can each be used for Raman excitation.

In general, Raman spectra of tissue are composed of relatively narrow bands, typically 10–20 cm^{-1} in width, which exhibit the presence of many biochemicals. The relative contributions of these biochemicals to the tissue Raman spectrum are proportional to their relative abundance in the tissue. This is the basis for the quantitative nature of the information Raman spectroscopy can provide for diagnosis. The quantitative nature of Raman spectra,

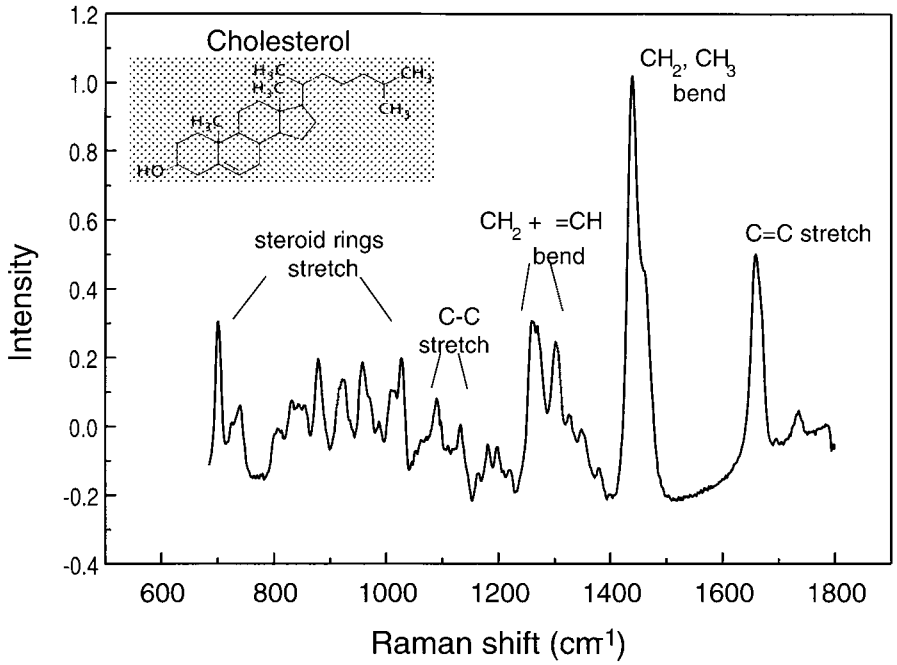


Figure 3. Near-IR Raman spectrum of cholesterol indicating typical vibrational bands. The background has been removed by subtracting a fourth-order polynomial fit to the raw spectrum, as described in the text.

combined with the ability to provide unique fingerprints of the biochemicals present in tissue, illustrate the potential of Raman spectroscopy to provide objective, quantitative diagnostic information for tissue analysis.

Although Raman spectroscopy had been widely used for chemical and molecular analysis for many years, its application to biomedical problems is relatively recent. Raman studies of biological tissue have been facilitated over the past 10 years by advancing technology, particularly in the areas of lasers and detectors. However, the real success of these studies at providing insight into the molecular basis of disease and showing the potential for medical diagnostic applications, has relied on advanced methods of analysing tissue Raman spectra. In order to exploit the full potential of Raman spectroscopy, methods of analysis that employ the full information content of the Raman spectrum, and not just the prominent peaks, must be used. The problem of extracting the full, useful, diagnostic information that tissue Raman spectra contain can be attacked at three related though distinct levels: statistical, chemical and morphological:

(a) Statistical analysis: Up to now, much of the work in this field has relied largely on empirical methods of analysis, such as correlating tissue type with prominent spectral bands. However, mathematical methods such as principal component analysis can be used to characterize the *full* range of spectral variations. The principal components can then be fitted to the observed Raman spectral lineshape. Correlating the fit coefficients with tissue type can be used to classify tissue by diagnostic category. In addition, features of the principal components can be used as a guide in identifying key biochemical or morphological constituents of the tissue or disease under study.

(b) Chemical analysis: Raman spectroscopy can provide information about the chemical composition of tissue. A Raman spectrum can be modelled as a superposition of the spectra of its chemical constituents, thereby providing quantitative chemical information. Chemical information is important, inasmuch as the onset of disease is accompanied by biochemical changes, and Raman spectroscopy can be an excellent method for detecting low level biochemical changes that may occur in either the cellular or extracellular compartments of tissue. At present, few techniques are capable of providing detailed *in vivo* biochemical analysis of tissues. Such techniques have an enormous potential, both in diagnosing disease and in understanding its origins and evolution.

(c) Morphological analysis: Quantitative methods to extract information about the microscopic constituents present in a tissue sample, i.e. the morphological structures, are also possible. Pathologists make a tissue diagnosis by assessing the presence, absence or relative abundance of such constituents, which include various types of cells, fibres, mineral deposits etc. Raman spectroscopy has the potential to provide such morphological information without tissue removal, in real time and in a quantitative and objective manner. This very interesting area is worthy of further exploration.

The goals of this review are to describe the various methodologies for tissue analysis and disease diagnosis using Raman spectroscopy. Statistical, chemical and morphological methods of analysis will be discussed in a variety of organs. The review is organized as follows: section 2 provides a general overview of the research in this field. Section 3 discusses instrumentation for Raman systems, and section 4 then discusses methods of modelling and analysis. Section 5 gives in-depth discussions of several applications, and is drawn from on-going research in our own laboratory. Applications to the diagnosis of atherosclerosis, breast cancer and Alzheimer's disease and measurement of blood analytes are presented. Conclusions and future prospects are discussed in section 6.

2. Application overview

Although the Raman effect was discovered over 70 years ago, Raman spectroscopy was not employed as an analytical tool until the invention of the laser. The initial Raman spectra of biological materials were measured with visible laser excitation, primarily using gas lasers. The first interpretable laser-excited Raman spectrum of a native protein was obtained from lysozyme by Lord and Yu (1970) at the MIT Spectroscopy Laboratory, using an He–Ne laser as the excitation source. As the techniques for biological and biomedical applications evolved, near-infrared (near-IR) laser excitation superceded excitation with visible light. This is because tissue background fluorescence is reduced when near-IR excitation light is used and because near-IR light can penetrate tissue deeply, of the order of millimetres, thus creating a 'diagnostic window' into tissue due to its minimal absorption. Initial near-IR Raman spectroscopic studies, in the early 1980s, almost exclusively employed Fourier transform (FT) spectrometers, which use interferometric methods to spectrally analyse the light (Diem 1993). FT Raman spectroscopy primarily uses 1064 nm light from an Nd:YAG laser as the excitation source, and a cooled InGaAs detector. FT spectroscopy has the so-called multiplex advantage, which allows all wavelengths to be collected simultaneously, and large light collection capability (the 'throughput advantage'), due to interferometric light collection. By working at 1064 nm in the near-IR, background fluorescence is almost completely suppressed. However, InGaAs detectors exhibit substantial noise, requiring long data collection times (typically 30 min) to acquire high-quality Raman spectra. The advent of low-noise charge coupled device (CCD) detector technology, highly efficient imaging spectrographs (which employ dispersion rather

than interference to spectrally analyse the Raman scattered light) and compact semiconductor laser excitation sources provided further impetus to applications of Raman spectroscopy for disease diagnosis.

Current instrumentation for Raman spectroscopy is discussed in section 3. The present section presents an overview of Raman spectroscopy for biomedical tissue analysis. The medical applications of Raman spectroscopy have been reviewed by several authors (Ozaki 1988, Nie *et al* 1990, Mahadevan-Jensen and Richards-Kortum 1996, Schrader *et al* 1997, Lawson *et al* 1997), and quantitative histochemical analysis for tissue diagnosis was reviewed by Manoharan *et al* (1996) and Richards-Kortum and Sevick-Muraca (1996).

2.1. Excitation wavelength strategies for biomedical Raman spectroscopy

When visible excitation is used, tissue exhibits strong, broadband fluorescence that can obscure the tissue Raman spectrum. Two strategies for reducing fluorescence interference are to use near-IR excitation or UV resonance excitation (figure 2).

Fluorescence decreases very rapidly at longer excitation wavelengths. This is because the lowest excited electronic states of molecules, the states from which fluorescence originates, have energies corresponding to visible wavelengths, and so they cannot be excited by lower-energy near-IR light. Therefore, most materials, including tissue, exhibit reduced fluorescence emission as the excitation wavelength increases into the near-IR region. Thus, fluorescence interference in tissue Raman spectra can be greatly reduced by using near-IR excitation. An alternate strategy is to use excitation wavelengths in the ultraviolet (UV). As discussed in section 2.5, background fluorescence is suppressed in tissue Raman spectra for excitation wavelengths below about 270 nm. An additional advantage of UV excitation light is that it can resonantly excite the electronic states of the molecule. In this case the Raman scattering for certain vibrations is enormously enhanced, providing larger signals. Disadvantages of UV resonance excitation are that the tissue penetration depth is of the order of microns, so this technique is only applicable for probing superficial abnormalities, and that UV light at these wavelengths can induce mutagenesis (Feld and Kramer 1991).

2.2. Visible excitation

Visible excitation Raman spectroscopy is particularly well suited for *in situ* cellular analysis because of the relatively weaker fluorescence of individual cells compared with intact tissue. Applications of Raman spectroscopy for cellular analysis were reviewed in 1986 (Butterfield 1986). Greve and co-workers used confocal Raman microscopy to examine DNA in chromosomes (Puppels *et al* 1990, 1991b, 1994) and pigments in granulocytes and lymphocytes, with very high spatial resolution (Puppels *et al* 1991a, b, 1993, Bakker Schut *et al* 1997). In confocal microscopy, the optics are arranged so that only a small, well-defined volume of tissue is sampled, thus giving rise to increased spatial resolution and rejection of background light (see, for example, Turrell and Corset 1996). Oxygen uptake in cells has been probed using spectral changes in haemoglobin modes (Hoey *et al* 1988). Employing argon ion laser-excited resonance Raman spectroscopy, they suggested that erythrocytes from patients with rheumatoid arthritis take up oxygen faster but in a less complete manner than those of normal patients. Spectral differences in normal and malignant hepatocytes have been reported by Hawi *et al* (1996). Larsson and Hellgren (1974) employed Raman scattering with 488 nm excitation light to study blood plasma samples from several patients, and observed differences in the carotenoid spectral regions between patients with cancer and normal patients.

Mineralized and collageneous tissues such as dentine (Rippon *et al* 1971), bovine Achilles tendon (Freshour and Koenig 1975), cornea (Goheen *et al* 1978) and ocular lens (Farrell and McCauley 1976, Yu *et al* 1985) were the first complex tissue matrices to be studied by visible-excitation Raman spectroscopy, because they exhibit relatively little fluorescence background. Raman measurements of intact lenses have been reported by many groups (Mizuno *et al* 1981, Yu *et al* 1982, 1987, Bot *et al* 1989). Protein aggregation and pigmentation that occur during lens aging and cataract formation were studied by Ozaki *et al* (1989). Confocal Raman microscopy was used to probe biochemical changes in the lens (Siebenga *et al* 1992, Smeets *et al* 1993) and hydration gradients across the cornea (Bauer *et al* 1998). Recently, confocal Raman microscopy has also been used for analysis of changes in cholesterol, phospholipid and protein in cataract lens, and the results compared with those of cytochemistry and transmission electron microscopy (Duindam *et al* 1995, 1998). Non-contact characterization of various layers of ocular tissues was studied by Jongsma *et al* (1997).

Gallstones and kidney stones are the other class of biological substances investigated extensively using visible excitation, because of the large Raman scattering cross section and relatively weak fluorescence from the inorganic substances in the stones (Ishida *et al* 1987). Tu and co-workers identified kidney stone types based on their chemical composition (Kodati *et al* 1990, 1991), and Ishida *et al* (1987) used Raman microspectroscopy to identify the molecular components in successive layers of gallstones.

Visible-excited Raman confocal microscopy for analysis of inclusions in pathological tissues was initially demonstrated by Abraham and Etz (1979). Since then, a number of studies have examined tissue from patients with mineral deposition diseases (Buitveld *et al* 1984, McGill *et al* 1991, Centeno *et al* 1994, Pestaner *et al* 1996) and prosthetic implants (Hahn *et al* 1997, Wolfarth, *et al* 1997, Luke *et al* 1997, Kalasinsky *et al* 1998), including chemical imaging of inclusions in pathological breast tissues (Schaeberle *et al* 1996).

The first Raman spectroscopy studies of cardiovascular tissue used visible argon ion laser excitation at 514.5 nm. The Raman spectrum was superimposed on a broad, featureless fluorescence background that could be subtracted from the sharp Raman features, and the Raman bands observed in calcified tissue and fatty plaques were assigned to hydroxyapatites and carotenoids respectively (Clarke *et al* 1987, 1988). Later, Klug *et al* (1992) confirmed this using red excitation light to minimize fluorescence.

In general, strong fluorescence background is a disadvantage for visible excitation Raman spectroscopy of biological tissue, even with 647.1 nm laser excitation (Yu *et al* 1979, Clarke *et al* 1987, 1988). Though 'pre-treatment' with continuous laser irradiation can reduce the fluorescence background (Clarke *et al* 1987, 1988), visible light excited Raman spectroscopy is limited in its usefulness for many applications of tissue diagnosis.

2.3. Near-IR excitation: FT Raman spectroscopy

To reduce intrinsic tissue fluorescence observed in visible excited Raman studies, the strategy of near-IR excitation, described above, has been adopted by several groups. In 1986, Hirschfeld and Chase (1986) showed that with 1064 nm Nd:YAG excitation, Raman spectra could be observed from materials in which Raman scattering was previously obscured by background fluorescence.

Initial studies of FT Raman spectroscopy focused on ocular lenses and calcified tissues. Yu and co-workers obtained fluorescence-free Raman spectra from pigmented animal and human lenses to study pathological processes such as senile cataract formation, lipid peroxidation and protein glycation (Nie *et al* 1990, Sebag *et al* 1994). The Raman spectral features of bone,

tooth enamel and dentine have been compared with those of their major components, such as hydroxy and carbonated apatite (Nie *et al* 1990, Rehman *et al* 1995).

Our laboratory reported the first interpretable FT Raman spectra of aortic tissues and later conducted an *in situ* quantitative biochemical analysis of normal and diseased tissues for diagnosing atherosclerosis (Rava *et al* 1991, Manoharan *et al* 1992, Baraga *et al* 1992a). Several other groups have also used near-IR FT Raman spectroscopy to collect spectra from arterial walls (Redd *et al* 1991, Liu *et al* 1992a).

The potential of Raman spectroscopy for probing chemical changes in cancerous tissues has also been investigated. Alfano and co-workers have conducted an exploratory survey of the near-IR Raman spectra of various lesions in gynaecological tissues (Liu *et al* 1992b). Raman spectra of normal and diseased breast tissues investigated with 1064 nm laser light exhibited differences in the amide modes (Alfano *et al* 1991). Spectral differences between normal human brain tissues and several kinds of tumours have also been reported (Mizuno *et al* 1992, 1994, Keller *et al* 1994). Mizuno *et al* (1994) suggested that relative changes in phospholipid to protein content and accumulation of mineral and carotenoid pigments in abnormal tissues are responsible for observed spectral differences. FT Raman spectra also have been measured and used to analyse stratum corneum and skin lesions both *in vitro* and *in vivo* (Barry *et al* 1992, Williams *et al* 1993, Fendel and Schrader 1998, Caspers *et al* 1998, Lawson *et al* 1998, Gniadecka *et al* 1998).

These studies make it clear that Raman spectroscopy can be employed as a powerful tool for probing biochemical changes in tissues *in situ*, and for providing important diagnostic information. However, the clinical utility of the FT Raman method is limited. Typical spectral data collection time is of the order of 30 min to 1 h, clearly impractical for *in vivo* clinical applications. Further, to obtain motion artefact-free spectra, clinical data collection times need to be less than 1 s. In addition, the FT Raman method requires average laser powers of 0.5 to 1 W, which is many orders of magnitude higher than the permissible light exposure limit for eye and skin (Sliney and Wolbarsht 1980). Finally, the throughput advantage of interferometer-based FT Raman spectroscopy is lost in instruments designed for clinical use, in which fibre-optic light delivery and collection is employed, due to the relatively small light collection capability of optical fibres.

2.4. Near-IR excitation: CCD based dispersive Raman spectroscopy

The technology for Raman tissue spectroscopy has now shifted to CCD-based dispersive spectrometers, with near-IR excitation. The main advantages of dispersive near-IR Raman spectroscopy are that compact solid state semiconductor lasers can be used for near-IR excitation, imaging spectrographs can be f-number matched with optical fibres for better throughput and cooled CCD detectors offer shot noise-limited detection. McCreery and co-workers suggested that dispersive near-IR Raman spectroscopy using a single-stage imaging spectrograph and CCD detector offers better sensitivity than FT Raman spectroscopy (Wang and McCreery 1989, Williamson *et al* 1989, Allred and McCreery 1990, McCreery 1995). Extensive studies conducted in our laboratory indicated that excitation wavelengths in the 800 nm range are optimal to minimize tissue fluorescence while best taking advantage of cooled CCD detectors (Baraga *et al* 1992b, Baraga 1992, Brennan 1995, Brennan *et al* 1997b). Though excitation wavelengths longer than 850 nm would further suppress fluorescence, they are not useful with silicon based CCD detectors because detector efficiency drops off sharply from 1000 to 1100 nm (corresponding to the Raman fingerprint region for 850 nm excitation). Though tissue background fluorescence still can be large, even with 850 nm excitation, the

statistical noise (shot noise) associated with this fluorescence is much smaller than the tissue Raman bands. Furthermore, since the fluorescence is broad and the Raman bands are sharp, the slowly varying fluorescence background can be removed by a variety of means. A common method is to subtract a low-order polynomial fitted to the raw spectrum. This removes the slowly varying fluorescence background and yields a Raman spectrum with zero integrated intensity (positive and negative intensity values). Another standard technique is to Fourier transform the spectrum, discard the low-frequency components, and then inverse transform the remaining high-frequency (Raman) components. Experimental and mathematical means to remove the fluorescence background to obtain good quality tissue Raman spectra have been detailed elsewhere (Baraga *et al* 1992b, Baraga 1992, Shreve *et al* 1992, Brennan 1995, Brennan *et al* 1997a).

Several groups have described near-IR Raman instrumentation with CCD detection for biological tissue analysis, both *in vitro* and *in vivo*. The first such application, demonstrated by our laboratory in 1992, used 810 nm laser excitation for *in vitro* laboratory analysis of aorta tissues (Baraga 1992, Baraga *et al* 1992b). Our subsequent work has shown that fluorescence interference in tissue Raman spectra can be reduced even further with 830 nm excitation without compromising CCD sensitivity (Brennan 1995, Brennan *et al* 1997b, Manoharan *et al* 1998). Recently, we described a near-IR Raman *in vitro* laboratory system and a prototype *in vivo* clinical system for tissue analysis (Brennan *et al* 1997b). Using laboratory systems of two designs, we have conducted extensive studies on the quantitative biochemical analysis of coronary artery (Brennan 1995, Brennan *et al* 1997a), biochemical-based diagnosis of coronary (Römer *et al* 1998a, b, Deinum *et al* 1999) and peripheral artery atherosclerosis (Salenius *et al* 1998), and analysis of peripheral blood gases and blood analytes (Berger *et al* 1999).

We have also surveyed the usefulness of near-IR Raman spectroscopy with CCD detection for studying cancerous changes in the colon, urinary bladder, breast and soft tissue sarcomas (Manoharan *et al* 1996). Our results suggest that Raman spectroscopy has less utility in probing cancerous alterations restricted to the epithelial surface, in hollow organs such as the colon and bladder, for example. In part, this is due to the fact that near-IR Raman spectroscopy is a relatively less sensitive detection method than fluorescence or absorption, even though much more spectral information is provided, and in part because near-IR Raman measurements sample deep into the tissue (of the order of 1 mm). Since early cancerous changes in epithelial tissues are subtle, and are primarily confined to the superficial layer of the tissue, the non-diagnostic Raman background from the underlying tissue can dominate the spectrum, reducing the useful signal-to-noise (S/N) ratio.

Nevertheless, promising studies of epithelial changes using near-IR Raman spectroscopy with CCD detection have been reported by other investigators. Puppels and co-workers have surveyed the potential of Raman spectroscopy for characterizing oesophageal cancer and various skin lesions using a fibre-optic Raman instrument (Puppels 1997, Caspers *et al* 1998). Richards-Kortum's group has investigated near-IR Raman spectroscopy to differentiate cervical precancers from non-malignant changes. The spectral differences were attributed to changes in collagen, nucleic acids and phospholipids (Mahadevan-Jensen *et al* 1996, 1998a). Recently, the same group has proposed a fibre-optic probe for *in vivo* studies of cervical precancers and collected initial spectra during colposcopy (Mahadevan-Jensen *et al* 1998b).

On the other hand, the Raman technique is an excellent tool for probing bulk biochemical changes in small solid lesions such as sarcomas and breast cancer. Recently, we reported the potential of near-IR Raman spectroscopy with CCD detection for differentiating malignant lesions from benign lesions in breast tissues (Manoharan *et al* 1998). The details of the instrumentation and applications are presented below in sections 3 and 5, respectively.

Applications of optical spectroscopy, including Raman spectroscopy, for gastrointestinal cancer diagnosis have been reviewed by Bohorfoush (1996). Shim and Wilson (1996, 1997) have also described a near-IR Raman system for clinical studies of tissues, and have examined the spectral effects of tissue drying, snap freezing, thawing and formalin fixation in resected mammalian tissues of various types. The results indicated minimal spectral artefacts due to *in vitro* tissue preparations except for tissue dehydration, which appeared to affect the amide modes of extracellular matrix proteins (Shim and Wilson 1996). McCreery's group used 785 nm laser light to obtain spectra of formalin-fixed normal breast tissue, infiltrating ductal carcinoma and fibrocystic disease (Frank *et al* 1995). The spectra of normal and diseased (benign and malignant) tissues were shown to be dramatically different, but differences between benign and malignant lesions were less pronounced (Frank *et al* 1994, 1995). The spectral differences in normal and diseased breast tissues have been attributed to the fact that normal tissue is mostly composed of fat and diseased tissue is composed of extracellular proteins. Spatial imaging of lipids and proteins in thin breast tissue sections using Raman microscopy has also been reported (Schaeberle *et al* 1996, Kline and Treado 1997).

2.5. UV resonance Raman spectroscopy

As mentioned above, another alternative for circumventing the fluorescence interference problem is to use deep ultraviolet light at wavelengths below 270 nm. When the excitation wavelength matches an electronic energy state of the scattering molecule, this process is known as UV resonance Raman (UVRR) spectroscopy. In biological materials, the high-lying electronic states excited by UV light often rapidly relax via non-radiative processes to lower-lying levels, thereby eliminating fluorescence or causing it to occur at much longer wavelengths. In contrast, Raman-scattered light is generated before relaxation can set in and must, by virtue of energy conservation, be emitted in the UV. The result is that fluorescence occurs outside the spectral region where the Raman bands are observed, so Raman spectra can be obtained without background fluorescence (Asher and Johnson 1984, Fodor *et al* 1986, Spiro 1988, Asher 1988, Nelson *et al* 1992). The principal advantage of UVRR spectroscopy in biomedical applications is that, by tuning the excitation frequency near to the absorption frequency, biochemical disease markers can be selectively observed in the Raman spectrum. Another advantage is that UV excitation can result in a tremendous increase in Raman intensity. This is due to resonance enhancement and the fourth-power frequency dependence of Raman scattering intensity.

UVRR spectroscopy has been extensively used to selectively probe biological macromolecules, organelles and cells, and has been reviewed (Asher 1988, 1993, Hudson and Sension 1989, Nelson *et al* 1992, Manoharan *et al* 1996). Nelson's group has pioneered the application of UV resonance Raman spectroscopy for quantitative *in situ* assay of chemical markers such as quinones, lipids, nucleic acids and proteins in micro-organisms, and for potential detection and identification of bacterial cells (Dalterio *et al* 1986, Manoharan *et al* 1990, 1993b, Nelson *et al* 1992). The utility of UVRR for detection of marine neurotoxins in phytoplankton has been recently reported by Nelson and co-workers (Yao *et al* 1997). UV resonance Raman spectroscopy has also been used to probe the structure of macromolecules in viruses (Grygon *et al* 1988, Wen *et al* 1997, Wen and Thomas 1998).

Normal and cancerous cells have been differentiated based on their biochemical make-up using UVRR. Turpin and co-workers have selectively excited nucleic acids from a single T84 cancer cell using 257 nm excitation (Sureau *et al* 1990). Recently, Richards-Kortum's group has reported the 257 nm-excited resonance Raman spectra of normal and malignant breast and cervical cell cultures (Yazdi *et al* 1999). The spectral differences between

normal and malignant cells have been due to changes in relative amounts of nucleic acid to proteins and nuclear hypochromism. Manoharan *et al* (1995) conducted the initial experiment on the use of UVRR for cancer analysis, using colon tissue specimens. This work was followed by a systematic investigation of normal and neoplastic colon using 240 and 251 nm excitation (Boustany 1997, Boustany *et al* 1999). The spectra were modelled as a linear combination of nucleotide, aromatic amino acid and lipid lineshapes. Neoplastic specimens in general showed a lower aromatic amino acid to nucleotide content ratio, compared with that of their corresponding normal samples. Raman microscopy has been used to confirm the origin of these spectral differences. This study shows that UVRR can be used to quantify biochemical changes in cells for differentiating tissue types (Boustany 1997, Boustany *et al* 1999).

The clinical application of UVRR spectroscopy may be limited because of the problems associated with the photomutagenicity of UV light. The same wavelengths that enhance the nucleic acid vibrational bands in cells and tissues are known to be mutagenic (Rasmussen *et al* 1989, Feld and Kramer 1991). However, the selectivity and sensitivity of the UVRR technique shows promise in analysing cell constituents *in vitro* as in cell cytometry and quantitative analysis of tissue constituents.

The substantial body of research summarized in this section has been facilitated by continuing advances in technology in areas of optics, lasers and spectroscopy. The following section outlines components and systems for modern biomedical applications of Raman spectroscopy.

3. Instrumentation for biomedical Raman spectroscopy

Near-IR Raman spectroscopy has become a potentially powerful tool for determining the biochemical composition of human tissue. Over the last decade, significant advances have occurred in the technology of every component of state-of-the-art laboratory Raman systems. With these advances, Raman spectroscopy has been extended to a wide range of applied problems outside the laboratory, in addition to traditional investigations of a more fundamental nature. For example, Raman spectroscopy is now widely used in many facets of applied materials analysis, process control and environmental monitoring. Today, state-of-the-art components for Raman spectroscopy can be incorporated into systems specifically designed for clinical use. Compact diode laser sources can provide stable narrow-line excitation over a range of wavelengths in the near-IR. High-quantum-efficiency, low-noise CCD detectors allow rapid multichannel data acquisition. Holographic optical components have become integral features of today's Raman systems. Holographic notch filters can attenuate excitation light by a factor of one million, and thereby reduce stray light and associated shot noise. Holographic laser line filters transmit the laser excitation wavelength in a narrow band with 85–95% efficiency, while blocking unwanted laser radiation, and holographic transmission gratings permit compact, high-throughput spectrograph design.

The performance criteria for clinical applications include convenience of use and the ability to access organs provided by optical fibre light delivery and collection, and rapid acquisition of high-quality spectral data, while maintaining excitation intensities sufficiently low to avoid tissue damage. Current near-IR FT Raman or scanning photomultiplier tube-based systems are not useful for clinical applications because of the long data collection times required (Brennan *et al* 1997b). By contrast, the high quantum efficiency and negligible dark noise of cooled, silicon-based CCD detectors, combined with high-throughput imaging spectrographs, permit multichannel acquisition of spectra with excellent S/N ratios in very short collection times.

Raman instruments built for laboratory research, with provision for variable excitation/detection wavelengths and the ability to collect very high-quality spectra, are important for applications development studies. Such systems can be used to optimize excitation wavelengths for a particular tissue type and establish limits on acceptable S/N ratio and data acquisition times. In vitro data taken with a laboratory system can be used to develop and test analytical models and algorithms. For example, a laboratory system that incorporates the capability of confocal Raman microscopy can acquire spectra from individual morphological components of the tissue for use in modelling. On the other hand, Raman systems built for clinical studies have entirely different requirements. They must be mobile, capable of collecting spectra via optical fibres within a few seconds (to minimize motional artefacts due to respiration and pulse), and must conform to hospital safety guidelines.

Several groups have developed laboratory-based systems for tissue Raman spectroscopy (see, for example, Frank *et al* 1994, 1995, Mahadevan-Jansen *et al* 1998a, Shim and Wilson 1997, Puppels *et al* 1990). In the following subsections we describe three systems used in our research. We have developed a versatile laboratory system for acquiring Raman spectra from biological tissues, a second laboratory system designed for acquiring spectra from fluids such as blood and a system designed for clinical investigations. These systems are described in detail below.

3.1. Laboratory Raman system /Raman microprobe for tissue specimens

We have developed a general-purpose laboratory system optimized for bulk tissue measurements and microscopy (figure 4). The excitation source consists of an argon ion pumped Ti:sapphire laser tunable from 720 to 1000 nm, with output power variable up to 1 W. Various holographic laser line filters are available for eliminating extraneous laser radiation outside the chosen excitation wavelength. Various holographic notch filters are used to suppress scattered excitation light while still permitting high transmission from about 50 cm^{-1} above or below the excitation frequency. The Raman-scattered light from tissue specimens is collected and collimated by means of an aberration corrected lens. A holographic notch filter at the appropriate excitation wavelength is positioned in the collimated beam before the focusing lens. The focusing lens, which is f-number matched to a 0.25 m imaging spectrograph, focuses the collected light into the spectrograph entrance slit. The spectrograph has an adjustable slit and a turret capable of holding three gratings for low or high spectral resolution measurements. The dispersed image of the Raman excitation region in the tissue is imaged on an near-IR antireflection coated, back-illuminated, deep-depletion CCD detector that is liquid nitrogen cooled.

This laboratory system also incorporates an optical microscope for obtaining tissue Raman spectra at a cellular level (Raman microprobe, figure 4). A prism is used to steer the laser beam either into the macro-optical system for bulk tissue measurements or into the microscope for microscopic tissue measurements. Excitation is accomplished by directing the laser beam onto the tissue section using the epi-illumination configuration of the microscope, in which the microscope objective both focuses the excitation light and collects the scattered light (Raman backscattering geometry). The microscope can be fitted with either phase contrast or normal objectives, and has appropriate dichroic filtering and imaging optics to allow registration of the focused incident laser spot ($\sim$2 μm diameter) with the white light transilluminated microscopic image of the tissue section. This can be displayed on a high-resolution video monitor or, alternatively, viewed through the microscope eyepiece with the laser beam blocked. With the white light source stopped, Raman-scattered light collected by the microscope objective is redirected to the spectrometer system along the same optical

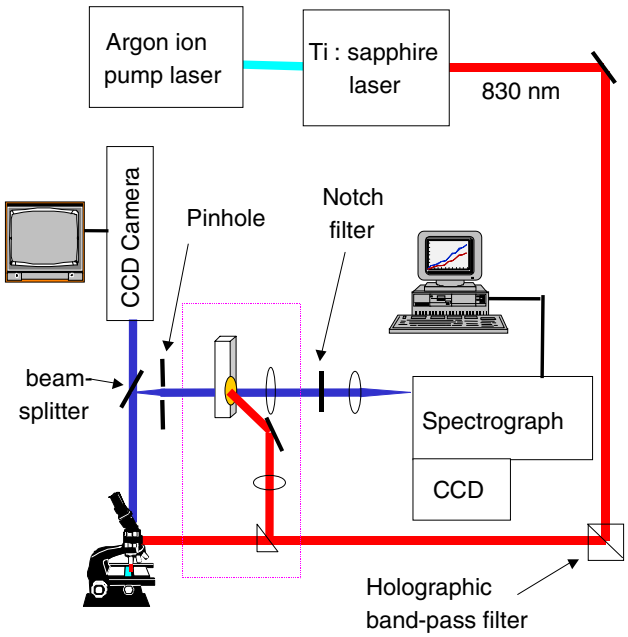


Figure 4. Schematic diagram of the laboratory Raman spectrometer/Raman microprobe system for tissue and tissue sections. Translating the prism into or out of the laser excitation path allows spectra to be collected from bulk tissue or from microscopic regions of thin tissue sections.

train used for the bulk tissue system (Brennan *et al* 1997b). This system has been used to collect Raman spectra from tissue sections as thin as 4 μm, as well as from bulk tissue specimens 4–5 mm thick. Spectrograph, microscope and data acquisition are computer controlled.

3.2. Laboratory Raman system for blood analytes

We have constructed a second laboratory instrument for Raman analysis of blood analytes. This instrument is specially designed for extracting chemical concentrations. In this case, the Raman spectra of the analytes to be measured are small compared with both the Raman spectrum from background analytes and the fluorescence background, and multivariate calibration must be used to extract the desired information (see section 4). This requires an absolutely reproducible light delivery/collection geometry, and light delivery and collection must be optimized, while maintaining spectral resolution.

This system has been designed to study how best to optimize delivery and collection geometries in liquid samples of varying turbidity—whole blood, blood serum, interstitial fluids and saline solution. For this, the delivery and collection geometries can be precisely varied. When used with multivariate techniques, physiological concentrations of analytes (e.g. glucose) at submillimolar levels can be measured. Noting that water is 55 molar, this represents a Raman concentration measurement sensitivity of 10 ppm, illustrating the potential sensitivity of Raman spectroscopy for precise, quantitative measurements.

The system uses a tunable diode laser operating at 830 nm and 500 mW power, a holographic imaging spectrograph, liquid nitrogen-cooled deep depletion CCD detector, and special collection optics (figure 5). The filtered light from the laser is focused onto the sample

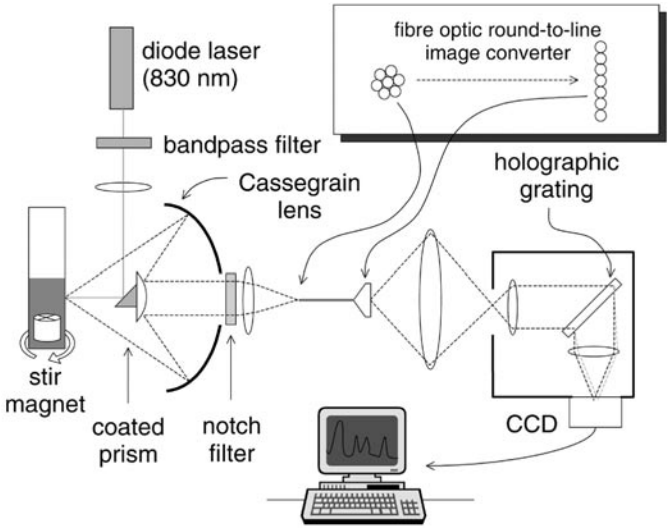


Figure 5. Schematic diagram of Raman spectroscopic system for the analysis of blood analytes. The reflective objective has a high collection efficiency. Inset: The collection efficiency and high throughput are maintained, while preserving resolution, by configuring the f-number matched fibre bundle into a linear array at the spectrograph entrance. This linear array of fibres forms the spectroscopic slit that determines resolution.

in a quartz cuvette containing a magnetic stir-bar to continuously agitate. A sample volume of about 1 ml was used, but smaller volumes can be achieved.

Collection geometry efficiency is an important consideration for quantitative Raman analysis of liquid samples. In clear samples, such as blood serum, the Raman signal originates along an excitation path extending deep into the sample, while turbid samples such as whole blood limit the penetration depth of the excitation light but provide a larger illumination spot, due to elastic scattering. In both cases, the Raman signal intensity is relatively small compared with the background fluorescence, even with 830 nm excitation, and the shot noise generated by the background can result in a low S/N ratio. To increase the efficiency of photon collection, and therefore improve the Raman S/N ratio, a microscopic reflective objective (Cassegrain) of high numerical aperture is incorporated with a special fibre optic bundle. In turbid media, the high numerical aperture allows imaging of a larger portion of the illumination spot. In clear media it accepts the Raman scattered light from deeper regions along the excitation path. The fibre optic bundle consists of 177 fibres arranged concentrically at the sample end and in a linear array at the spectrograph entrance slit (figure 5, inset). This provides high collection efficiency and efficient coupling into the spectrograph without sacrificing resolution. This combination of fibre and imaging optics has a much higher effective numerical aperture than typical conventional or fibre optical systems. Ongoing design improvements are aimed at optimizing the system for measurements in turbid media (Berger 1998).

3.3. Clinical Raman system

A number of investigators have used different configurations to acquire *in vivo* Raman spectra. Using a 785 nm diode laser with fibre optic excitation and conventional optics collection, Mahadevan-Jansen *et al* (1998b) could detect epithelial cervical precancer with good sensitivity and specificity. Shim and Wilson (1997) reported an *in vivo* Raman system, and using an

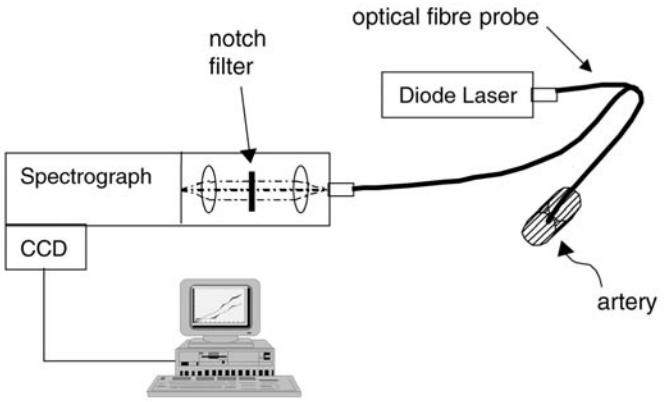


Figure 6. Clinical Raman system. This system is designed for rapid data acquisition, portability and safety in a hospital environment.

optical fibre probe with 785 nm excitation collected spectra from human skin, buccal cheek epithelium, fingernail and tooth. Puppels and co-workers (Puppels 1997, Puppels *et al* 1998) have demonstrated *in vivo* Raman spectroscopy of oesophagus using a commercially developed optical fibre probe.

Our clinical Raman system has been built to be portable, simple to operate and fully functional in a hospital catheterization laboratory or operating room. It incorporates a single-mode, diode laser excitation source operating at 830 nm directly coupled to an optical fibre probe (figure 6). The probe serves as both light delivery source and optical collector, and it is directly coupled to a high-throughput f-number matched holographic imaging spectrograph, providing 10 cm^{-1} resolution. The detector is a liquid nitrogen-cooled back-thinned deep-depletion CCD having 1024 $\times$ 256 pixels. Design considerations to conform with hospital guidelines have been followed: the entire system is enclosed to avoid stray light, and it fits on a small transportable cart that can be placed close to the patient. The optical fibre probe is designed to withstand hospital sterilization procedures and the system power supply is isolated to keep ground leakage current below 70 μA.

A detailed description of this system can be found in Brennan *et al* (1997b). Preliminary spectra of peripheral arteries have been obtained *in vivo* with this system, using data acquisition times of only a few seconds (Salenius *et al* 1998).

3.4. System components

The remainder of this section describes current components and technology for Raman systems suitable for biomedical applications, both *in vitro* and *in vivo*.

3.4.1. CCD considerations. A CCD spectroscopic detector can be thought of as a rectangular array of photosensitive elements or pixels arranged in horizontal rows and vertical columns, a configuration known as a focal plane array (FPA). In spectroscopic applications, the spectral or wavelength direction corresponds to the horizontal rows and the column pixels are usually summed, i.e. 'binned', vertically, providing intensity at each wavelength. CCD detectors have the manufacturing and operational advantages of being fabricated on a monolithic silicon chip. Typical dimensions of CCDs for spectroscopic applications are 1024 $\times$ 256 pixels, each of which is 25 μm on a side, thus covering an area of about 25 mm $\times$ 6 mm. Such arrays

operate at typical readout rates of 50 kHz. CCD technology has improved to such an extent that quantum efficiencies of 90% can be achieved from the visible to the near-IR, a significant wavelength range. Different types of CCDs, front illuminated, thinned back-illuminated and front- or back-illuminated deep depletion, are used for different applications. For the near-IR region, deep-depletion CCDs are best. The depletion region penetrates at least 30 μm into the silicon, thus providing a larger volume to generate photoelectrons, which results in increased quantum efficiency.

A problem that can arise with back-thinned CCDs in the near-IR region (for wavelengths longer than about 750 nm) is the so-called etaloning effect. Due to the larger transparency of silicon at these wavelengths, photons arising from the broad background fluorescence of the tissue are not fully absorbed in the shortened path length of the thinned silicon wafer. The result is that the thinned silicon wafer acts as an etalon for the back-reflected fluorescence, producing sharps peaks in the signal from this otherwise featureless background. Since these peaks are of similar linewidth to the tissue Raman features, they are exceedingly difficult to correct. In Princeton Instruments deep-depletion CCDs (Princeton Instruments, Trenton, NJ), a special surface treatment is used that randomizes the silicon wafer thickness on a wavelength scale. When the CCDs are used for spectroscopy purposes and the pixels are binned vertically, the etaloning effect is effectively eliminated.

In the near-IR region, liquid nitrogen-cooled, silicon-based CCD detectors are efficient out to wavelengths of about 1100 nm, and have extremely low dark noise (5 electrons per pixel per hour) and low read noise (4–7 electrons at 50 kHz). However, shot noise due to tissue fluorescence at these wavelengths still limits Raman sensitivity. Other types of detectors show reasonably high quantum efficiency from 1000 nm out to about 2000 nm, where tissue fluorescence is substantially eliminated. Indium gallium arsenide (InGaAs), germanium (Ge) and indium antimonide (InSb) single-element and linear array detectors can attain quantum efficiencies of about 70% at these longer wavelengths (Chase and Talmi 1991, Barbillat *et al* 1993). However, although tissue fluorescence is suppressed at these longer wavelengths, devices based on these materials typically exhibit levels of read noise three to four orders of magnitude greater than that of silicon-based CCDs at 1000 nm, resulting in measurements limited by detector noise rather than by shot noise. Recently, we conducted a feasibility study using a specially prepared palladium silicide (PdSi) FPA detector (Brennan *et al* 1996). The results are promising, but while the readout noise is not greater than 50 electrons rms (compared with 37 000 electrons rms for Ge and InGaAs linear arrays, and 2000 electrons rms for InSb FPA detectors (Brennan *et al* 1996, Treado *et al* 1994)), the quantum efficiency of the PdSi FPA detector is only 10% at these wavelengths. Nevertheless, working with samples of human aortic atherosclerotic plaque, we were able to collect Raman spectra using 1064 nm excitation. More research must be conducted to optimize design parameters before practical applications can be considered.

3.4.2. Optical filters. In order to take advantage of compact, high-throughput holographic spectrographs and multichannel CCD detection, scattered light at the laser excitation wavelength, typically one million times more intense than the Raman spectrum, must be removed from the detected signal with extremely high extinction efficiency. The advent of holographic notch filters for this purpose has eliminated the need for cumbersome, low-throughput, multistage scanning spectrometers, and was critical for enabling rapid spectral data acquisition using the new technologies. Holographic notch filters can have a very steep-edged blocking bandwidth at the excitation wavelength, with contrast ratios greater than one million, while transmitting 90% of the light with a relatively flat baseline above and below the excitation wavelength. For laboratory systems, where excitation wavelength

versatility is desired, holographic laser line filters are used. Holographic laser line filters block unwanted emission from the laser that would otherwise swamp the Raman spectrum, while transmitting 85–95% of the excitation line. For clinical systems, filter design and construction are paramount for useful fibre optic probes (see below).

3.4.3. Spectrographs. Clinical applications of Raman spectroscopy require rapid acquisition of high-quality spectral data using low-power laser excitation. The emergence of the CCD as the most effective detector element for biomedical Raman spectroscopy applications has driven spectrograph design innovations. CCDs measuring 25 mm in the spectral direction allow multichannel acquisition of the entire range of the Raman fingerprint region (400–2000 cm^{-1}) in the shortest possible time. However, in order to take advantage of the structural detail of tissue Raman spectra, spectral resolution of 8–10 cm^{-1} is required. Fibre optic f-number matched holographic spectrographs provide both the high throughput and flat image field at the detector plane needed for sensitivity at low laser power and spectral resolution throughout the range of interest. Holographic transmission grating designs provide more than 80% throughput, and permit a compact optical layout while providing the required flat image at the CCD detector.

3.4.4. Optical fibre probes. For most medical applications, all components of the fibre optic probe must be of suitable dimensions for incorporation into needles, endoscopes, angioscopes and other clinical devices. Typically, this is of the order of 1 to 2 mm in diameter. To access remote internal organs, the length of the fibre optic elements comprising the probe must be of the order of several metres. Finally, the materials and construction of the probe must be able to withstand hospital instrument sterilization procedures. These requirements lead to mechanical and optical design problems that have not yet been fully resolved.

Optical design problems include spectral interference from the optical fibres themselves. Fibre fluorescence and absorption can be minimized by using readily available, high-purity, low-hydroxyl fibres. However, the dominant optical design problem is generation of Raman scattered light in the optical fibres, which interferes with the Raman signals generated in the tissue sample under investigation. Raman scattering is generated both in the excitation fibres, by the excitation laser light, and in the collection fibres, by laser light scattered from the tissue. Since the pathlength traversed by the laser light in the optical fibres is of the order of metres, the intensity of the fibre Raman signal can be much greater than that of the tissue Raman signal. Simple polynomial fitting or background subtraction techniques, such as are used for removing fluorescence background, are inadequate for removing the highly structured quartz fibre Raman background interference. Rather, appropriate filters are needed to remove the fibre Raman signal. These filters must be positioned at the distal tip of the probe, and therefore must be of submillimetre dimensions. The need to produce dielectric coatings and assemble submillimetre, coated, optical elements with the precision and consistently high quality demanded for clinical Raman applications continues to drive research and optics manufacturing technologies. Our laboratory has been engaged in developing such probes for some time, and probes with dielectric filters of such dimensions with acceptable performance characteristics for *in vivo* application should become available in the near future. Probes with reduced fibre Raman interference have been produced commercially and used in less demanding clinical settings, such as in the oral cavity (Shim and Wilson 1997).

When the optical fibre probe issues are resolved, Raman spectroscopy will be compatible with access via needles, cannulas, angioscopes and endoscopes for *in vivo* detection and imaging. The instrumentation will have reached a threshold at which, in a relatively short time period, Raman spectroscopy can be used to obtain real time, biochemical information about

many organs in the human body, leading to a number of non-invasive and minimally invasive diagnostic techniques.

4. Modelling methods

In current medical practice, tissue analysis provides the basis for definitive disease diagnosis. There are two distinct approaches to tissue analysis: quantitative and qualitative. The qualitative approach usually takes the form of disease classification, which relies on sorting tissue into discrete categories using more or less subjective criteria. Pathologists study haematoxalin and eosin stained biopsy specimens under the microscope, looking for characteristic morphological structures or patterns in the tissue. The presence or absence of certain biochemicals, assessed by a variety of specific histochemical stains, can also provide important diagnostic information. The quantitative approach can take the form of chemical analysis, for example determining the concentration of important blood analytes (such as glucose or cholesterol), or quantitative histology, in which the morphological components of the tissue of interest are counted and characterized quantitatively. These techniques provide quantitative measures of severity of disease and organ malfunction.

Raman spectroscopy can be used to perform both types of analysis: disease classification and chemical and morphological quantitation. The purpose of this section is to describe various mathematical techniques that can be used for extracting information from Raman spectra for these types of analyses. Specific examples of implementation of many of these techniques in different disease settings will be found in section 5.

A Raman spectrum consists of numerous bands, each one due to a specific molecular vibration. Some amount of diagnostic information can be obtained by simply following the changes in one (or several) of these bands, and then correlating them with biochemical or morphological changes in the tissue. However, to make full use of the richness of Raman spectroscopy, more sophisticated modelling techniques that incorporate the full range of data available are required. A Raman spectrum can include as many as a thousand intensity points, each representing a different wavenumber. Each intensity point contains information. Extracting useful, usually biochemical, information from such complex data can be accomplished by using a set of mathematical techniques referred to as multivariate analysis by statisticians, or as 'chemometrics' by analytical chemists. Such techniques enable us to model our data, and thus use to our advantage the plethora of information contained in the spectra.

Raman tissue spectra can be modelled using linear analysis. Two assumptions are associated with this type of modelling. One is that the Raman spectrum of a mixture of biochemicals is simply a linear superposition of the mixture's component spectra. The second is that there is a linear relationship between signal intensity and biochemical component concentration. Manoharan *et al* (1992) conducted a Raman spectroscopy study in mixtures similar in composition to that of biological tissue, and found that both of these conditions hold within experimental limits, allowing the use of Raman spectroscopy to extract precise biochemical concentrations from a mixture. Based on these assumptions, one can model the mixture using a linear superposition of basis spectra. These basis spectra can be the Raman spectra of the pure chemical constituents in the mixture, the Raman spectra of individual morphological components of the tissue, or even mathematically derived lineshapes.

There are two important steps in the process of developing a model, whether to determine concentrations or make diagnoses (Martens and Naes 1989). The first step, calibration, uses data to determine the values of parameters employed in the model. The second step, validation, applies the model and tests it on new, independent data. If the predictions on the new data set

are good, then the model can be considered robust. Otherwise, it must be reconsidered. A model will never perfectly represent reality, and the validation set allows us to establish how good a job it does. This helps protect against artefacts of the initial data set, and 'overtraining'. However, the calibration–validation process usually requires a large data set. In practice, two independent data sets are not always available. In such situations, one can use cross-validation methods to make efficient use of a single data set. One example is the leave-one-out method, in which one sample at a time is left out of the calibration step. The model is developed using the remaining data and then applied to the 'left-out' sample. The predictions made on the 'left-out' data can be compared with the reference value to evaluate the model. By cycling through left-out samples in this way, a relatively small data set can be used efficiently and without bias to establish the robustness of a model.

4.1. Quantitative measurements

One desired use of Raman spectroscopy is to obtain quantitative information from a tissue sample, either biochemical or morphological. Initially, for the purposes of discussion, we will confine ourselves to measurement of biochemical concentrations. As stated above, chemical concentrations can be extracted by using a linear combination of basis spectra to model the data. But how can we do this in practice? Consider a Raman spectrum, $s(\lambda)$, with n wavelengths (say 1000) at which Raman intensity values are measured. Since, by assumption, the tissue response is assumed to be linear, the concentration, c, must be obtainable by multiplying each of these intensity values by a particular number, $b(\lambda)$, and taking their sum:

$$c = \sum_{\lambda} s(\lambda)b(\lambda). \tag{1a}$$

Thus, development of an algorithm to extract concentration information from a Raman spectrum is equivalent to determining the correct values for the 1000 $b(\lambda)$. It is convenient to employ vector notation. We can write

$$c = \boldsymbol{s} \cdot \boldsymbol{b} \tag{1b}$$

where $\boldsymbol{s}$ is a vector with 1000 orthogonal elements, in which each element represents the signal intensity at a specific wavelength. From this point of view, the set of numbers $b(\lambda)$ is also a vector, $\boldsymbol{b}$, which can be called the calibration vector. Note that $b(\lambda)$ also takes the form of a spectrum, with the same number of wavelengths as $s(\lambda)$.

When applying these principles to a set of Raman spectra it is convenient to use matrices. In such a case, a set of spectra taken of different mixtures (same wavelength components, different concentrations) is written as a matrix $\mathbf{S}$, in which each row represents an individual spectrum (vector). (Here, vectors are represented by lower-case bold while matrices are represented by upper-case bold.) $\mathbf{B}$ is now also a matrix, with the jth row representing the calibration vector of the jth chemical constituent:

$$\mathbf{C} = \mathbf{S} \cdot \mathbf{B}^{\mathrm{T}}. \tag{2}$$

$\mathbf{C}$ is now a matrix of concentrations of the j chemical components for each spectrum, and $\mathbf{B}$ is also a matrix, with each row representing the spectrum of the jth $\boldsymbol{b}$ vector. Observing the rules of matrix multiplication, in equation (2) we must use $\mathbf{B}^{\mathrm{T}}$, the transpose of $\mathbf{B}$, i.e. the matrix obtained by interchanging its rows and columns. (For a review of linear algebra, see Leon (1990).)

The task of multivariate analysis is to determine the appropriate calibration vector $\boldsymbol{b}$ for each chemical. There is one and only one 'true' $\boldsymbol{b}$ vector for a given data set; however, a number of different techniques can be used to estimate it. Explicit techniques are those in

which the linear model is completely defined, and therefore the linear superposition properties can be directly applied. Some examples, which will be discussed here, are ordinary least-squares (OLS) and classical least-squares (CLS) (Haaland and Thomas 1988). In other techniques, called implicit techniques, not all of the basis spectra are known, nor can they be calculated (Geladi and Kowalski 1986). Principal component regression (PCR) and partial least-squares (PLS) are examples of implicit techniques. Explicit methods are more robust, but they require knowledge of the composition of the mixture to be completely characterized. A hybrid technique recently developed in our laboratory, hybrid linear analysis (HLA), is another method, which is useful when some but not all of the basis spectra are known (Berger *et al* 1998).

As mentioned above, explicit analysis can be used when all of the chemicals present in the mixture are known, and the spectrum of each is also known. For example, OLS can be used to determine the relative concentrations of glucose, urea and ethanol in water. In this case, the Raman spectrum of each pure chemical component is easily obtained, and the Raman spectral data set can be expressed in terms of the concentration values, $\mathbf{C}$, and the matrix of known pure chemical component spectra, $\mathbf{P}$:

$$\mathbf{S} = \mathbf{C} \cdot \mathbf{P} \tag{3}$$

with $\mathbf{P}$ a matrix in which each row represents the spectrum of the jth pure component basis spectrum.

The task of OLS is to obtain the concentrations, $\mathbf{C}$, from the observed Raman spectra, $\mathbf{S}$, and the basis spectra of the pure components, $\mathbf{P}$. This can be accomplished by means of a straightforward linear algebra manipulation, to obtain

$$\mathbf{C} = \mathbf{S} \cdot \mathbf{P}^{\mathrm{T}} \cdot (\mathbf{P} \cdot \mathbf{P}^{\mathrm{T}})^{-1} \tag{4}$$

in which $\mathbf{P}^{-1}$ is the inverse of $\mathbf{P}$, constructed so that $\mathbf{P}P^{-1} = \mathbf{I}$, with $\mathbf{I}$ being the identity matrix (i.e. the matrix with ones down the diagonal running from upper left to lower right, and zeros elsewhere). Equation (4) gives the most accurate possible values for $\mathbf{C}$, given the imperfections due to noise in the known Raman spectra of the individual chemical components, and is equivalent to least-squares fitting the data, hence the name ordinary least-squares. Comparing equations (2) and (4), we have

$$\mathbf{B}^{\mathrm{T}} = \mathbf{P}^{\mathrm{T}} \cdot (\mathbf{P} \cdot \mathbf{P}^{\mathrm{T}})^{-1}. \tag{5}$$

Thus, in OLS, once the spectra of the pure components are known, one can construct the calibration matrix $\mathbf{B}$, from which all of the concentrations of an unknown sample can be readily obtained using equation (2).

In CLS, a closely related technique, the spectra of the chemical constituents are unknown, but their concentrations are known. An example of this situation might occur when dealing with biochemicals which cannot be isolated from the mixture in their biologically active form for spectroscopic analysis, but whose concentrations might be determined using other analytical chemistry techniques. In this case, one can solve for $\mathbf{P}$ using linear least-squares fitting and then determine the $\boldsymbol{b}$ vector of each chemical using OLS (see Haaland and Thomas (1988) for details).

In order for explicit methods to be employed, prior knowledge of all the chemical components of the mixture is required (either the concentration or the Raman spectrum of each chemical component). Implicit methods, such as PCR and PLS, do not require as much prior knowledge. Both of these techniques compute $\boldsymbol{b}$ directly from a set of calibration spectra in which only the concentration or basis spectrum of the one constituent of interest is specified. All of the other constituents in the samples may vary freely from sample to sample. An example in which these techniques are used is in determining the concentrations of analytes, such as

glucose or triglycerides in blood (see section 5.4). Remarkably, the $\boldsymbol{b}$ vector for the analyte of interest can be obtained and, even more remarkably, it is the same (except for calibration noise and uncertainties) as the corresponding $\boldsymbol{b}$ vector obtained using OLS or CLS!

PCR and PLS are often called data compression techniques, because they minimize the number of parameters needed to characterize the measured spectra. Although a spectrum may have 1000 points, it will not contain 1000 pieces of independent information, since a well-defined relationship exists among the spectral values of the underlying chemical constituents. Thus, in principle, only a small number of basis spectra are needed to explain all the variations in a data set. Each measured spectrum can then be described by the fit coefficients of these basis spectra.

Principal component analysis (PCA), the first step in PCR, creates a set consisting of a small number of specially chosen basis spectra, called principal components, which can accurately characterize all of the spectral changes in a set of spectral data. These principal components can be calculated using singular-valued decomposition (Wold *et al* 1987). The first principal component accounts for the maximum variance in the data, the second principal component for the next greatest variance, and so on, until additional principal components describe only noise. These last principal components are discarded. A linear combination of the principal components can then be used to model the measured spectra. Because of the way in which principal components are generated, the fits of the principal components to all of the spectra in the data set will be excellent. The fit coefficients in this linear model are often called scores. The principal component scores can then be correlated with the concentrations of a specific chemical in a least-squares step known as regression. This process, in which principal components are used to estimate the $\boldsymbol{b}$ vector, is called principal component regression (PCR).

In PLS, this process is refined to incorporate concentration information into the process of determining lineshapes (Geladi and Kowalski 1986). The first basis spectrum is still chosen to explain the maximum variance in the data, but the fit coefficient is constrained to be proportional to the concentration of the particular component being measured, rather than a free parameter, as it is in PCR. The lineshape that is derived is thus directly related to the concentration of this component. This process can be repeated for each component of interest. Geladi and Kowalski (1986) described the mathematical algorithm of partial least squares (PLS) analysis and Haaland and Thomas (1988) showed how PLS analysis of spectroscopic data could be applied to concentration measurements. Kramer (1998) showed the performance of PLS by simulating concentration measurements from spectral data.

The most appropriate analysis technique to use in a given situation depends on the information available. OLS and CLS are the best techniques to use if the spectral components of the system are completely characterized, while PCR, PLS or similar techniques are required when calibration information about only one or more spectral components is available.

Hybrid linear analysis (HLA), a multivariate method used to extract the concentration of a particular constituent, is a hybrid technique that combines the advantages of both the explicit and implicit methods (Berger *et al* 1998). Consider our earlier example of measuring the concentration of glucose in blood. Instead of basing our determination of $\boldsymbol{b}$ solely on the information produced by the calibration step, with HLA we can incorporate additional physical information in the form of the known glucose spectrum. During the calibration step, the spectrum of the desired component, in our example glucose, is subtracted from the mixture spectra in proportion to its known concentration (measured using a reference technique). PCA is then used to characterize the residual spectra (the original spectra minus the weighted glucose spectra). The calibration vector, $\boldsymbol{b}$, is then calculated using the portion of the pure chemical spectrum that does not overlap with the principal components. HLA thus gives a refined estimate for $\boldsymbol{b}$; hence, concentration accuracy is improved.

For the purposes of HLA, a Raman spectrum can easily be obtained in the case of glucose, but not in the case of less well chemically defined blood analytes, such as triglycerides. Many blood analytes are loosely defined categories of chemicals, which encompass a wide range of species with similar biological properties. Triglycerides have the same basic structures but can have many different fatty-acid side chains, each chemically very similar to the next, but with a slightly different Raman spectrum. In such situations, it is difficult to define and measure a single 'true' Raman spectrum and multivariate techniques such as PLS are suggested.

So far, we have discussed techniques that enable us to extract quantitative information from Raman spectra. These methods are most commonly applied to determining chemical concentrations. However, the Raman spectra of specific morphological features, obtained using Raman confocal microscopy, can also be used as basis spectra. An example of this type of analysis technique will be found in section 5.1.

In summary, the multivariate methods discussed above all have the goal of finding the $\boldsymbol{b}$ vectors of the system. The different methods use different basis vectors to carry out the analysis. OLS and CLS use chemical or morphological spectra as the building blocks of the model, PLS and PCR use principal components as statistical building blocks, and HLA combines statistical techniques with chemical or morphological information.

4.2. Disease classification

Multivariate methods can also provide information that can be used to classify tissue according to disease categories, which are used to direct medical treatment. One can, of course, simply seek to correlate prominent features in the Raman spectra with the disease diagnoses, for example the spectral intensities at several specific wavelengths. However, multivariate analysis provides more powerful tools for spectral classification. In particular, the three types of lineshapes discussed above, chemical, morphological and principal components, can be used as basis spectra to model the data. The scores (or fit coefficients), of these basis spectra can also be used as parameters, referred to as input variables, for developing diagnostic algorithms.

Many techniques have been developed to aid researchers in producing decision algorithms to classify data according to specific parameters. These include discriminant analysis (Fisher and Van Belle 1993), logistic regression (Fisher and Van Belle 1993, Pagano and Gauvreau 1993) and neural networks (Robb and Munk 1990). (Numerous other methods can be found in Sharma (1996).) Discriminant analysis and logistic regression are very similar, linear techniques. Both assume that the samples are binomially distributed within each classification. Discriminant analysis, which applies Bayes' theorem, is the more restrictive in its assumptions. It allows only for normally distributed input variables, whereas logistic regression allows for discrete (e.g. only integer values) or categorical (e.g. male or female) variables. This means that logistic regression can incorporate binary information into a model. A third commonly used classification method, artificial neural networks, simulates the processing of signals in the brain. In such a system information is processed by a 'network' of interconnected units according to a pattern established during the calibration process. Neural networks do not require the data to be normally distributed, and can be used to describe nonlinear systems as well as linear ones. However, it is especially important to perform proper calibration and validation of neural-network-based decision algorithms because physical interpretation of the coefficients is difficult and robustness is not always easy to achieve.

Logistic regression is well suited to the purpose of developing classification algorithms from Raman spectral data, correlating linear variables with disease classifications obtained from a pathological examination. Logistic regression allows us to monitor the model as it is

being built, but allows us to classify our samples according to discrete categories. As it is used extensively in section 5, we shall discuss it here in further detail.

The underlying assumption of logistic regression is that the probability, p_i, that the ith sample in a set of spectral data belongs to a particular category, for example that the tissue of interest is malignant, is described by the logistic function, which is of the form

$$p_i = \left[1 + \exp\left(\alpha + \sum_j \beta_j x_j\right)\right]^{-1} \tag{6}$$

with x_j the score associated with the jth basis spectrum, β_j the corresponding amplitude (or weighting coefficient) and α a constant offset. This probability varies from zero to one in a sigmoidal fashion. It is small for large values of x_j and approaches unity for small values, hence it has the appropriate form for probabilistic modelling. The parameters β_j and α are to be determined.

Logistic regression introduces the 'likelihood function', L, defined as the probability of observing a specific set of diagnoses in a set of samples. This can be calculated as the product of the probabilities, p_i, that each individual sample belongs to a given category:

$$L = p_1 \times p_2 \times \cdots \times p_n. \tag{7}$$

The first step in implementing logistic regression is to decide on a particular basis set (e.g. principal components), and apply these basis vectors to a set of calibration data to obtain a set of scores, x_j, for each of the i samples. To determine the values for α and the β_js, the maximum likelihood principle is used, which states that the values for the parameters of the likelihood function should be chosen to maximize the probabilities predicted for obtaining the specific set of diagnoses made for the sample set.

The expression for L is obtained by inserting equation (6) into equation (7), and the values of α and the β_js are then found by applying this expression to the calibration data set (using the already determined x_js for each of the i samples in the calibration set). The values of α and the β_js are chosen so as to maximize the agreement between L and the known classifications of the calibration data set. Computer programs such as STATA (STATA Corporation) and SAS (SAS Institute Inc.) can be conveniently used to perform this calculation. Once α and the β_js have been obtained, equation (6) can be used to predict the probability that a new sample falls into a given category.

Section 5 contains case studies of various applications of Raman spectroscopy for chemical analysis and medical diagnosis. In each case, the modelling principles discussed here will be applied.

5. Diagnostic applications of Raman spectroscopy: case studies

The theme of this article is the potential of Raman spectroscopy to provide accurate diagnosis of human disease. Four important clinical applications of Raman spectroscopy are detailed in this section: diagnosis of atherosclerosis, breast cancer and Alzheimer's disease and measurement of blood analytes. The discussions are based on current projects in our laboratory that employ the experimental tools and analytical methods discussed in the earlier sections. In each case we discuss the medical considerations, instrumental design, data acquisition, and data analysis.

5.1. Atherosclerosis

More than 58 million Americans have some form of cardiovascular disease, to which one out of every 2.4 deaths in the United States, or 900 000 deaths annually, can be attributed. Of these,

half are caused by coronary artery disease. Coronary heart disease remains the number one cause of death in the United States for both men and women. Almost 14 million Americans have a history of coronary artery disease and over a million will have a new or recurrent heart attack during the year—one event every 30 s. More than half of those stricken will die within an hour of the onset of the acute symptoms. The estimated direct and indirect cost of caring for patients with coronary artery disease in the United States in 1998 was greater than 95 billion dollars (American Heart Association 1998).

'Hardening of the heart arteries', the development of atherosclerotic obstructions in the blood vessels (coronary arteries) of the heart muscle, underlies this enormous national health problem. Despite decades of research, the atherosclerotic disease process remains incompletely understood, making it difficult for clinicians to predict the likelihood of a sudden cardiac event in individual patients. Recent medical thinking supports two important hypotheses: (a) the development of atherosclerotic plaques in the arterial wall is an active process involving multiple chemical moieties, not the passive process it was once thought to be, and (b) the biochemical composition of atherosclerotic plaques is the most important indicator of future cardiac complications (Mitchinson 1982, Wexler *et al* 1996, Demer 1997, Mach *et al* 1997). Thus, a better understanding of this disease requires that techniques be developed to allow the collection and analysis of chemical information from atherosclerotic plaques in living patients.

At present, most of the diagnostic techniques commonly used to study patients with atherosclerotic vascular disease do not assess the biochemical composition of the vessel wall. These techniques include angiography, nuclear magnetic resonance imaging (MRI), electron beam computed tomography (EBCT) and intravascular ultrasound (IVUS). Angiography is a catheter-based procedure used to determine the degree of obstruction present in a given artery. A radio-opaque dye solution is injected into the coronary arteries via a catheter, and x-ray images are taken. This simply shows the degree of stenosis (i.e. narrowing of the vessel lumen) present in the arteries. No biochemical information is obtained.

MRI is a non-invasive technique that uses radio waves to probe the nuclear magnetism of the water component of biological tissue. Imaging information based on tissue density is generated by applying a magnetic field gradient to the tissue. Although widely used in medical imaging, there is difficulty in applying this technique to coronary arteries because they are so small (only a few millimetres in diameter). Despite this obstacle, MRI is now able to distinguish the various layers of coronary arteries, as well as the morphological structures, such as the fibrous cap and necrotic core, found in atherosclerotic plaques (Toussaint *et al* 1996). However, MRI has difficulty detecting small but clinically significant microcalcifications within atherosclerotic plaques, and it does not provide significant biochemical information (Wexler *et al* 1996).

EBCT is a non-invasive tomographic technique in which a beam of electrons is passed through the patient's body and structures are imaged by measuring the attenuation of the beam. Because electrons have a short wavelength, EBCT has a relatively high spatial resolution. However, it is not very site specific, and improvements in reproducibility are required to enable longitudinal patient studies (i.e. repeat patient studies over time) (Wexler *et al* 1996). IVUS, the most accurate and quantitative technique currently available, uses acoustical waves delivered via catheter and reflected from the arterial wall, to image various layers of the artery wall and the structures within atherosclerotic plaques (Peters *et al* 1994). Information is even provided as to whether or not the obstructing lesion is atherosclerotic or, conversely, due to a thrombus (Fishbein and Siegal 1996). Indeed, this technique is increasing in usage, but again, microcalcifications are insufficiently diagnosed and biochemical information is lacking (Wexler *et al* 1996).

Two newer experimental techniques that do provide significant biochemical information about the atherosclerotic plaque are based on the optical properties of biological tissue. Fluorescence spectroscopy has been demonstrated *in vitro* as an accurate tool for classifying arterial tissue into three disease categories: non-atherosclerotic, non-calcified plaque and calcified plaque, based largely on the presence of ceroids (Richards-Kortum *et al* 1989, Bartorelli *et al* 1991). However, the fluorescence properties of many biochemicals found in tissue are similar, and thus there is a limited amount of information in the broad, indistinct signal obtained from this technique, making it difficult to analyse tissue composition. On the other hand, Raman spectroscopy exhibits sharp spectral features unique to each Raman active chemical, thus yielding a fingerprint for those biochemicals present in the tissue of interest. Significant work has been done *in vitro* to demonstrate the utility of this technique. *In vivo* work will soon begin, and holds great promise (Puppels *et al* 1998). The *in vivo* application of Raman spectroscopy to the study of atherosclerosis will be performed using optical fibre probes incorporated into catheters used during standard angiography.

After reviewing the general pathology of atherosclerosis and some associated biochemistry, the methods used to apply Raman spectroscopy to arterial tissue will be elucidated. Recent *in vitro* results will then be surveyed and subsequently discussed, along with prospects for *in vivo* studies and possible future directions and applications.

5.1.1. Anatomy, pathology and biochemistry. Atherosclerosis is a disease that affects the large and medium sized vessels of the arterial tree. The two major types of vessels of interest are elastic and muscular arteries. The gross anatomy of both types is similar. These arteries are composed of three layers: the tunica intima, which borders the lumen through which the blood passes; the tunica media, which provides structural support, metabolic activity and synthesis of the extracellular matrix; and the tunica adventitia, which is a network of connective tissue anchoring the artery, supporting nerves and the vasa vasorum, the network of microvessels which carries blood and provides oxygen and nutrients to the artery tissue. The media is normally bounded by two thin, elastin-rich connective structures, the internal and external elastic lamina, which separate it from the intima and adventitia, respectively. The adventitia is not affected in atherosclerosis and thus requires no description other than noting that it is primarily composed of collagen-rich connective tissue and adipose tissue (Rubin and Farber 1994).

The normal intima of elastic and muscular arteries is composed of a single, continuous endothelial layer and the internal elastic lamina. The difference between muscular and elastic arteries lies in the media. The larger vessels of the body are elastic. The aorta is an example of an elastic artery. Here, the media contains smooth muscle cells and a dense network of elastic fibres that are composed primarily of the protein elastin. The small and medium-sized arteries, including the coronaries, make up the muscular class with which we will primarily be concerned. The media of these vessels is lacking the elastic fibres found in larger arteries. They still exhibit an internal and external elastic lamina. However, the former is fenestrated and allows for the migration of smooth muscle cells into the intima (Rubin and Farber 1994).

Proliferation of smooth muscle cells and their subsequent migration into the intima of the vessel are the initial signs of atherosclerotic plaque development. Along with this change, there is also an accumulation of lipids from the bloodstream in the arterial wall. Blood derived mononuclear phagocytes (i.e. macrophages) then begin to infiltrate the arterial wall and engulf the lipids. These become the so-called foam cells. As the disease progresses, these foam cells die and release their phagocytosed lipids. Eventually, a large core of extracellular lipid accumulates underneath the endothelium, which becomes reinforced with collagen fibres, producing a fibrous cap. At the same time, the elastic laminae begin to fragment (Mitchinson

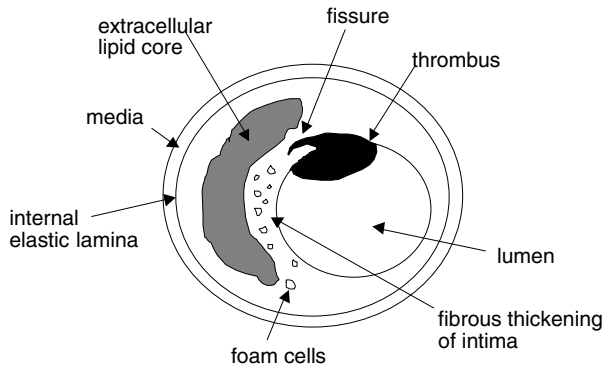


Figure 7. Schematic diagram of an advanced atherosclerotic lesion exhibiting intimal thickening, stenosis and thrombus formation. (Figure adapted from Stary *et al* (1995), with permission, Lippincott, Williams & Wilkins.)

1982, Rubin and Farber 1994). Figure 7 shows a schematic representation of a complicated atherosclerotic lesion.

Advanced lesions begin to exhibit calcifications, mostly in the form of calcium hydroxyapatite, probably using lipids or debris from necrotic foam cells and smooth muscle cells as a nucleus (Wexler *et al* 1996, Demer 1997). Recent developments show that the calcification process is active, rather than passive as once thought, thus suggesting a reparative purpose (Wexler *et al* 1996). Most likely, the calcification is an attempt by the artery to support a progressing lesion and prevent rupture.

The pathology just described leads to a gradual narrowing of the lumen, which can be monitored by the existing techniques. The current focus of research on atherosclerosis is on unstable plaques, composed of a thin fibrous cap overlying a large core of necrotic lipid material. Such plaques comprise the greatest threat for heart attack, because when the fibrous cap ruptures there is subsequent thrombosis and acute occlusion (i.e. complete obstruction). Evidence suggests that such plaques are those of only moderate stenosis (less than 50% narrowing) and that the instability is due to the chemical composition of the core and possibly the presence of a partially calcified plaque (Fishbein and Siegal 1996, Mann and Davies 1996, Wexler *et al* 1996).

One primary chemical concern is the accumulated lipids (Mann and Davies 1996, Demer 1997). However, several proteins, such as osteopontin, osteonectin and gamma-carboxylated protein involved in the calcification process, are found in developing plaques. Proteases, such as metalloproteinase (MMP), contribute to the weakening of the fibrous cap and elastic lamina and have also been identified in plaques that have ruptured (Mann and Davies 1996, Wexler *et al* 1996, Mach *et al* 1997). Once the plaque has ruptured, the growth factors expressed by macrophages and smooth muscle cells are exposed to the circulating blood and form a thrombus via platelet aggregation. This acute occlusion is potentially life threatening and accounts for a large fraction of deaths due to coronary artery disease (Fishbein and Siegal 1996, Wexler *et al* 1996).

5.1.2. Methods. Our laboratory has employed three main methods to analyse Raman spectra collected from arterial tissue (see section 4). The earliest developed technique speaks directly to the current clinical concerns: chemical modelling of the spectra based on the Raman signals of individual constituents. A subsequent method, based on principal component analysis, provides a non-parametric means to study the disease. Finally, if one makes a few very

plausible assumptions, a Raman spectroscopic morphological model of atherosclerosis should be possible. This would provide the cardiologist with quantitative information about the microscopic composition of the arterial wall during cardiac catheterization, at a level consistent with histological techniques performed by a clinical pathologist, but without the need for tissue removal. Thus, a morphological model would enable longitudinal studies of patients and evaluation of various interventional methods.

The *in vitro* data used for the modelling were collected from coronary arteries obtained from explanted hearts. The tissue was snap frozen in liquid nitrogen and stored at $-80\,^{\circ}$C. Just prior to data collection, the bulk intact arteries were allowed to passively thaw and, subsequently, were kept moist with phosphate buffered saline to prevent desiccation. The intact arteries were investigated with the laboratory Raman system discussed in section 3. The tissue was excited with 350 mW of 830 nm laser light for 10–100 s. After data collection, the examined 1 mm^3 spot was marked with India ink and then fixed in formalin. These samples were then subjected to standard histology, i.e. examination to determine their disease state. Two data sets were collected so that a validation study could be performed. The calibration set comprises 97 samples, while the validation set contains 68 samples.

The basis spectra of the chemical model were obtained by extracting various chemicals, including cholesterol, cholesterol esters, triglycerides, phospholipids and calcium salts and proteins from arterial tissue (Brennan *et al* 1997a). The morphological model was developed using the confocal microscopy set-up of the Laboratory system. Examining 6 μm thick sections of the intact arteries under 63$\times$ magnification the major morphological structures of the arterial wall can be appreciated. Laser light is focused on each type of structure, and the resulting Raman spectra are collected and used as a basis set to model the spectra from the bulk arteries.

In vivo data will be collected using the clinical system described in section 3 (Brennan *et al* 1997b), which allows for direct coupling between a portable Raman apparatus and optical fibre probes. The established *in vitro* models will then be tested with the *in vivo* Raman spectra of arterial tissue to determine the effects of environmental factors such as blood and probe positioning stability.

5.1.3. Results. Typical spectra of various stages of atherosclerosis in coronary artery tissue are shown in figure 8. The slowly varying background fluorescence has been subtracted by fitting to a fourth-order polynomial. The molecular vibrations associated with various peaks are labelled. For example, calcium hydroxyapatite present in calcified lesions exhibits a pronounced peak at 960 cm^{-1} that is readily identified. Similar distinguishing features, though more subtle spectroscopically, are present to mark the course of the disease.

Initial *in vitro* studies of human arteries focused on comparison of histochemical analysis using Raman spectroscopy with standard laboratory techniques such as histochemical staining and electrophoresis (Manoharan *et al* 1992). Excellent correlation with standard methods showed that Raman spectroscopy could quantify the relative percentage weights (i.e. concentrations) of the major proteins (collagen and elastin), cholesterol and cholesterol esters in human aorta. Further chemical modelling of atherosclerosis has been extensively demonstrated as an *in vitro* technique. Brennan and co-workers (Brennan 1995, Brennan *et al* 1997a) derived Raman basis spectra from seven biochemicals found in normal and atherosclerotic arteries, and used these to model the Raman spectra of homogenized coronary arteries. Römer *et al* (1998a) then used this model to accurately classify coronary artery lesions into three distinct disease categories: non-atherosclerotic, non-calcified plaque and calcified plaque, employing logistic regression (section 4). This was accomplished using only two pieces of the chemical information provided by the Raman measurements: the relative percentage weights of total cholesterol (esterified and non-esterified) and the amount of calcium salts (figure 9). Figure 10

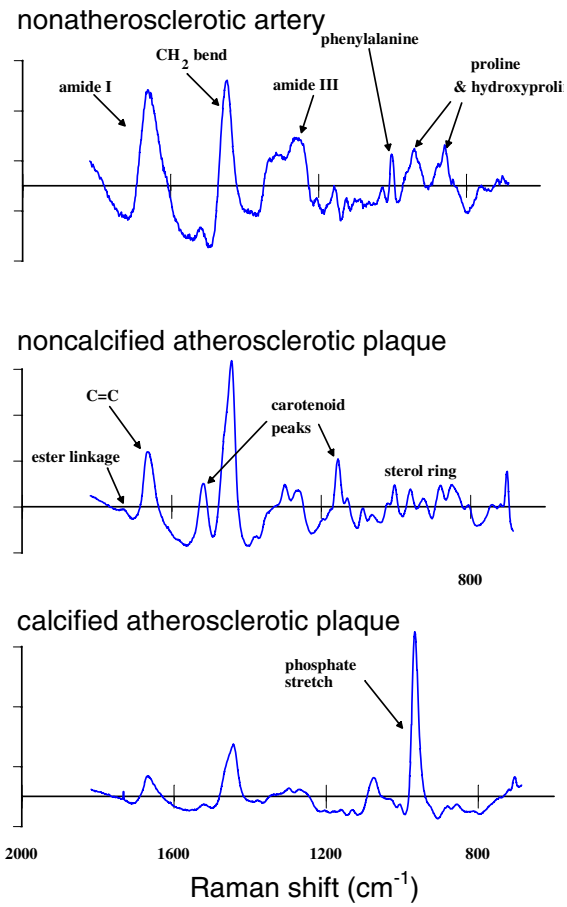


Figure 8. Examples of Raman bands from coronary artery tissue, excited with 830 nm laser light, in various stages of atherosclerosis. Note the prominent phosphate stretch of calcium hydroxyapatite at 960 cm^{-1}, present only in the calcified plaque spectrum. (From Brennan *et al* (1997a) with permission, Lippincott, Williams & Wilkins.)

demonstrates how well the chemical basis set can model the Raman spectra from intact coronary artery. Raman spectra of coronary arteries exhibiting various stages of disease are plotted along with the model fit and the residual (the actual spectrum minus the model fit).

Very recently, Salenius *et al* (1998) have demonstrated that this chemical model is robust and not limited to coronary arteries. Applying similar techniques to femoral and carotid arteries, these investigators have also been able to distinguish normal artery from calcified and non-calcified plaques. This study lends credence to the likely initial *in vivo* studies which, for ethical reasons, will need to focus on arteries other than coronaries to allow for histochemical comparison.

The power of a purely statistical method was demonstrated by Deinum *et al* (1999), who used principal component analysis (section 4) to characterize arterial spectra. The same data set of Römer *et al* (1998a) was used. These investigators were able to accurately classify coronary arteries into the same three classes mentioned above. This prospective study showed an accuracy of greater than 90% using two principal components to diagnose coronary artery disease.

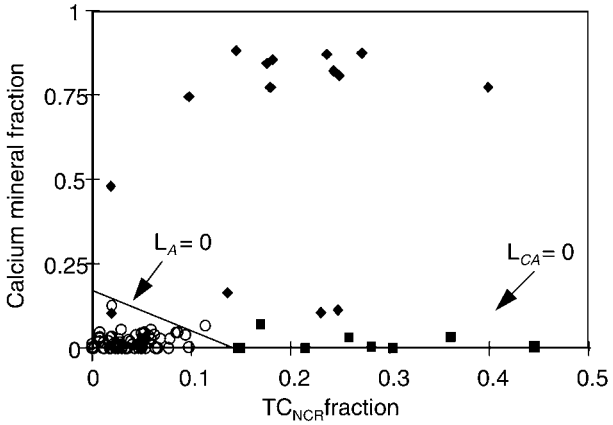


Figure 9. Diagnostic algorithm of Römer *et al* (1998a), based on the total cholesterol (TC_{NCR}) and calcification relative weight fractions. In this decision plot the disease state of a given artery specimen can be classified: (○) non-atherosclerotic tissue, (■) non-calcified atherosclerotic plaque, (♦) calcified atherosclerotic plaque. (With permission, Lippincott, Williams & Wilkins.)

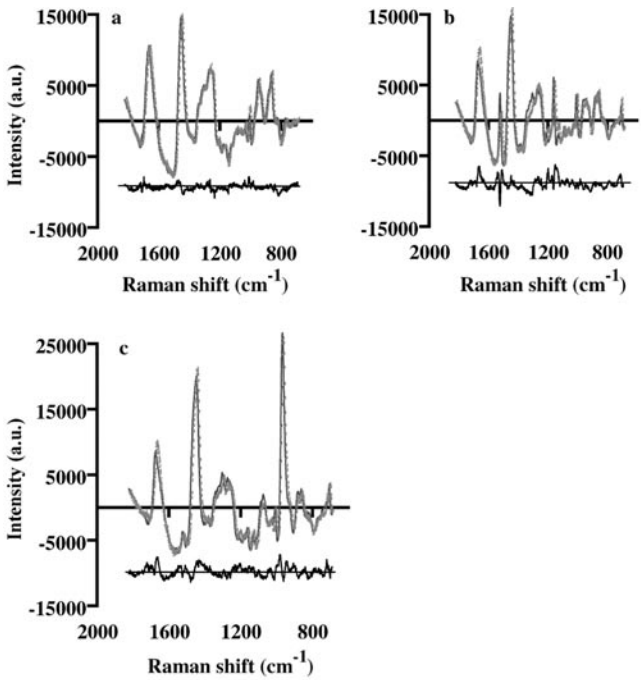


Figure 10. Chemical model fits (curve) and spectral data (dots) for coronary arteries in various stages of atherosclerosis: (*a*) non-atherosclerotic tissue, (*b*) non-calcified atheromatous plaque and (*c*) calcified plaque. The residuals are plotted below the fits. (From Brennan *et al* (1997a) with permission, Lippincott, Williams & Wilkins.)

To demonstrate the interrelationships between the principal component and chemical techniques, Deinum *et al* (1999) then analysed the two diagnostic principal components in terms of the chemical basis spectra of Brennan *et al* (1997a), using least-squares fitting. It was

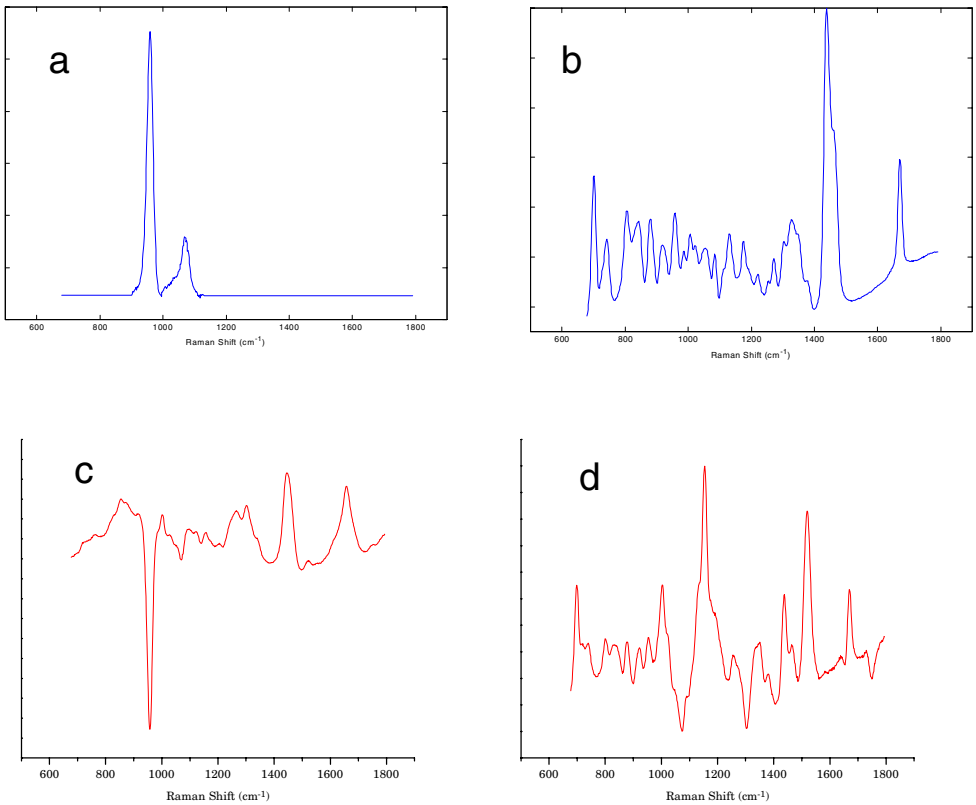


Figure 11. Diagnostic chemical basis spectra from Brennan *et al* (1997a): (With permission, Lippincott, Williams & Wilkins): (*a*) calcification and (*b*) cholesterol, compared with the diagnostic principal components of Deinum *et al* (1999) (with permission, Society for Applied Spectroscopy): (*c*) and (*d*). Note the prominent spectral feature at 960 cm^{-1} in both (*a*) and (*c*), and the similar vibrational structure detail between 700 and 1000 cm^{-1} in (*b*) and (*d*).

determined that the two principal components were primarily composed of (a) cholesterol and β-carotene (positively correlated chemicals) and (b) calcium salts and protein (negatively correlated constituents). Figure 11 shows the similarity between the peaks in the diagnostic basis spectra of these two studies. This technique is powerful because it provides an additional link to determining which and how many of the chemicals are diagnostically significant. This type of analysis is especially crucial for diseases such as breast cancer, where the chemical make-up of the tissue is even less well understood.

Despite the turbidity of biological tissue, there is a relatively low cross section for elastic scattering of the excitation light in the near-infrared. This, combined with the deep sampling volume, makes it possible to use near-infrared Raman spectroscopy to image and map arteries, as suggested by Baraga *et al* (1992a). Römer *et al* (1998b) have recently done just this using APOE*3 Leiden transgenic mice, a well-established model for atherosclerosis in humans (Groot *et al* 1996). Sampling at a rate of 0.5 mm increments, they examined the first 7–8 mm of aorta from mice that were fed diets intended to induce different disease states. They then created 0.5 mm grid maps of the relative percentage weights of total cholesterol and calcium salts found in the aorta as determined by a chemically derived Raman spectral model. Upon mapping the dissected aorta, they determined that total cholesterol and calcium salt deposits

occurred in a reproducible manner (within the first 5 mm of the aortic arch) in all diseased mice. They also established, in concert with findings from other methods, that calcium salt deposits occur only within the area of plaque formation (Wexler *et al* 1996).

The final type of modelling to be discussed relates to the morphological features of arterial tissue. When pathologists use histology to diagnose atherosclerosis, they are concerned with the morphological features of the tissue section of interest. Features such as the number of foam cells or the size of the necrotic core are taken into consideration when classifying the tissue. Thus, in order to diagnose atherosclerosis in a manner similar to that used by pathologists, but using spectroscopy, it is logical to focus on these morphological features. As mentioned above, this paradigm rests on a few plausible assumptions. First, morphological modelling, like chemical modelling, is dependent on the linearity of the Raman effect. Second, it must be assumed that there is little patient-to-patient spectral variation for a given morphological structure. Finally, there must be sufficient differences between the Raman spectra of the different morphological structures to allow differentiation at a level comparable to histological examination.

Indeed, recent results from our laboratory, currently being prepared for publication, suggest that there is little variation in the Raman spectra of any given class of morphological structures within a single patient or among different patients. Therefore, a mean spectrum averaged over many such features can be taken, for example, to represent a 'typical' smooth muscle cell. Moreover, we have also found that there are significant spectral differences between most of the morphological features found in normal coronary arteries and atherosclerotic plaques.

We have recently collected Raman spectra from the major morphological structures found in coronary arteries, including internal and external elastic lamina, adventitial fat, smooth muscle cells, foam cells, necrotic core, cholesterol crystals, carotene containing crystals, calcifications and collagen fibres from the fibrous cap. By modelling Raman spectra from intact arteries with this basis set of morphological structures, it should be possible to correlate the fitting parameters for each structure with the information provided microscopically, to render a morphology-based spectroscopic diagnosis. Figure 12 shows how well this morphological basis set can model artery spectra exhibiting various degrees of atherosclerosis. Furthermore, by relating the morphological spectral basis vectors to the chemical basis vectors discussed earlier, it should also be possible to monitor the individual stages of atherosclerosis on a chemical level. The beauty of this linear technique lies in the correspondence of one model with another. Thus, just as PCR was linked to the chemical model, a successful morphological model could be united with both of these, as well.

5.1.4. Discussion. Although to date there have been no results published regarding *in vivo* Raman spectroscopy of atherosclerosis in humans, strong groundwork to pursue such studies has been laid. Indeed, clinical trials are about to begin at several institutions.

No other diagnostic technique currently available holds the potential to diagnose disease on a chemical level to the extent that Raman spectroscopy can. In fact, if the morphological approach discussed above is successful, it could overcome some of the limitations of current histological methodology. For example, Raman spectroscopy samples the entire volume of arterial wall in one go, whereas only representative section of tissues are routinely examined histologically, a process which is prone to sampling error (Römer *et al* 1998a). In addition, processing and sectioning of diseased arterial tissue for histological examination requires initial decalcification, followed by dehydration using lipolytic organic solvents, resulting in loss of chemical information critical to understanding the disease process (Salenius *et al* 1998).

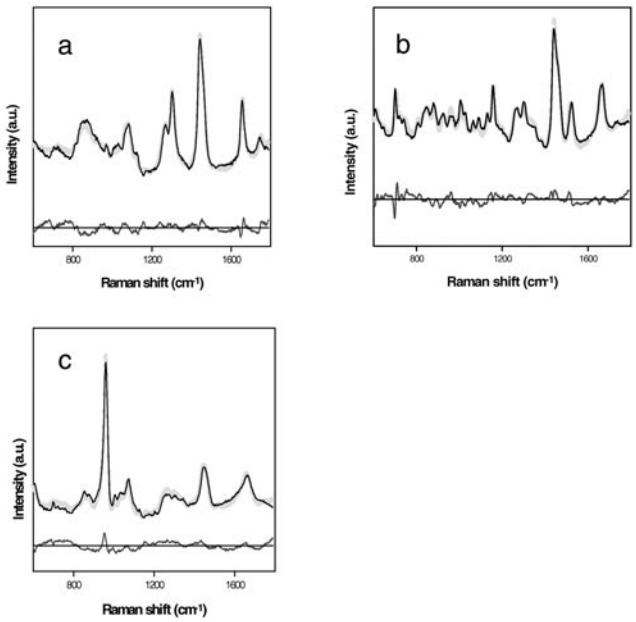


Figure 12. Morphological model fits (dots) and spectral data (curve) for coronary arteries exhibiting various stages of atherosclerosis: (*a*) non-atherosclerotic tissue, (*b*) non-calcified atheromatous plaque and (*c*) calcified plaque. The residuals are plotted below the fits (unpublished data).

Because *in vivo* Raman spectroscopy does not require the removal of tissue, its success might open several new avenues of research. First, it would allow for the evaluation of various non-interventional medical therapies such as pharmacological, diet or exercise regimens. The effectiveness of these therapies, not only in reversing luminal stresses but also in modifying the chemical or morphological composition of the plaque, could be studied by monitoring critical lesions over the course of treatment. In this manner it would be possible to determine the stability of the plaque, and then monitor its progression or regression upon treatment. This type of information could be used to decide which types of treatments are most successful for a given disease state. For instance, there is evidence that certain chemical pathways that contribute to plaque rupture could be blocked pharmacologically (Mach *et al* 1997). It would then be possible to screen a patient prior to prescription, use the spectroscopic information to guide the choice of therapy, and reduce the incidence of acute myocardial infarction and coronary artery thrombosis.

Second, *in vivo* Raman spectroscopy would allow for evaluating risk factors for the restenosis that occurs following percutaneous intervention. It has been found that restenosis following balloon and laser angioplasty is due to smooth muscle cell proliferation, but further study is necessary to determine if this is true for other types of surgical intervention, such as atherectomy and stenting (Manoharan *et al* 1996). The chemical composition of a plaque and the presence of macrophages have also been linked to the probability of restenosis (Moreno *et al* 1996, Salenius *et al* 1998). At least some of these variables can be quantitated by Raman analysis, which would allow for better selection of interventional or non-interventional therapy.

In conclusion, Raman spectroscopy holds the promise of performing accurate, quantitative spectral analysis, which could allow detailed chemical and morphological study of the

atherosclerotic process *in vivo*. Subsequently, it could allow screening of patients to provide the most effective medical therapy or intervention, and evaluation of their progress as the treatment proceeds.

5.2. Breast cancer

Breast cancer is the most common form of malignant tumour found among women in the western world. The National Cancer Institute (1997) reports that 184 000 new cases are diagnosed each year. Early detection can reduce mortality by 40–50% (Marchant 1994). Currently, breast cancer screening is performed by a combination of annual clinical breast examinations, in which doctors search for palpable lesions, and x-ray mammography, in which suspicious focal density changes in the breast tissue can be detected. In cases where the tissue is particularly dense throughout, ultrasound may also be used to help locate suspicious regions. If a lesion is found during the examination, tissue biopsy follows. Biopsy ranges from the fine needle aspiration of single cells to the surgical removal of the entire suspicious mass by excisional biopsy. Seventy to ninety per cent of the suspicious lesions detected using mammography are determined to be benign upon biopsy. Conversely, mammography fails to detect 20% of all malignant lesions. The biopsy procedure and wait for diagnosis is a traumatic experience, both physically and emotionally, even for those women whose outcomes are negative. Furthermore, a recent study found that over a 10 year period, during which subjects received a median of four mammograms and five clinical breast examinations, 31.7% of all women experienced at least one false positive from either test (Elmore *et al* 1998). Consequently, each year a large number of unnecessary breast biopsies are performed. The complete diagnostic procedure, from start to finish, may take several months and include multiple biopsies.

The patient trauma, time delay and high medical costs associated with biopsy could be reduced if a minimally invasive optical spectroscopy technique for real-time diagnosis of suspicious lesions were available. Such a technique would allow physicians to diagnose a lesion much more accurately and immediately upon detection during the screening examination using an optical fibre probe incorporated into a needle device. This Raman needle probe would be inserted into the breast where a suspicious lesion was found, a spectrum would be acquired, and, following real-time analysis, a diagnosis would be given. The probe could be inserted multiple times to examine different areas of the breast with minimal damage to the tissue, unlike current needle biopsy techniques, which must remove cells or tissue for analysis and whose use must be limited. The method of initial detection could be mammography or any of the new imaging techniques that are currently being developed, such as photon migration (Manoharan *et al* 1998).

5.2.1. Anatomy, pathology and biochemistry. The anatomy of the breast is deceptively simple. The mature breast is a large secretory gland composed of 15 to 25 lobes. Each lobe is autonomous and empties into its own excretory duct, which connects to the nipple (figure 13). The lobes themselves are divided into smaller units, called lobules, which in turn are made up of alveoli. The lobular units and the ducts that connect them are separated from the surrounding connective tissue (called stroma) by a basement membrane. Lobular and ductal elements consist of single layers of epithelial and myoepithelial cells. Despite this apparent simplicity, the breast undergoes many changes throughout a woman's life, both progressive (due to puberty, pregnancy and menopause) and cyclical (due to menstruation). These changes, regulated by hormones, are noticeable at a microscopic level in the epithelium, the myoepithelium and the stroma (Sternberg 1997). Because of these constant changes, there

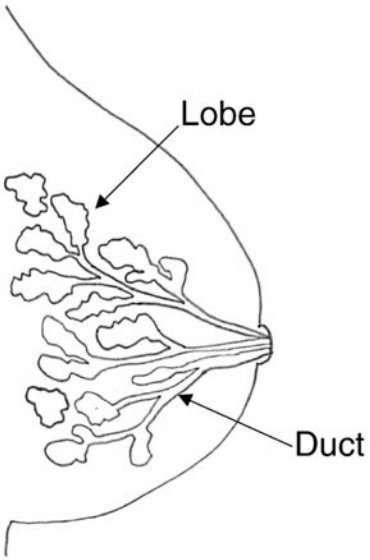


Figure 13. Anatomy of the breast, schematic diagram.

are many opportunities for disease. Although breast pathology is extremely diverse, breast disease is usually divided into two main categories: benign and malignant. Most benign breast lesions are part of a spectrum of fibrocystic changes, whereas malignant lesions are primarily adenocarcinomas, either ductal or lobular in origin.

Malignant transformations in tissue are associated with complex biochemical changes that should provide a basis for spectroscopic detection. For example, increased cell proliferation and metabolic activity result in changes in the concentrations and oxidation states of species such as NADH, $FADH_2$ and ATP, and generally decreased amounts of storage lipids. Differences in phosphocholine, phosphocreatine, NADH, and ATP concentrations in malignant breast cancer cell lines compared with normal breast epithelium cells have been observed using NMR spectroscopy (Singer *et al* 1995). Microcalcifications, a type of mammographically detected lesion, are markers of cellular degradation and necrosis and are chemically composed of calcium oxalate and calcium hydroxyapatite mineral deposits (Radi 1989). Even alterations in the extracellular matrix (stromal desmoplasia) in breast cancer are connected with increased levels of structural proteins such as collagen and elastin (Yeo *et al* 1991). Morphological cellular changes associated with malignancy include loss of differentiation, nuclear enlargement, hyperchromatism (darker staining of cell nucleus), pleomorphism (variation in size and shape of cells and cell nuclei), and atypical mitoses. These changes are reflected in increases of nucleoproteins and nucleic acids. Finally, as malignant epithelial cells break through the basement membrane into the surrounding stroma, proteoglycans (such as chondroitin sulfate) are released to facilitate tumour cell invasion of the stroma (Alini and Losa 1991). Because Raman spectroscopy can provide detailed information about the chemical composition of tissue, several groups have recently begun to explore its potential to diagnose breast cancer.

5.2.2. Background. The purpose of this section is to provide a brief overview of recent research exploring the use of Raman spectroscopy to diagnose breast cancer (see also section 2). Alfano *et al* (1991) were the first to record FT-Raman spectra of human breast tissue, using

1064 nm excitation. They looked for differences in the characteristic peaks of 14 breast tissue specimens (three normal, four benign and seven malignant). They observed that while normal tissue showed Raman bands at 1078, 1300, 1445 and 1651 cm^{-1}, malignant tissue only showed bands at 1445 and 1651 cm^{-1}. Benign tumours had characteristic bands at 1240, 1445 and 1659 cm^{-1}. They found that the relative intensities of the 1445 and 1651 cm^{-1} bands could be correlated with disease classification, although they were unable to associate these spectral differences with biochemical changes in the tissue.

Since then, several other groups have studied Raman spectra of breast tissue using various excitation wavelengths. Redd *et al* (1993) examined the Raman spectra of normal breast tissue using 406.7, 457.9 and 514.5 nm excitation, finding differences in peak intensities between normal and malignant tissue which they attributed to contributions by fatty acids and β-carotene. Wavelength dependence studies (406–830 nm) to determine the best excitation wavelength for observing Raman bands and reducing the fluorescence background of normal breast tissue were published by Frank *et al* (1994). They found that Raman peaks associated with lipids were strongest with 782–830 nm excitation, while carotenoid features were observed with 488–515 nm excitation due to resonance enhancement.

In subsequent studies, Frank *et al* (1995) used 784 nm excitation to obtain Raman spectra of normal tissue, benign tissue (fibrocystic disease), and malignant tissue (infiltrating ductal carcinoma). They noted a shift in the 1439 cm^{-1} band in normal tissue to 1450 cm^{-1} in infiltrating ductal carcinoma. They attributed this change to increased protein concentrations in malignant samples. Using the area ratio of 1654/1439 cm^{-1} bands, they were able to differentiate easily between infiltrating ductal carcinoma and normal tissue. However, they were unable to demonstrate statistically significant differences between infiltrating ductal carcinoma and fibrocystic change.

5.2.3. Materials and methods. Raman spectra were collected using the Ti:sapphire based laboratory system described in section 3. The excitation wavelength was 830 nm. Each tissue site was irradiated with 80 mW of laser power (1 mm diameter spot size) for 100 s. The tissue samples were obtained from excisional biopsy and mastectomy specimens, frozen in liquid nitrogen and stored until use. After data acquisition, the site where the laser was focused was marked with India ink and the tissue was fixed in formalin. The tissue was then diagnosed by a pathologist.

Prior to multivariate analysis, the spectra themselves were processed. A white light source was used to correct for the spectral response of the system. Both the known Raman spectrum of indene and a neon calibration lamp were used for wavelength calibration. The slowly varying fluorescence background was removed by fitting a fourth-order polynomial to each spectrum.

5.2.4. Results and discussion. Typical near-IR Raman spectra of normal, benign, and malignant breast tissue, acquired with our laboratory system, are shown in figure 14. Consistent with the earlier studies cited above, we found that normal tissue spectra were dominated by Raman bands characteristic of fatty acids (1657, 1442 and 1300 cm^{-1}), while the most prominent bands in the spectra of benign and malignant tissue are related to structural protein modes (1667, 1452, 1260, 890 and 820 cm^{-1}). The work published by both Alfano and Frank suggests that one should be able to build an elementary decision algorithm based on the shift in the 1442 cm^{-1} peak (due to the changes in chemical environment of the CH_2 bending mode) and the ratio of the 1445 cm^{-1} and 1650 cm^{-1} bands. Unfortunately, these spectral features are insufficient to separate all three disease categories (figure 15). This figure uses data collected in our laboratory from 61 specimens taken from 13 patients (15 normal, 15

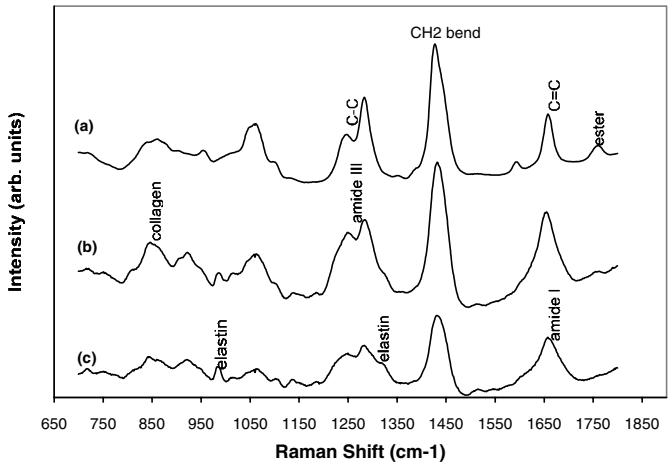


Figure 14. Raman spectra of (*a*) normal, (*b*) benign and (*c*) malignant breast tissue. (From Manoharan *et al* (1998) with permission, American Society for Photobiology.)

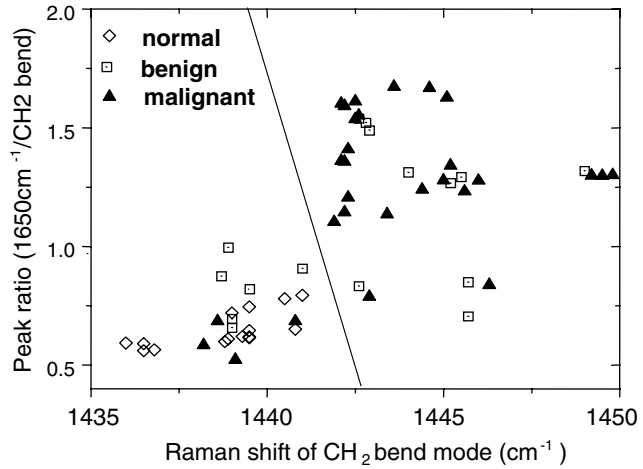


Figure 15. Decision plot for classifying normal, benign and malignant breast tissue based on spectral features suggested by the Alfano *et al* and Franck *et al*. The line is drawn to emphasize the separation between normal and malignant samples. (From Manoharan *et al* (1998) with permission, American Society for Photobiology.)

benign and 31 malignant) (Manoharan *et al* 1998). The decision line was drawn to maximize the separation between the categories. As can be seen, these spectral features separate normal and malignant samples with high sensitivity and specificity. However, the benign samples are not well separated.

Differentiation of benign and malignant lesions is, of course, crucial for clinical purposes. Both benign and malignant breast lesions have a high protein content due to stromal changes, making them difficult to differentiate using these specific spectral features. Normal breast tissue is easy to distinguish because of the strong Raman signals due to the abundance of fat.

We have used principal component analysis to examine the degree to which consistent spectral differences exist between benign and malignant tumours. Instead of building an

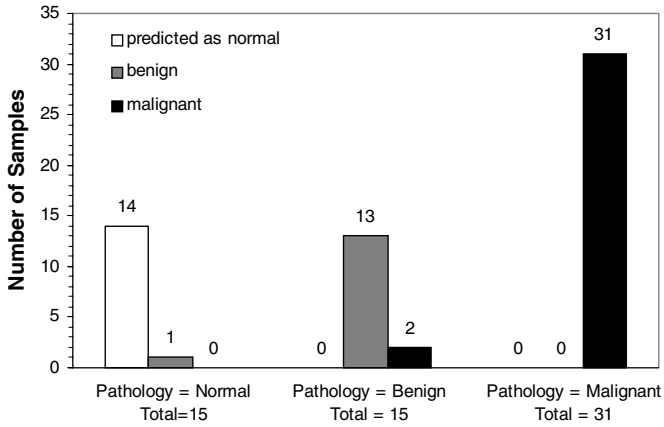


Figure 16. Comparison of predicted and histological classifications of breast tissue using principal component analysis and logistic regression. (From Manoharan *et al* (1998) with permission, American Society for Photobiology.)

algorithm based on a few spectral features, principal component analysis enables us to use the full range of data to develop a decision algorithm using logistic regression (section 4). In a preliminary study, we used principal component analysis in conjunction with logistic regression to analyse the same data presented above (Manoharan *et al* 1998). Four of the twelve principal components were determined to be statistically significant. Using the fit coefficients of these principal components, we were able to estimate the probability that each sample belonged to a given diagnostic category (normal, benign or malignant). The resulting predicted classification versus the pathological diagnosis is shown in figure 16. Fourteen of 15 normal, 13 of 15 benign, and 31 of 31 malignant samples were correctly diagnosed. We note that only two of the fifteen benign samples were misclassified, with no false negatives. This is particularly promising when considering the potential of utilizing the technique as a clinical tool for reducing the number of unnecessary excisional breast biopsies.

These studies demonstrate that Raman spectra contain the necessary information for differentiating normal, benign and malignant breast tissue. However, researchers have yet to interpret this spectral information at a detailed biochemical level. In order to do this, our laboratory is undertaking biochemical and morphological modelling of breast tissue Raman spectra.

Currently, we are collecting Raman microimages of breast tissue. Our Raman confocal microscope is able to take Raman spectra every few micrometres, creating a spectral map of a section of breast tissue (6 μm thick). The ability to do spectral mapping will allow us to understand the biochemistry and morphology which gives rise to the observed spectra. For example, by correlating spectral maps with stained sections of the same tissue we can identify the morphological origins of the spectral features that we observe. Alternatively, by comparing the spectra of particular morphological structures, such as ductal epithelial cells, found in malignant tissue with those found in normal or benign tissue, we can also learn about how these morphological structures change as a result of the disease process. Finally, by collecting Raman microimages of a variety of tissue types, we can build a library of spectra, morphological in origin, which can be used to generate basis spectra for modelling (see section 4). Our model can then be used for diagnostic purposes, using techniques which are similar to those employed for arteries (see section 5.1).

The future of Raman diagnosis of breast cancer depends on our ability to understand the origins of the spectral features that principal component analysis indicates exist. Modelling, whether it be chemical or morphological, is the key to uncovering the answer. Once we are confident we understand the chemical and biological origins of our spectral features, we can begin to develop an algorithm for diagnosing breast cancer *in vivo*. This will require the development of suitable optical fibre probes, compatible with existing needle devices, as discussed in section 3.

5.3. Alzheimer's disease

Alzheimer's disease (AD) is the most common form of dementia. The impact of Alzheimer's disease is expanding with the expanding percentage of elderly people in our population. It is estimated that unless a cure or prevention is found, 14 million Americans will be afflicted by 2050. According to the Alzheimer's Disease and Related Disorders Association, AD costs patients, their families and society more than $100 billion per year in health care and related costs.

At present, a definitive diagnosis of AD is only possible by histological examination of brain tissue under the microscope, a procedure conducted almost exclusively at autopsy. Lack of an accessible, non-invasive technique for diagnosing AD is the greatest impediment in the search for its treatment and prevention. Our results show that Raman spectroscopy can differentiate brain tissue of AD patients from that of normal patients, and therefore may provide a basis for a new diagnostic technique.

Neurons, with their network of interconnecting cell processes (axons and dendrites) are the communication system of the brain. The grey matter is composed of neuronal cell bodies and glia, the specialized 'connective tissue' of the brain. Neurons range widely in size and shape in different areas of the brain. The archtypal neuron is the pyramidal neuron of the temporal cerebral cortex, an area frequently affected by AD. These neurons may measure up to 5 μm in width at the base and 10–25 μm in length. The white matter is composed primarily of axonal processes.

Alzheimer's disease is a terminal form of dementia resulting from progressive degeneration of the neurons. Its cause is unknown. The most striking feature of the gross anatomy of the disease is the overall atrophy of the brain's grey matter. Microscopically, the known neuropathological hallmarks of AD are senile plaques and neurofibrillary tangles (figure 17). Both are composed of abnormally aggregated proteins. Senile plaques are extracellular deposits of amyloid beta protein, aggregated into fibrils and surrounded by abnormal neurons and glia. Neurofibrillary tangles are intracellular aggregates formed by paired helical filaments of highly phosphorylated tau protein, a protein that normally regulates structural stability in axons. Senile plaques typically occupy a roughly spherical volume of brain tissue about 20–50 μm in diameter while neurofibrillary tangles vary in dimension depending on the type of neuron involved. In addition to the neurofibrillary tangles and senile plaques, the AD afflicted brain contains a dense array of distorted neurites, known as neuropil threads or dystrophic neurites, which are found throughout the interwoven array of nerve endings known as the neuropil.

Although cerebral amyloid beta protein accumulation by itself in the form of senile plaques is insufficient for the development of dementia, it may induce neuronal degeneration in the form of neurofibrillary tangles and neuropil threads. The abundance of these structures correlates with the severity of dementia in AD (McKee *et al* 1991, Kowall and McKee 1993). These neuropathological markers can provide detectable spectroscopic signatures.

Vibrational spectroscopy techniques, such as FT-IR, have been applied to the classification of AD brain tissue (Choo *et al* 1995), structural characterization of synthetic amyloid beta protein fragments (Fabian *et al* 1993), tau protein and AD-paired helical fragments (Schweers

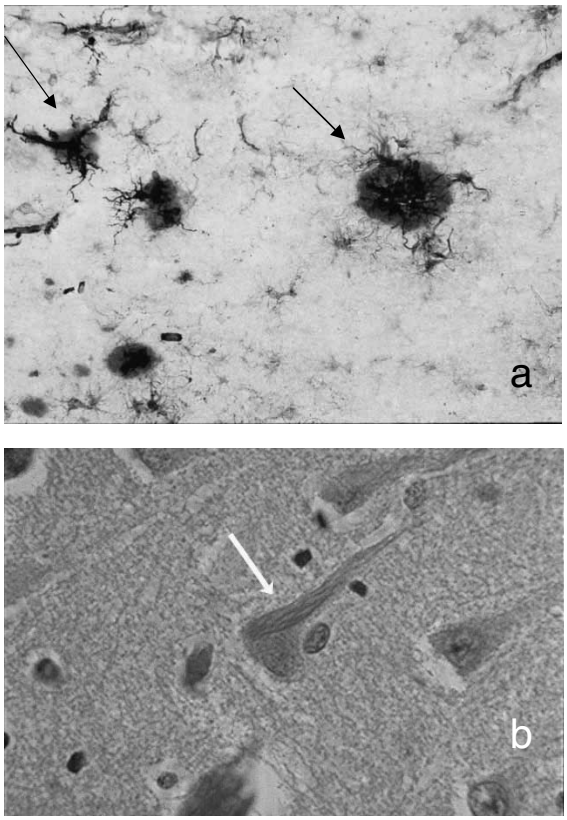


Figure 17. Photomicrographs showing the pathologic hallmarks of AD in brain tissue. The arrows indicate (*a*) senile plaques (with permission of Dr Haruhiko Akiyama, Division of Neuropathology, Tokyo Institute of Psychiatry) and (*b*) neurofibrillary tangles (with permission of Dr Edward Klatt, Department of Pathology, School of Medicine, University of Utah). The senile plaques shown in (*a*) are 20–50 μm in diameter and the neurofibrillary tangle shown in (*b*) is about 3 μm wide and 20 μm in length.

et al 1994) and microspectroscopy of endocrine amyloid (O'Leary and Levin 1985). One conclusion from these studies was that infrared spectroscopy could potentially be used in the pathological diagnosis of AD from autopsy tissue (Choo *et al* 1995; Fabian *et al* 1993). However, the penetration depth of infrared light in tissue is quite small, 10 μm or less. Therefore, infrared absorption spectroscopy cannot be used for non-invasive diagnosis of AD *in vivo*. Raman spectroscopy provides information about molecular vibrations that is similar to that provided by infrared spectroscopy. In addition, the near-IR light used for Raman excitation can penetrate quite deeply into tissue, of the order of centimetres. Thus, near-IR Raman spectroscopy is a potentially attractive technique for probing and diagnosing AD.

Near-IR FT-Raman spectra of control and tumour brain tissue have been reported (Mizuno *et al* 1994). We have used near-IR Raman spectroscopy to examine the differences between normal brain tissue and AD brain tissue. The vibrational frequencies and relative intensities of the spectral features for white and grey matter of normal brain tissue agree with those obtained using FT-Raman spectroscopy (Mizuno *et al* 1994). Raman spectra of AD brain tissue show distinct differences from normal tissue spectra that can be used to distinguish AD from normal brain.

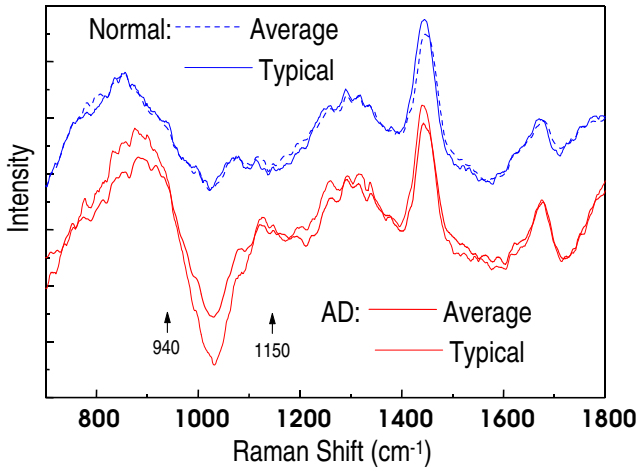


Figure 18. Average NIR Raman spectra of AD and normal temporal cortex grey matter. Typical spectra of AD and normal grey matter are also shown.

5.3.1. Materials and methods. Post mortem tissue specimens of human temporal cortex and white matter were harvested at autopsy, flash frozen on dry ice, and stored at $-80\,^{\circ}$C prior to spectroscopy. Selected specimens were matched with respect to age, autopsy interval, race and gender, and classified as AD or normal control based on histological examination. All neuropathological specimens diagnosed as AD met the standard diagnostic criteria for definite AD established by the Consortium to Establish a Registry for Alzheimer's Disease (Mirra *et al* 1994, Hyman and Trojanowski 1997). The control brain specimens were histologically normal, without neurofibrillary changes. Immediately prior to acquiring spectra, all specimens were brought to room temperature. The thawed, unfixed specimens were used without further treatment.

The near-IR Raman scattering spectra for white and grey matter specimens of AD brain and normal brain were measured in the vibrational spectral region from 600 to 1800 cm^{-1}. The specimens were excited at 830 nm using the tissue Raman laboratory system described in section 3 (figure 4). Each tissue sample was irradiated with 50 mW of laser power (100 μm spot size). Spectra were collected with a data acquisition time of 80 seconds. Differences between the Raman spectra of normal and diseased white matter are evident, but the most striking distinctions are seen upon comparing spectra from AD and normal grey matter specimens.

5.3.2. Results and discussion. Figure 18 shows Raman spectra of temporal cortex grey matter from AD (17 samples from four brains) and normal (9 samples from one brain) brains. Tissue background fluorescence has been removed, and the spectra have been intensity calibrated with a standard tungsten lamp. The average spectra for the nine control specimens and the 17 AD specimens are compared with a single representative spectrum of each tissue type to indicate reproducibility. The strongest Raman scattering arises from well-characterized bands involving protein and lipid vibrational transitions between 1200 and 1700 cm^{-1}. Based on isolated protein investigations, the structure of the band at about 1670 cm^{-1} (amide I) probably corresponds to the aggregation state of amyloid β-protein observed in infrared absorption studies of brain tissue (Choo *et al* 1995). The low-frequency shoulder on the 1670 cm^{-1} band is clearly evident in the AD tissue spectrum. This structure is typical of the β-pleated sheet conformation observed for amyloid β-protein in senile plaques. The intensity ratio of the

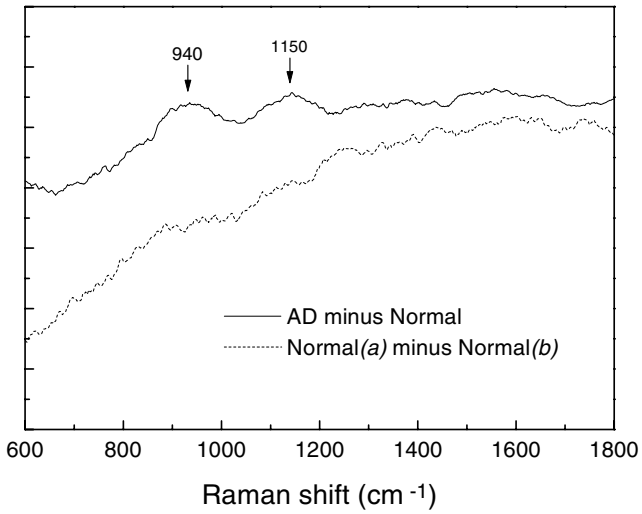


Figure 19. Raman difference spectra. Upper curve, AD minus normal; lower curve, difference of two different normal controls.

bands at 1450 and 1670 cm^{-1} characterizes the relative lipid-to-protein content of the tissue and has been used to monitor disease-related changes in tissue composition. These bands are somewhat similar in structure, relative intensity and frequency for both the normal and AD tissue spectra. This similarity is due to an overlap of the contributions to these specific bands from all protein and lipid scatterers, both disease-related and normal. However, the complex band structure in the AD grey matter spectrum between 800 and 1000 cm^{-1} and between 1000 and 1200 cm^{-1} in figure 18 is absent from the normal grey matter spectrum. This distinction becomes evident upon examining the difference spectra of figure 19.

Figure 19 shows the Raman difference spectra obtained from normalized, intensity-calibrated data that include background fluorescence. The upper curve is obtained by subtracting the spectrum of a normal grey matter specimen from that of an AD grey matter specimen. The lower curve is obtained by subtracting the same normal grey matter spectrum from a different normal grey matter specimen. The overall slope and curvature of each difference spectrum is due to differences in the tissue fluorescence background, and is not significant. The similarity of the diseased and control spectra between 1200 to 1700 cm^{-1} is indicated by the absence of features in this region of the AD (AD minus normal) difference spectrum of figure 19 (upper curve). However, in the region 800 to 1200 cm^{-1}, the AD difference spectrum shows two distinct features centred at about 940 cm^{-1} and 1150 cm^{-1}. The corresponding region in the control difference spectrum (normal (*a*) minus normal (*b*), lower curve) is featureless. We have not yet assigned the Raman transitions at 940 cm^{-1} and 1150 cm^{-1}, but the difference spectra indicate that they are absent from the normal control tissue spectra. Raman activity is known to occur for phospholipid structural changes (*trans* versus *gauche* isomerism) at about 1130 cm^{-1}. This band becomes more intense as *trans* structure increases, with a concomitant loss in fluidity or flexibility for these molecules. It is tempting to associate these Raman features with the loss of fluidity of amyloid deposits as they undergo diffuse to fibrillated structural changes in the formation of senile plaques. Sharp features in FT-Raman spectra near 940 and 1150 cm^{-1} show the presence of carotenoids in other brain diseases (Mizuno *et al* 1994).

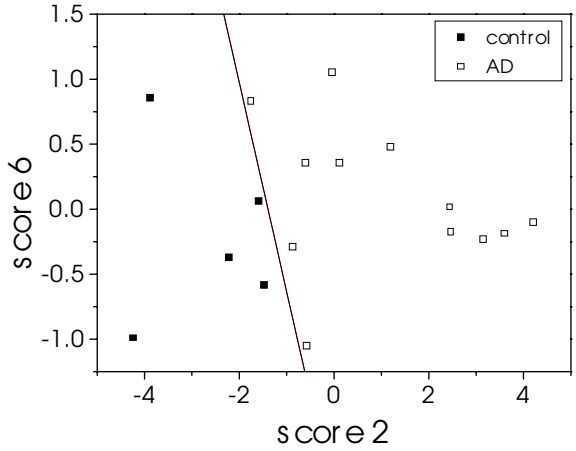


Figure 20. Binary decision result of principal component analysis of Raman spectra from AD (□) and normal (■) temporal cortex specimens.

A subset of the above spectral data was studied using principal component analysis (see section 4). The Raman spectra were pre-processed by normalization to unit integrated intensity, followed by mean centring. Principal components were generated for Raman spectra of 5 normal and 12 AD brain specimens. Eight principal components had features above the noise level. Logistic regression indicated that principal component scores 2 and 6 together could be used to correctly classify all specimens as AD or normal at a statistical significance of $p < 0.0002$. By plotting score 2 versus score 6 for each Raman spectrum, and graphing the line of equal probability that a specimen is spectroscopically diagnosed as AD or normal, brain tissue specimens could be spectroscopically diagnosed as normal or AD with no misclassifications for this limited data set (figure 20). Note that most of the diagnostic information is carried in principal component 2. This result was validated using the leave-one-out method of internal cross validation (section 4).

Can the encouraging results of these initial *in vitro* studies be realized in a clinical setting without brain tissue removal? To be of practical value for clinical implementation, Raman spectral diagnosis of AD would require non-invasive or minimally invasive optical access to the brain. In fact, non-invasive optical access to a region of the brain that exhibits the hallmark pathological features of Alzheimer's disease may be afforded by the olfactory cavity. An optical fibre probe inserted into the nostril and positioned at the olfactory epithelium, located approximately 7 cm inside the nasal passageway (figure 21), can provide near-IR laser illumination of the olfactory bulb and the base of the frontal lobe, with minimal interference from intervening tissue. The cortical tissue of these brain regions exhibit pathology hallmarks of AD (Kovacs *et al* 1996, 1998). Furthermore, the thin bone of the cribriform plate is a perforated structure, allowing penetration of the septal olfactory nerves into the inner chamber of the nose, thus providing an unobstructed optical path to the brain. In addition, biopsy specimens of the olfactory mucosa, the mucus-covered membrane of olfactory epithelium, have shown evidence of Alzheimer's disease (Yamagishi *et al* 1994). Thus, an optical fibre contact probe can directly access these structures and may offer easier access for implementing Raman spectroscopy for AD diagnosis in a clinical setting.

In addition, as technological developments extend Raman measurements further into the short-wave IR region (1000–2000 nm), transcranial measurements may become feasible. For example, with excitation at 1064 nm, the Raman fingerprint region extends to about 1300 nm.

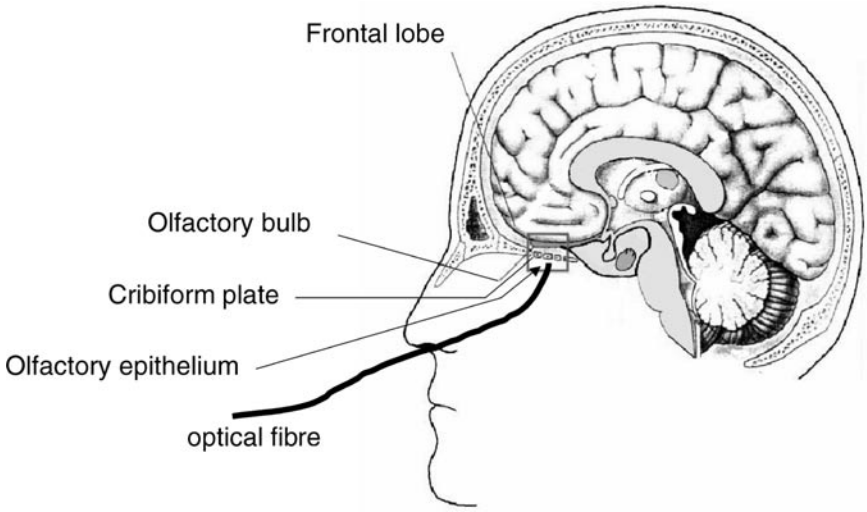


Figure 21. Cross-sectional diagram of the head showing the olfactory tract (with permission of Nicole Cooper, University of Colorado). Optical access to the brain for Alzheimer's disease diagnosis may be afforded by an optical fibre inserted into the nostril, guided through the nasal cavity, and positioned at the olfactory epithelium.

Although water absorption has increased by about a factor of 2.5 at this wavelength compared with 975 nm (the extent of the fingerprint region with 830 nm excitation), tissue fluorescence is largely eliminated with 1064 nm excitation. Since brain tissue fluorescence with 830 nm excitation is one to two orders of magnitude greater than the Raman signal, removing its associated shot noise with 1064 nm excitation could still improve the brain tissue Raman S/R ratio by an order of magnitude, even considering the increased absorption at these longer wavelengths due to water. The feasibility of this approach has been demonstrated in artery (section 3.4.1) and warrants further consideration in future developments of optical spectroscopic diagnosis of AD.

5.4. Raman analysis of blood analytes

Determination of blood analyte concentrations is important for diagnosis and clinical management of many medical conditions. For example, blood glucose is a critical indicator of diabetes. Increased blood serum concentrations of urea and creatinine may indicate renal dysfunction. Cholesterol and triglycerides are related to cardiovascular disease risk. Total protein and albumin concentrations can be affected by many factors, including inflammatory disease and nutrition. Conventional clinical chemistry techniques for determining blood analyte concentrations include electrophoresis, high-performance liquid chromatography, immunochemical techniques, photometry, absorption spectroscopy and flame spectroscopy (Henry *et al* 1974). These techniques share several disadvantages, the first of which is the requirement to withdraw blood for *in vitro* analysis. Blood withdrawal can be painful and apt to generate errors. For example, diabetics find performing the daily, multiple, finger-prick blood samplings required to determine blood glucose levels the recommended three to five times a day not only painful, but also a deterrent to self-management of their condition (Waynant and Chenault 1998). In addition, the physical and chemical techniques required to prepare blood samples for conventional clinical analyses are complex and time-consuming.

Blood withdrawal also carries the risk of infection and transmission of disease by blood-borne pathogens such as HIV and hepatitis. Continuous monitoring of blood analyte levels, such as in an intensive care setting or an ongoing medical emergency, compound these problems by requiring repetitive blood withdrawal, with little time to wait for conventional analyses.

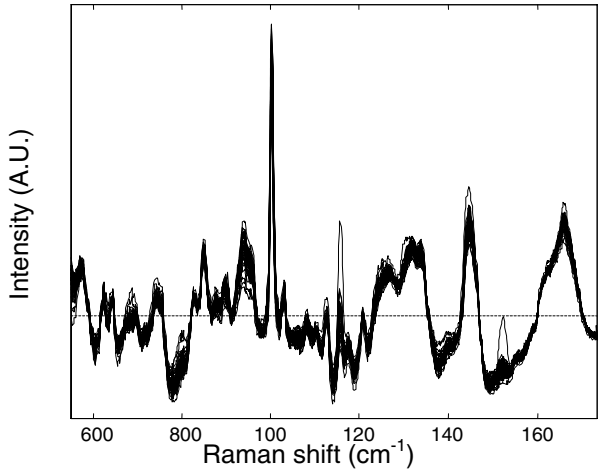
The potential medical benefits and prospective market if these problems could be solved have driven study of many alternative blood analysis techniques based on different approaches. Marbach *et al* (1993) measured blood glucose by diffuse reflectance spectroscopy. Dou *et al* (1996) quantitatively analysed glucose in blood plasma and serum by anti-Stokes Raman spectroscopy. Koo *et al* (1998) measured glucose in blood serum by Stokes Raman spectroscopy. Coté *et al* (1992) showed the possibility of predicting glucose concentration by measuring the polarization rotation of the light transmitted through the sample. Tarr and Steffes (1993) devised a stimulated Raman system to measure the glucose concentration in the aqueous humour to avoid absorption and scattering problems in whole blood.

These optical methods measure the photon interaction with blood without chemical modification, thus offering the possibility of reagentless blood analysis. Furthermore, using wavelengths of light at which absorption by haemoglobin, melanin, and water is insignificant allows deeper penetration of photons (Itzkan and Izatt 1994). Operating in this 'diagnostic window', from about 700–1300 nm, may enable transcutaneous measurement of blood analytes, thereby obviating blood withdrawal.

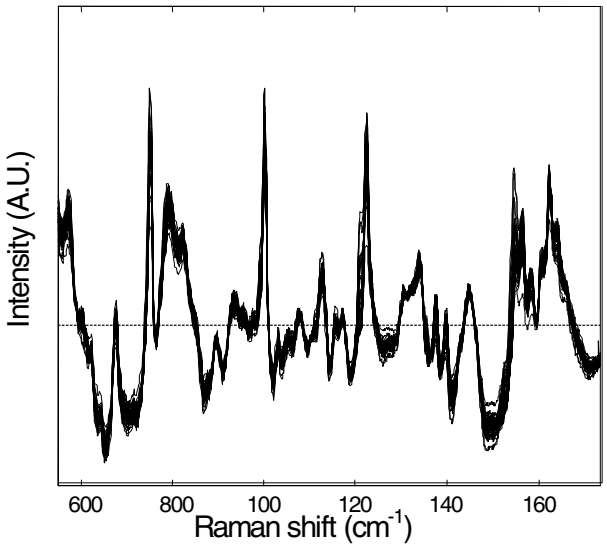
Raman spectroscopy is particularly suitable for implementing the advantages of optical blood analysis techniques. The relatively well-resolved fundamental vibrational transitions of the Raman spectrum, unique for each chemical species in a given environment (Campbell and Durek 1984), provide more detailed spectral information than the broad overtone bands of NIR absorption spectra. This is particularly important for complex biological systems such as blood. Despite the relatively small cross sections for Raman scattering, instrumentation advances have promoted rapidly increasing interest in, and have begun to reveal the benefits and capabilities of, Raman spectroscopy for blood analysis.

5.4.1. Methods. Recently Berger *et al* (1999) conducted experiments that demonstrate the ability of Raman spectroscopy to measure blood analytes in both blood serum and whole blood. Sixty-nine blood samples were drawn from different patients and divided into three samples: EDTA-anticoagulated whole blood, blood plasma and blood serum. Chemical analyses were performed on the blood plasma and serum samples, using a standard reference clinical laboratory assay. The blood samples were then stored at 4 °C until the Raman spectroscopy assays were performed. For each blood sample, 30 successive Raman spectra were taken with a 10 s integration time per spectrum. This allowed us to vary the total integration time of the analysed signal from 10 s (one spectrum) to 5 min (sum of 30 spectra). Thus, it was possible to study improvement of prediction accuracy as a function of signal integration time. Details of the instrumentation are presented in section 3.2.

5.4.2. Results. Raman spectra of the 69 whole blood and blood serum samples are shown in figure 22. The background due to protein fluorescence was removed by polynomial fitting and subtraction. Spectral changes from sample to sample appear to be small. In particular, analytes with low physiological concentrations make no obvious contributions to the observed spectra. To illustrate this, figure 23 compares the Raman spectrum of a typical blood serum specimen with that of glucose in saline solution at a typical physiological concentration of 100 mg dl^{-1}. The background slope is fluorescence, and the small peaks are Raman signals from blood proteins. Raman glucose peaks cannot be seen in either spectrum, because of its



(*a*)



(*b*)

Figure 22. Raman spectra of (*a*) blood serum and (*b*) whole blood samples from 69 patients after linear background subtraction. Spectral changes from sample to sample appear to be insignificant.

weak signal intensity and the low concentration at physiological levels. Clearly, the weak contribution of the Raman spectrum of glucose at physiological levels to the observed blood serum spectrum would be impossible to analyse using traditional empirical techniques, so a more sophisticated data analysis technique is required.

To quantify the contributions of various blood analytes to the observed Raman spectra of blood serum and whole blood, Berger *et al* (1999) used PLS, which is described in section 4. To determine the concentrations of the individual analytes in a blood sample, PLS requires a training set of spectra to calibrate a linear model. This calibration set is composed of spectra of samples whose constituent chemical concentrations are known *a priori*. It is used to build

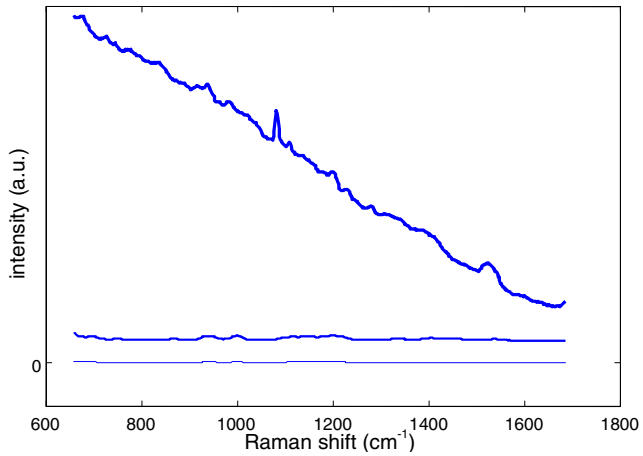


Figure 23. Comparison of Raman spectra of (top curve) blood serum and (bottom) glucosein saline solution at physiological concentration, 100 mg dl^{-1}. Because spectral features of glucose at this concentration are barely visible, Berger (1998) plotted glucose spectrum (middle) at a higher concentration, 1000 mg dl^{-1}, which is beyond the physiological range.

a mathematical model to describe the spectral contribution of each constituent analyte to the sample spectra. This model can then be applied to analyse the spectra of samples for which the constituent analyte concentrations are unknown. The errors in the concentrations derived from this 'calibration' data set can be determined by comparison with the concentrations determined independently by reference clinical laboratory methods. These prediction errors are then compared with the error levels in the calibration data set to determine the robustness of the linear model.

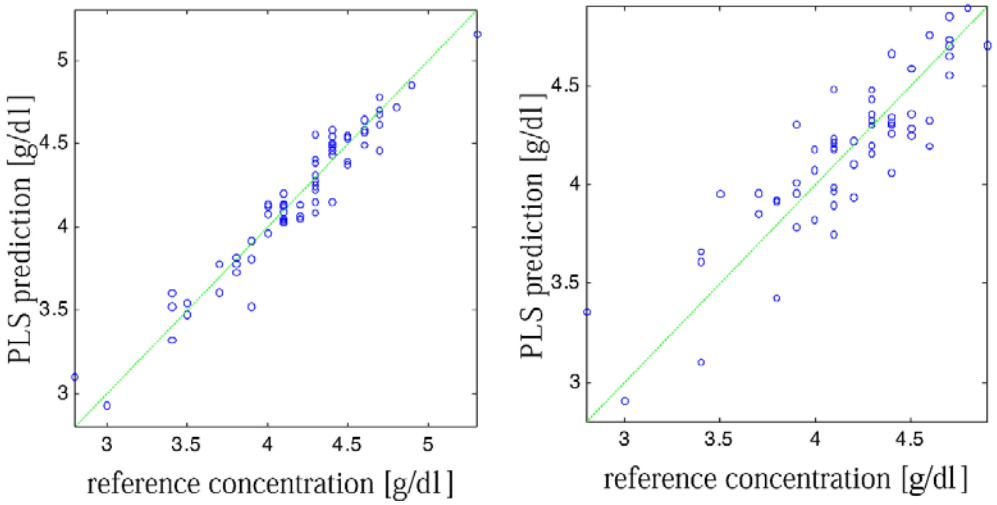
Berger *et al* (1999) applied PLS analysis to the Raman spectral data acquired for the 69 blood specimens in the study, and showed that accurate concentration information could be extracted from the minute spectral changes described above. Several blood analytes were measured in blood serum and whole blood. The blood analyte concentrations predicted by PLS were compared with reference concentrations measured by the standard clinical laboratory technique. Predictions of a typical analyte, albumin in blood serum and whole blood, are shown in figure 24. Prediction errors for analytes measured in serum and in whole blood are listed with their normal reference ranges in table 1. To reduce the prediction errors, Koo *et al* (1999) recently applied a newly developed analysis technique, HLA, which is described in section 4, and showed that the prediction errors were further reduced by about 20% for glucose in blood serum and whole blood.

Achieving the smallest prediction error possible is, of course, desirable, but may not be necessary in all clinical settings. Therefore, determining the medically acceptable level of concentration prediction error is not straightforward. W L Clarke *et al* (1987) suggested the method of error grid analysis to evaluate the clinical accuracy of a glucose measurement system. Instead of simply quoting the prediction error, the Clarke error grid takes into account the potential adverse impact on patient outcome if medical decisions are based on an inaccurate concentration determination. For example, for blood glucose, small errors have more clinical significance near physiological levels than at extremely high or low levels. Glucose predictions obtained using HLA analysis are plotted on the error grid in figure 25. Error grid analysis for evaluating methods of predicting analyte concentrations other than glucose are not currently

Table 1. Prediction errors of analytes in blood serum and whole blood.

Analyte	Serum	Blood	Normal range†
Glucose	21 mg dl^{-1}	N/A	63–110 mg dl^{-1}
Cholesterol	10 mg dl^{-1}	N/A	144–275 mg dl^{-1}
Total Protein	0.20 g dl^{-1}	0.37 g dl^{-1}	6.6–8.3 g dl^{-1}
Albumin	0.12 g dl^{-1}	0.22 g dl^{-1}	3.76–5.49 g dl^{-1}
Triglyceride	26 mg dl^{-1}	80 mg dl^{-1}	30–135 mg dl^{-1}
Urea	3.0 mg dl^{-1}	6.4 mg dl^{-1}	25.3–47.8 mg dl^{-1}
Haematocrit	N/A	1.5%	36–55%

† Normal ranges from Lentner (1985). Normal ranges in serum for all analytes except haematocrit. Normal ranges for haematocrit in whole blood.



(*a*) (*b*)

Figure 24. Raman prediction of albumin concentration in (*a*) blood serum and (*b*) whole blood, versus reference concentration.

available. However, it is helpful to compare the prediction error and the normal reference range to see how useful the prediction would be. A standard deviation of the prediction errors which is less than 5% of the mean normal value is generally acceptable for clinical laboratories.

Table 1 shows that all the analytes are predicted with higher error in whole blood. This is partly due to interference by Raman signals from scatterers in whole blood not present in blood serum, such as blood cells, and partly to the lower S/N ratio attained for the whole blood specimens. Because of the turbidity of whole blood, as compared with clear blood serum, the penetration depth of light is reduced. With our current collection optics, this results in a smaller sampling volume and reduced Raman signal intensity, whereas the noise level (shot limited) remains approximately constant. Computer simulation showed that the S/N ratio plays a significant role in the prediction errors of blood analyte concentrations (Berger and Feld 1997), so it is not surprising to find that the observed prediction errors for whole blood are larger than those observed for blood serum.

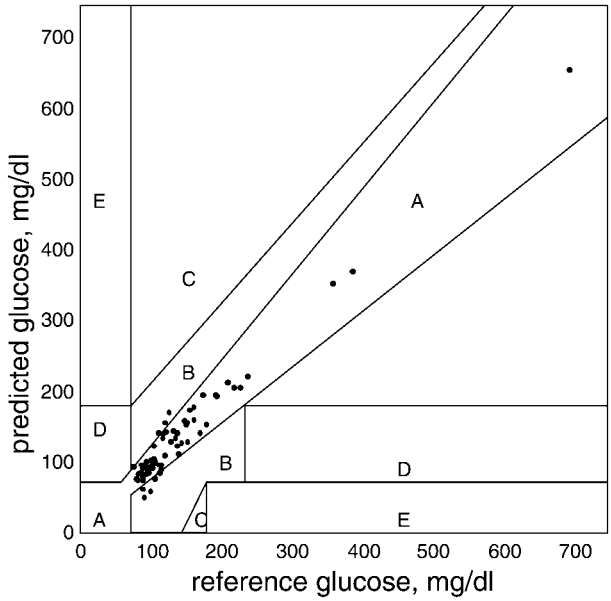


Figure 25. HLA predictions of glucose concentration in blood serum, plotted on a Clarke error grid (Koo *et al* 1999, copyright © 1999 Diabetes Technology and Therapeutics). Zone A represents 'clinically accurate' data; zone B covers 'harmless' errors; predictions in zone D could lead to 'dangerous failure to detect and treat'; those in zone C could result in 'overtreatment' and those in zone E could result in 'undertreatment' (W Clarke *et al* 1987).

5.4.3. Discussion and future directions. If prediction errors can be further reduced, near infrared Raman spectroscopy could be applied to the clinical analysis of whole blood. Of the two error sources mentioned above, interference is more difficult to prevent with current multivariate calibration techniques, unless the calibration data set contains sufficient information on the chemical interferents present in whole blood. The S/N ratio can be improved by increasing the laser power or the signal integration time, although long integration times are not desirable when rapid analysis results are critical and too much laser power may damage blood cells and proteins in the sample. Another way to increase the S/N ratio is to increase the optical throughput of the system. Further investigation of the relationship between the S/N ratio and concentration prediction error will help to determine the optimal integration time and minimum S/N ratio threshold, allowing us to determine the optimal throughput of the system.

For transcutaneous measurement of blood analytes, both elastic and Raman scattering from the skin are potential concerns. Elastic scattering limits the photon penetration depth while increasing the Raman spot size. As a result, the light is concentrated in a volume within the tissue determined by optical diffusion. With current collection optics, this results in reduced signal intensity and a lower S/N ratio. The consequence is an increase in prediction error.

Thus, understanding Raman scattering in complex turbid media such as the blood–skin matrix is critical if non-invasive, transcutaneous blood analysis by Raman spectroscopy is to be achieved. Implementing alternative collection geometries may also necessitate developing appropriate multivariate calibration methods to allow the study of concentration measurements in turbid media.

Another problem is that Raman signals generated in tissue compartments other than blood will interfere with the analysis of the blood analyte concentrations, and multivariate techniques cannot be used to resolve such interference. Since it may be impossible to distinguish

the spectrum of the same chemical from one compartment to another, this problem must be solved optically.

In conclusion, Raman spectroscopy has been successfully employed to accurately measure blood concentrations of glucose, urea, total protein, albumin, cholesterol and triglycerides. However, further research is required to make this technique practical in clinical laboratories, and to develop transcutaneous monitoring schemes.

6. Conclusion: some observations and future directions

At the beginning of the 1990s, when Raman spectroscopy was first being considered as a clinical tool for diagnosing disease in biological tissue, there was a considerable skepticism that useful techniques could be developed. Critics doubted that Raman signals could be observed at levels of laser power that would not produce severe tissue damage. In addition, there was concern that the background tissue fluorescence was so enormous that it would not be possible to discern observable Raman signals. It was also argued that even if Raman spectra could be obtained, the variety of biochemical moieties present in biological tissue was so great that all that would be observed would be a broad background composed of Raman contributions from a multitude of different molecules.

As it turns out, these obstacles were overcome within a relatively short time, and well-defined Raman spectral bands are indeed observable. And so, the clinical use of Raman spectroscopy as a tool for real-time diagnosis of disease and *in situ* evaluation of biological tissue is nearly at hand. This review has summarized the current status of this rapidly developing field. We have reviewed the literature, discussed the recent technology developments, described modern methods of spectral analysis for extracting diagnostic chemical and morphological information, and described several in-depth applications in various stages of development.

We shall begin this concluding section with a summary of what has been achieved so far. We shall then discuss the issues to be addressed in implementing Raman spectroscopy in various clinical applications. Finally, we shall comment on some future directions.

6.1. Accomplishments

What has been achieved so far? In less than a decade, investigators studying a variety of tissues with a broad range of techniques have accumulated a substantial body of *in vitro* data. Taken as a whole, these results clearly demonstrate the correlation between Raman spectra and the biological and chemical hallmarks of disease, and make it clear that Raman spectroscopy is a potentially important clinical tool.

Over this period, the technology has progressed from visible-laser to FT-Raman to CCD-based Raman instruments, and the superiority of near-IR CCD-based Raman instrumentation for biomedical applications has been established. This approach provides high quality Raman spectra with good S/N ratio in clinically acceptable data-acquisition times, using acceptable levels of laser excitation power at a benign wavelength from the point of view of tissue damage. In addition, it permits the ubiquitous background fluorescence of biological tissues to be brought under control. Without these advances, clinical implementation of Raman spectroscopy would not be possible.

In addition, multivariate techniques have been developed for analysing the full Raman spectrum to measure concentrations and classify disease. The examples provided in section 5 illustrate the application of these techniques to the diagnosis of atherosclerosis, breast cancer and Alzheimer's disease, and measurement of blood analytes, and suggest how the technology for such applications can be implemented in clinical settings.

6.2. Clinical implementation of Raman spectroscopy

Although substantial advances in biomedical Raman spectroscopy have been made, a number of additional steps will be required to realize the development of clinical tools based on this new technology. First, most clinical applications of Raman spectroscopy require optical fibre probes. As discussed in section 3, such probes must be free of Raman background, or at least this background must be substantially reduced.

Raman probes will need to be incorporated into catheters, cannulas and needles. Probes of particular design will be required for each specific application, and in each application there will be different problems to address. For example, in the case of atherosclerosis, side-looking contact probes that displace blood will be needed to examine arterial walls. In the case of breast cancer, needle probes of large effective cylindrical sensing range will be important, requiring development of innovative needle configurations. In the case of diagnosing Alzheimer's disease and measuring blood analytes, probes with transcutaneous capabilities will have to be developed. Various issues will have to be addressed, such as developing light collection geometries with improved S/N, eliminating interference from unwanted compartments of tissue and, in the case of blood chemicals, relating the information obtained in the blood–tissue interface to the standard information currently measured in blood itself.

In the near term, further advances in Raman instrumentation can also be expected. Several avenues are worthy of investigation. CCD detector technology continues to improve, and may soon include sensitive detection in the 1000–2000 nm region, as well as increased quantum efficiency and array size, reduced read-out noise, and reduced cost. In addition, high-throughput spectrometers specially designed to couple to optical fibres can be developed. These developments will further improve S/N and, at the same time, reduce excitation laser power and/or signal collection time. Such advances will not only enhance the diagnostic capabilities of Raman spectroscopy, but also may enable monitoring of therapeutic interventions and developing insight to drug behaviour through *in vivo* molecular pharmacology.

6.3. Future directions

The last millimetre of the distal tip of the probe is perhaps the most important part of the entire Raman system, and tip design is an important area for future development. By properly designing the probe–tissue interface, one can obtain calibrated intensity information (Cothren *et al* 1986) and depth-ranging information. Such designs must take into account the turbidity of biological tissue. Biological tissue is distinct from most other spectroscopic media in that it is highly turbid, due to elastic scattering by the rich variety of microscopic structures that comprise the tissue. Thus, even when absorption is negligible, light can penetrate only a few millimetres into the tissue, arranging itself in a diffuse blob. Up to now, this has been viewed as a great problem, causing spectral distortions and limiting light collection and depth resolution. With properly designed Raman probes, however, it may be possible to turn elastic scattering into an advantage or, alternately, to devise ways of eliminating it. This is of interest because Raman signals have the potential to penetrate much more deeply into tissue, perhaps providing a transcutaneous means to probe Alzheimer's disease through the skull.

As discussed above, great progress has been made in developing modelling techniques for classifying tissue and extracting concentration information. This exciting new area of research can be profitably extended in a number of ways. Extraction of morphological information is one very important example. The ability to provide the clinician with information about the morphological composition of the tissue in the cubic millimetre at the tip of the probe in a fraction of a second during surgery or a diagnostic procedure is of enormous potential significance, and will essentially bring the insights of pathology into the operating room and

catheterization laboratory. Problems such as the modelling of layered tissue will have to be addressed. Another interesting observation, indicated by the blood analysis study of section 5, is that current Raman technology can extract concentrations of the order of a few parts per million. And yet, in most other modelling studies, accuracy of the order of only a few per cent is currently obtained. This indicates that with new techniques and new insights, it should be possible to develop models with far greater accuracy. Moreover, there is no reason to assume that Raman technology should be limited to concentration accuracies of a few parts per million. With improved technology and understanding, further improvements in accuracy are possible.

The combination of Raman diagnostic techniques with other technologies should also be explored. For example, in the case of atherosclerosis, combining Raman spectroscopy with intravascular ultrasound is worthy of exploration (Römer *et al* 1998c). Raman spectroscopy can also serve as an important feedback element in combined 'see and treat' systems. For example, in the case of laser angiosurgery (Litvack *et al* 1990), it should be possible to use Raman spectroscopy for feedback control of tissue removal by laser ablation. Precision feedback-controlled laser microsurgery has the potential to remove atherosclerotic plaques without the trauma associated with balloon angioplasty, which is believed to be responsible for its high rate of restenosis (Rozenman *et al* 1995). Past attempts to use laser light to remove arterial plaques, a technique known as 'laser angioplasty', have also resulted in a high rate of restenosis (Ghazzal *et al* 1995). This probably resulted from excessive laser damage, which Raman-controlled laser microsurgery could avoid.

Of course, feedback-controlled laser microsurgery would not be limited to atherosclerosis. Raman-guided laser brain surgery and other applications that require high precision can also be envisioned. In addition, real-time Raman diagnosis can provide feedback signals for other treatment modalities, including cryo-, thermal, radiofrequency and other types of ablation and even those involving drug delivery.

Another interesting area of consideration is the use of Raman spectroscopy for functional imaging. Once the problems of probe delivery are resolved, Raman spectroscopy can be used in an endoscopic configuration to create endoscopic images of the chemical or morphological properties of the tissue. Deep tissue Raman imaging may also be possible, once problems of turbidity are better understood. In addition, it may be possible to combine Raman spectroscopy with existing imaging techniques such as x-ray CT and magnetic resonance imaging or, as discussed above, intravascular ultrasound. In this case, the chemical and morphological information provided by Raman spectroscopy would complement the tissue density information provided by the existing imaging technologies.

In conclusion, this article has reviewed the substantial accomplishments of biomedical Raman spectroscopy over the past few years. The technology has tremendous clinical potential, and there is substantial room for future advances. We anticipate that within the next five years, the remarkable prospects for *in vivo* Raman spectroscopy will begin to emerge as practical clinical tools.

Acknowledgments

Much of the work described in this article has benefited from the research of many collaborators and colleagues in the field of biomedical optics and spectroscopy. We also thank Dr Abha Sur for her contributions to the biographical sketch of Professor Sir C V Raman. Finally, we acknowledge support of the MIT Laser Biomedical Research Center by the National Institutes of Health, under grant number P41-RR02594.

References

Abraham J L and Etz E S 1979 Molecular microanalysis of pathological specimens *in situ* with a laser Raman microprobe *Science* **206** 716–18

Alfano R R, Liu C H, Sha W L, Zhu H R, Akins D L, Cleary J, Prudente R and Clemer E 1991 Human breast tissues studied by IR Fourier transform Raman spectroscopy *Lasers Life Sci.* **4** 23–8

Alini M and Losa G A 1991 Partial characterization of proteoglycans isolated from neoplastic and nonneoplastic human breast tissues *Cancer Res.* **51** 1443–7

Allred C D and McCreery R L 1990 Near-infrared Raman spectroscopy of liquids and solids with a fiberoptic sampler, diode laser, and CCD detector *Appl. Spectrosc.* **44** 1229–31

American Heart Association 1998 *1998 Heart and Stroke Statistical Update* (Dallas, TX: American Heart Association)

Asher S A 1988 UV resonance Raman studies of molecular-structure and dynamics—applications in physical and biophysical chemistry *Ann. Rev. Phys. Chem.* **39** 537–58

——1993 UV resonance Raman-spectroscopy for analytical, physical, and biophysical chemistry 1, 2 *Anal. Chem.* **65** A59–A66, A201–A210

Asher S A and Johnson C R 1984 Raman-spectroscopy of a coal liquid shows that fluorescence interference is minimized with ultraviolet excitation *Science* **225** 311–13

Bakker Schut T C, Puppels G J, Kraan Y M, Greve J, van der Maas L L and Figdor C G 1997 Intracellular carotenoid levels measured by Raman microspectrocopy: comparison of lymphocytes from lung cancer patients and healthy individuals *Int. J. Cancer* **74** 20–5

Baraga J J 1992 *In situ* chemical analysis of biological tissue: vibrational Raman spectroscopy of human atherosclerosis *PhD Dissertation* Massachusetts Institute of Technology

Baraga J J, Feld M S and Rava R P 1992a *In situ* optical histochemistry of human artery using near infrared Fourier transform Raman spectroscopy *Proc. Natl Acad. Sci.* **89** 3473–7

——1992b Rapid near-infrared Raman spectroscopy of human tissue with a spectrograph and CCD detector *Appl. Spectrosc.* **46** 187–90

Barbillat J, DaSilva E and Hallaert J L 1993 Multi-channel micro-Raman spectroscopy with near-infrared excitation II *J. Raman Spectrosc.* **24** 53–62

Barry B W, Edwards H G M and Williams A C 1992 Fourier-transform Raman and infrared vibrational study of human skin—assignment of spectral bands *J. Raman Spectrosc.* **23** 641–5

Bartorelli A L, Leon M B, Almagor Y, Prevosti L G, Swain J A, McIntosh C L, Neville R F, House M D and Bonner R F 1991 *In vivo* human atherosclerotic plaque recognition by laser-excited fluorescence spectroscopy *J. Am. Coll. Cardiol.* **17** 160B–168B

Bauer N J C, Wicksted J P, Jongsma F H M, March W F, Hendrikse F and Motamedi M 1998 Non-invasive assessment of the hydration gradient across the cornea using confocal Raman spectroscopy *Invest. Opthalmol. Vis. Sci.* **39** 831–5

Berger A J 1998 Measurement of analytes in human serum and whole blood samples by near-infrared Raman spectroscopy *PhD Dissertation* Massachusetts Institute of Technology

Berger A J and Feld M S 1997 Analytical method of estimating chemometric prediction error *Appl. Spectrosc.* **51** 725–32

Berger A J, Koo T-W, Itzkan I and Feld M S 1998 An enhanced algorithm for linear multivariate analysis *Anal. Chem.* **70** 623–7

Berger A J, Koo T-W, Itzkan I, Horowitz G and Feld M S 1999 Multicomponent blood analysis by near-infrared Raman spectroscopy *Appl. Opt.* **38** 2916–26

Bohorfoush A G 1996 Tissue spectroscopy for gastrointestinal diseases *Endoscopy* **28** 372–80

Bot A C C, Huizinga A, de Mul F F M, Vrensen G F J M and Greve J 1989 Raman microspectroscopy of fixed rabbit and human lenses and lens slices—new potentialities *Exp. Eye Res.* **49** 161–9

Boustany N 1997 Characterization of mucosal dysplasia with ultraviolet resonance Raman spectroscopy *PhD Thesis* (Massachusetts Institute of Technology)

Boustany N, Crawford J M, Manoharan R, Dasari R R and Feld M S 1999 Analysis of nucleotides and aromatic amino acids in normal and neoplastic colon mucosa by ultraviolet resonance Raman spectroscopy *Lab. Invest.* **79** 1201–14

Brennan J F 1995 Near infrared Raman spectroscopy for human artery histochemistry and histopathology *PhD Dissertation* (Massachusetts Institute of Technology)

Brennan J F, Beattie M E, Wang Y, Cantella M J, Tsaur B-Y, Dasari R R and Feld M S 1996 PdSi focal-plane array detectors for short-wave infrared Raman spectroscopy of biological tissue: a feasibility study *Appl. Opt.* **35** 5736–9

Brennan J F, Römer T J, Lees R S, Tercyak A M, Kramer J R and Feld M S 1997a Determination of human coronary artery composition by Raman spectroscopy *Circulation* **96** 99–105

Brennan J F, Wang Y, Dasari R R and Feld M S 1997b Near-infrared Raman spectrometer systems for human tissue studies *Appl. Spectrosc.* **51** 201–8

Buitveld H, De Mul F F M, Mud J and Greve J 1984 Identification of inclusions in lung tissue with a Raman microprobe *Appl. Spectrosc.* **38** 304–6

Butterfield D A 1986 Spectroscopic methods in degenerative neurological diseases *CRC Crit. Rev. Clin. Neurobiol.* **2** 169–240

Campbell I D and Dwek R A 1984 *Biological Spectroscopy* (New York: Benjamin/Cummings)

Caspers P J, Lucassen G W, Wolthuis R, Bruining H A and Puppels G J 1998 *In vitro* and *in vivo* Raman spectroscopy of human skin *Biospectroscopy* **4** S31–39

Centeno J A, Ishak K, Mullick F G, Gahl W A and O'Leary T J 1994 Infrared microspectroscopy and laser Raman microprobe in the diagnosis of cystinosis *Appl. Spectrosc.* **48** 569–72

Chase B and Talmi Y 1991 The use of a near-infrared array detector for Raman spectroscopy beyond one micron *Appl. Spectrosc.* **45** 929–31

Choo L P, Mansfield J R, Pizzi N, Somorjai L, Jackson M, Halliday W C and Mantsch H H 1995 Infrared spectra of human central nervous system tissue: diagnosis of Alzheimer's disease by multivariate analyses *Biospectroscopy* **1** 141–8

Clarke R H, Hanlon E B, Isner J M and Brody H 1987 Laser Raman spectroscopy of calcified atherosclerotic lesions in cardiovascular tissue *Appl. Opt.* **26** 3175–7

Clarke R H, Wang Q and Isner J M 1988 Laser Raman spectroscopy of atherosclerotic lesions in human coronary artery segments *Appl. Opt.* **27** 4799–800

Clarke W L, Cox D, Gonder-Frederick L A, Carter W and Pohl S L 1987 Evaluating clinical accuracy of systems for self-monitoring of blood glucose *Diabetes Care* **10** 622–8

Coté G L, Fox M D and Northrop R B 1992 Noninvasive optical polarimetric glucose sensing using a true phase measurement technique *IEEE Trans. Biomed. Eng.* **39** 752–5

Cothren R M, Hayes G B, Kramer J R, Sacks B A, Kittrell C and Feld M S 1986 A multifiber catheter with an optical shield for laser angiosurgery *Lasers Life Sci.* **1** 1–12

Dalterio R A, Nelson W H, Britt D, Sperry J and Purcell F J 1986 A resonance Raman microprobe study of chromobacteria in water *Appl. Spectrosc.* **40** 271–3

Deinum G, Rodriguez D, Römer T J, Fitzmaurice M, Kramer J R and Feld M S 1999 Histological classification of Raman spectra of human coronary artery atherosclerosis using principal component analysis *Appl. Spectrosc.* **53** 938–42

Demer L L 1997 Lipid hypothesis of cardiovascular calcification *Circulation* **95** 297–8

Diem M 1993 *Introduction to Modern Vibrational Spectroscopy* (New York: Wiley)

Dou X M, Yamaguchi Y, Yamamoto H, Uenoyama H and Ozaki Y 1996 Biological applications of anti-stokes Raman spectroscopy: quantitative analysis of glucose in plasma and serum by a highly sensitive multichannel Raman spectrometer *Appl. Spectrosc.* **50** 1301–6

Duindam J J, Vrensen G F, Otto C and Greve J 1998 Cholesterol, phospholipids, and protein changes in focal opacities in the human eye lens *Invest. Ophthalmol. Vis. Sci.* **39** 94–103

Duindam J J, Vrensen G F, Otto C, Puppels G J and Greve J 1995 New approach to assess the cholesterol distribution in the eye lens: confocal Raman microspectroscopy and filipin cytochemistry *J. Lipid Res.* **36** 1139–46

Elmore G, Barton M B, Moceri V M, Polk S, Arena P J and Fletcher S W 1998 Ten-year risk of false positive screening mammograms and clinical breast examinations *New Eng. J. Med.* **3883** 1089–96

Fabian H, Choo L-P, Szendrei G I, Jackson M, Halliday W C, Otvos L Jr and Mantsch H H 1993 Infrared spectroscopic characterization of Alzheimer plaques *Appl. Spectrosc.* **47** 1513–18

Farrell R and McCauley R 1976 On corneal transparency and its loss with swelling *J. Opt. Soc. Am.* **66** 342–5

Feld M S and Kramer J R 1989 Laser angiosurgery: a biomedical system using photons to diagnose and treat atherosclerosis *Future Directions in Interventional Cardiology* ed J H K Vogel and S B King (St Louis, MO: Mosby) pp 179–87

——1991 Mutagenicity and the XeCl excimer laser—a relationship of consequence *Am. Heart J.* **122** 1803–5

Fendel S and Schrader B 1998 Investigation of skin and skin lesions by NIR-FT Raman spectroscopy *Fresenins J. Anal. Chem.* **360** 609–13

Fishbein M C and Siegal R J 1996 How big are coronary atherosclerotic plaques that rupture? *Circulation* **94** 2662–6

Fisher L and Van Belle G 1993 *Biostatistics: A Methodology for the Health Sciences* (New York: Wiley)

Fodor S P A, Rava R P, Copeland R A and Spiro T G 1986 H-2 Raman-shifted YAG laser ultraviolet Raman spectrometer operating at wavelengths down to 184 nm *J. Raman Spectrosc.* **17** 471–5

Frank C J, McCreery R L and Redd D C B 1995 Raman spectroscopy of normal and diseased human breast tissues *Anal. Chem.* **67** 777–83

Frank C J, Redd D C B, Gansler T S and McCreery R L 1994 Characterization of human breast biopsy specimens with near-IR Raman spectroscopy *Anal. Chem.* **66** 319–26

Freshour B G and Koenig J L 1975 Raman scattering of collagen, gelatin, and elastin *Biopolymers* **14** 379–91
Geladi P and Kowalski B R 1986 Partial least-squares regression: a tutorial *Anal. Chim. Acta* **185** 1–17
Ghazzal Z M, Burton E, Weintraub W S, Litvack F, Rothbaum D A, Klein L and King S B III 1995 Predictors of restenosis after excimer laser coronary angioplasty *Am. J. Cardiol.* **75** 1012–14
Gniadecka M, Faurskov Nielsen O, Christensen D H and Wulf H C 1998 Structure of water, proteins, and lipids in intact human skin, hair and nail *J. Invest. Dermatol.* **110** 393–8
Goheen S C, Lis L J and Kauffman J W 1978 Raman spectroscopy of intact feline corneal collagen *Biochem. Biophys. Acta* **536** 197–204
Groot P H E, van Vlijmen B J M, Benson G M, Hofker M H, Schiffelers R, VidgeonHart M and Havekes L M 1996 Quantitative assessment of aortic atherosclerosis in APOE*3 Leiden transgenic mice and its relationship to serum cholesterol exposure *Arterioscler. Thromb. Vasc. Biol.* **16** 926–33
Grygon C A, Perno J R, Fodor S P A and Spiro T G 1988 Ultraviolet resonance Raman spectroscopy as a probe of protein-structure in the FD virus *Biotechniques* **6** 50–5
Haaland D M and Thomas E V 1988 Partial least-squares methods for spectral analysis 1. Relation to other quantitative calibration methods and the extraction of qualitative information *Anal. Chem.* **60** 1193–202
Hahn D W, Wohlfarth D L and Parks N L 1997 Analysis of polyethylene wear debris using micro Raman spectroscopy: a report on the presence of beta-carotene *J. Biomed. Mater. Res.* **35** 31–7
Hawi S R, Campbell W B, Kajdacsy-Balla A, Murphy R, Adar F and Nithipatikom K 1996 Characterization of normal and malignant human hepatocytes by Raman microspectroscopy *Cancer Lett.* **110** 35–40
Hawi S R, Nithipatikom K, Wohlfeil E R, Adar F and Campbell W B 1997 Raman microspectroscopy of intracellular cholesterol crystals in cultured bovine coronary artery endothelial cells *J. Lipid. Res.* **38** 1591–7
Henry R J, Cannon D C and Winkelman J W 1974 *Clinical Chemistry* (New York: Harper & Row)
Hirschfeld T and Chase B 1986 FT-Raman spectroscopy—development and justification *Appl. Spectrosc.* **40** 133–7
Hoey S, Brown D H, McConnell A A, Smith W E, Marabani M and Sturrock R D 1988 Resonance Raman spectroscopy of hemoglobin in intact cells: a probe of oxygen uptake by erythrocytes in rheumatoid arthritis *J. Inorg. Biochem.* **34** 189–99
Hudson B and Sension R J 1989 *Raman Spectroscopy : Sixty Years on Vibrational Spectroscopy and Structure* vol 174, ed H D Bist, J R Durig and J F Sullivan, p 363
Hyman B T and Trojanowski J Q 1997 Consensus recommendations for the postmortem diagnosis of Alzheimer disease from the National Institute on Aging and the Reagan Institute Working Group on diagnostic criteria for the neuropathological assessment of Alzheimer's disease *J. Neuropathol. Exp. Neurol.* **56** 1095–7
Ishida H, Kamoto R, Uchida S, Ishitani A, Ozaki Y, Iriyama K, Tsukie T, Shibata K, Ishihara F and Kameda H 1987 Raman microprobe and Fourier transform-infrared microsampling studies of the microstructure of gallstones *Appl. Spectrosc.* **41** 407–12
Itzkan I and Izatt J A 1994 Medical use of lasers *Encyclopedia of Applied Physics* vol 10 (Portland, OR: VCH) pp 33–59
Jongsma F H M, Erckens R J, Wicksted J P, Bauer N J C, Hendrikse F, March W F and Motamedi M 1997 Confocal Raman spectroscopy system for noncontact scanning of ocular tissues: an *in vitro* study *Opt. Eng.* **36** 3193–9
Kalasinsky V F, Johnson F B and Ferwerda R 1998 Fourier transform IR and Raman microspectroscopy of materials in tissue *Cell Mol. Biol.* **44** 141–4
Keller S, Schrader B, Hoffman A, Schrader W, Metz K, Rehlaender A Pahnke J, Ruwe M and Budach W 1994 Application of near-infrared Fourier-transform Raman spectroscopy in medical research *J. Raman Spectrosc.* **25** 663–71
Kline N and Treado P J 1997 Raman chemical imaging of breast tissue *J. Raman Spectrosc.* **28** 119–24
Klug D D, Singleton D L and Walley V M 1992 Laser Raman spectrum of calcified human aorta *Lasers Surg. Med.* **12** 13–7
Kodati V R, Tomasi G E, Turmin J L and Tu A T 1991 Raman spectroscopic identification of phosphate-type kidney stones *Appl Spectrosc.* **45** 581–3
Kodati V R, Tu A T and Turmin J L 1990 Raman spectroscopic identification of uric-acid-type kidney stone *Appl. Spectrosc.* **44** 1134–6
Koo T-W, Berger A J, Itzkan I, Horowitz G and Feld M S 1998 Measurement of glucose in human blood serum using Raman spectroscopy *IEEE LEOS Mag.* **12** 18
——1999 Reagentless blood analysis by near-infrared Raman spectroscopy *Diabetes Technol. Therapeu.* **1** 153–7
Kovacs I, Torok I, Zombori J and Kasa P 1996 Neuropathologic changes in the olfactory bulb in Alzheimer's disease *Neurobiology* **4** 123–6
——1998 Cholinergic structures and neuropathologic alterations in the olfactory bulb of Alzheimer's disease brain samples *Brain Res.* **789** 167–70
Kowall N and McKee A 1993 The histopathology of neuronal degeneration and plasticity in Alzheimer's disease *Adv. Neurol.* **59** 5–33

Kramer R 1998 *Chemometric Techniques for Quantitative Analysis* (New York: Dekker)
Larsson K and Hellgren L 1974 A study of the combined Raman and fluorescence scattering from human blood plasma *Experientia* **30** 481–3
Lawson E E, Anigbogu A N, Williams A C, Barry B W and Edwards H G 1998 Thermally induced molecular disorder in human stratum corneum lipids compared with a model phospholipid system: FT Raman spectroscopy *Spectrochim. Acta A: Mol. Biomol. Spectrosc.* **54** 543–58
Lawson E E, Barry B W, Williams A C and Edwards H G 1997 Biomedical applications of Raman spectroscopy *J. Raman Spectrosc.* **28** 111–17
Lentner C 1985 *Geigy Scientific Table* (Basel, Switzerland: Ciba-Geigy)
Leon L J 1990 *Linear Algebra with Applications* (New York: Macmillan)
Litvack F *et al* 1990 Percutaneous excimer laser coronary angioplasty *Am. J. Cardiol.* **66** 1027–32
Liu C H *et al* 1992b Raman, fluorescence and time-resolved light-scattering as optical diagnostic techniques to separate diseased and normal biomedical media *J. Photochem. Photobiol. B: Biol.* **16** 187–209
Liu C H, Glassman W L S, Zhu H R, Akins D L, Deckelbaum L I, Stetz M L, O'Brien K, Scott J and Alfano R R 1992a Near-IR Fourier transform Raman spectroscopy of normal and atherosclerotic human aorta *Lasers Life Sci.* **4** 257–64
Lord R C and Yu N T 1970 Laser-excited Raman spectroscopy of biomolecules I. Native lysozyme and its constituent amino acids *J. Mol. Biol.* **50** 509
Luke J L, Kalasinsky V F, Turnicky R P, Centeno J A, Johnson F B and Mullick F G 1997 Pathological and biophysical findings associated with silicone breast implants: a study of capsular tissues from 86 cases *Plast. Reconstr. Surg.* **100** 1558–65
Mach F, Schonbeck U, Bonnefoy J Y, Pober J S and Libby P 1997 Activation of monocyte/macrophage functions related to acute atheroma complication by ligation of CD40 *Circulation* **96** 396–9
Mahadevan-Jensen A, Mitchell M F, Ramanujam N, Malpica A, Thompsen S, Utzinger U and Richards-Kortum R 1998a NIR Raman spectroscopy for *in vitro* detection of cervical precancers *Photochem. Photobiol.* **68** 123–32
Mahadevan-Jensen A, Mitchell M F, Ramanujam N, Utzinger U and Richards-Kortum R 1998b Development of a fiber optic probe to measure NIR Raman spectra of cervical tissue *in vivo Photochem. Photobiol.* **68** 427–31
Mahadevan-Jensen A and Richards-Kortum R 1996 Raman spectroscopy for the detection of cancers and precancers *J. Biomed. Opt.* **1** 31–70
Mann J M and Davies M J 1996 Vulnerable plaque: relation of characteristics to degree of stenosis in human coronary arteries *Circulation* **94** 928–31
Manoharan R, Baraga J J, Feld M S and Rava R P 1992 Quantitative histochemical analysis of human artery using Raman spectroscopy *J. Photochem. Photobiol. B: Biol.* **16** 211–33
Manoharan R, Baraga J J, Rava R P, Dasari R R, Fitzmaurice M and Feld M S 1993a Biochemical analysis and mapping of atherosclerotic human artery using FT-IR microspectroscopy *Atherosclerosis* **103** 181–93
Manoharan R, Ghiamati E, Chadha S, Nelson W H and Sperry J F 1993b Effect of cultural conditions on deep UV resonance Raman spectra of bacteria *Appl. Spectrosc.* **47** 2145–50
Manoharan R, Ghiamati E, Dalterio R A, Britton K A, Nelson W H and Sperry J F 1990 UV resonance Raman spectra of bacteria, bacterial spores, protoplasts and calcium dipicolinate *J. Microbiol. Methods* **11** 1–15
Manoharan R, Shafer K, Perelman L, Wu J, Chen K, Deinum G, Fitzmauice M, Myles J, Crowe J, Dasari R R and Feld MS 1998 Raman spectroscopy and fluorescence photon migration for breast cancer diagnosis and imaging *Photochem. Photobiol.* **67** 15–22
Manoharan R, Wang Y, Dasari R R, Singer S, Rava R P and Feld M S 1995 UV resonance Raman spectroscopy for detection of colon cancer *Lasers in Life Sci.* **6** 217–27
Manoharan R, Wang Y and Feld M S 1996 Review: histochemical analysis of biological tissues using Raman spectroscopy *Spectrochem. Acta A* **52** 215–49
Marbach R, Koschinsky T H, Gries F A and Heise H M 1993 Noninvasive blood glucose assay by near-infrared diffuse reflectance spectroscopy of the human inner lip *Appl. Spectrosc.* **47** 875–81
Marchant G J 1994 Contemporary management of breast cancer *Obstet. Gynecol. Clin. North Am.* **21** 555–60
Martens H and Naes T 1989 *Multivariate Calibration* (Chichester: Wiley)
McCreery R L 1995 *Applications of Raman Spectroscopy* ed T Laserna (New York: Wiley)
McGill N, Dieppe P A, Bowden M, Gardiner D J and Hall M 1991 Identification of pathological mineral-deposits by Raman microscopy *Lancet* **337** 77–8
McKee A C, Kosik K and Kowall K N 1991 Neuritic pathology and dementia in Alzheimer's disease *Ann. Neurol.* **30** 156–65
Mirra S S, Gearing M, McKeel D W Jr, Crain B J, Hughes J P, van Belle G and Heyman A 1994 Inter-laboratory comparison of neuropathology assessments in Alzheimer's disease: a study of the consortium to establish a registry for Alzheimer's disease (CERAD) *J. Neuropathol. Exp. Neurol.* **53** 303–15 (Erratum *J. Neuropathol. Exp. Neurol.* 1994 **53** 425)

Mitchinson M J 1982 Insoluble lipids in human atherosclerotic plaques *Atherosclerosis* **45** 11–15

Mizuno A, Hayashi T, Tashibu K, Maraishi S, Kawauchi K and Ozaki Y 1992 Near-infrared FT-Raman spectra of the rat-brain tissues *Neurosci. Lett.* **141** 47–52

Mizuno A, Kitajima H, Kawauchi K, Muraishi S and Ozaki Y 1994 Near infrared Fourier transform Raman spectroscopic study of human brain tissues and tumours *J. Raman Spectrosc.* **25** 265–9

Mizuno A, Ozaki Y, Kamada Y, Miyazaki H, Itoh K and Iriyama K 1981 Direct measurement of Raman spectra of intact lens in a whole eyeball *Curr. Eye Res.* **1** 609–13

Moreno P R *et al* 1996 Macrophage infiltration predicts restenosis after coronary intervention in patients with unstable angina *Circulation* **94** 3098–102

National Cancer Institute 1997 Breast Cancer *National Institute of Health News Release* March 7

Nelson W H, Manoharan R and Sperry J F 1992 UV resonance Raman studies of bacteria *Appl. Spectrosc. Rev.* **27** 67–124

Nie S M, Bergbauer K L, Kuck J F R and Yu N T 1990 Near-infrared Fourier-transform Raman-spectroscopy in human lens research *Exp. Eye Res.* **51** 619–23

O'Leary T J and Levin I W 1985 Secondary structure of endocrine amyloid: infrared spectroscopy of medullary carcinoma of the thyroid *Lab. Invest.* **53** 240–2

Ozaki Y 1988 Medical application of Raman spectroscopy *Appl. Spectrosc. Rev.* **24** 259–312

Ozaki Y, Kaneuchi F, Iwamoto T, Yoshiura M and Iriyama K 1989 Nondestructive analysis of biological-materials by FT-IR-ATR 1 Direct evidence for the existence of collagen helix in lens capsule *Appl. Spectrosc.* **43** 138–41

Pagano M and Gauvreau K 1993 *Principles of Biostatistics* (Belmont, CA: Duxbury)

Pestaner J P, Mullick F G, Johnson F B and Centeno J A 1996 Calcium oxalate crystals in human pathology: molecular analysis with the laser Raman microprobe *Arch. Pathol. Lab. Med.* **120** 537–40

Peters R J G, Kok W E, Havenith M G, Rijsterborgh H, van der Wal A C and Visser C A 1994 Histopathologic validation of intracoronary ultrasound imaging *J. Am. Soc. Echocardiogr.* **7** 230–41

Puppels G J 1997 Development of *in vivo* diagnostic tools based on Raman spectroscopy *24th Ann. Conf. FACSS Conf. (Providence, RI, 1997)* (Santa Fe, NM: FACSS)

Puppels G J, Demul F F M, Otto C, Greve J, Robertnicoud M, Arndtjovin D J and Jovin T M 1990 Studying single living cells and chromosomes by confocal Raman microspectroscopy *Nature* **347** 301–3

Puppels G J, Garristen H S P, Kummer J A and Greve J 1993 Carotenoids located in human lymphocyte subpopulations and natural killer cells by Raman microspectroscopy *Cytometry* **14** 251–6

Puppels G J, Garritsen H S P, Segersnolten G M J, Demul F F M and Greve J 1991a Raman microspectroscopic approach to the study of human granulocytes *Biophys. J.* **60** 1046–56

Puppels G J, Olminkhof J H, Segers-Nolten G M, Otto C, de Mul F F and Greve J 1991b Laser irradiation and Raman spectroscopy of single living cells and chromosomes: sample degradation occurs with 5145 nm but not with 660 nm laser light *Exp. Cell Res.* **195** 361–7

Puppels G J, Otto C, Greve J, Robert-Nicoud M, Arndt-Jvin D J and Jovin T M 1994 Raman microspectroscopic study of low pH induced changes in DNA structure of polytene chromosomes *Biochemistry* **33** 3386–95

Puppels G J, van Aken T, Wolthuis R, Caspers P J, Bakker Schut T, Bruining H A, Romer T J, Buschman H P J, Wach M L and Robinson J S J 1998 *In vivo* tissue characterization by Raman spectroscopy *Proc. BIOS/SPIE* **3257** 78–83

Radi M 1989 Calcium oxalate crystals in breast biopsies *Arch. Pathol. Lab. Med.* **113** 1367–9

Rasmussen R E, Hammer-Wilson M and Berns M W 1989 Mutation and sister chromatid exchange induction in Chinese hamster ovary (CHO) cells by pulsed excimer laser radiation at 193 nm and 308 nm and continuous UV radiation at 254 nm *Photochem. Photobiol.* **49** 413–18

Rava R P, Baraga J J and Feld M S 1991 Near-infrared Fourier-transform Raman-spectroscopy of human artery *Spectrochim. Acta A* **47** 509–12

Redd D C, Yue K T, Martin L G and Kaufman S L 1991 Raman spectroscopy of human atherosclerotic plaque: implications for laser angioplasty *J. Vasc. Interv. Radiol.* **2** 247–52

Redd D C B, Feng Z C, Yue K T and Gansler T S 1993 Raman spectroscopic characterization of human breast tissues: implications for breast cancer diagnosis *Appl. Spect.* **47** 787–91

Rehman I, Smith R, Hench L L and Bonfield W 1995 Structural evaluation of human and sheep bone and comparison with synthetic hydroxyapatite by FT Raman spectroscopy *J. Biomed. Mater. Res.* **29** 1287–94

Richards-Kortum R, Rava R P, Fitzmaurice M, Tong L L, Ratliff N B, Kramer J R and Feld M S 1989 A one layer model of laser-induced fluorescence for diagnosis of disease in human tissue: applications to atherosclerosis *IEEE Trans. Biomed. Eng.* **36** 1222–31

Richards-Kortum R and Sevick-Muraca E 1996 Quantitative optical spectroscopy for tissue diagnosis *Ann. Rev. Phys. Chem.* **47** 455–606

Rippon W P, Koenig J L and Walton A 1971 Laser Raman spectroscopy of biopolymers and proteins *Agri. Food Chem.* **19** 692–7

Robb E W and Munk M E 1990 A neural network approach to infrared spectrum interpretation *Mikrochim. Acta* **1** 131–55

Römer T J, Brennan J F, Fitzmaurice M, Feldstein M L, Deinum G, Myles J L, Kramer J R, Less R S and Feld M S 1998a Histopathology of human coronary atherosclerosis by quantifying its chemical composition with Raman spectroscopy *Circulation* **97** 878–85

Römer T J, Brennan J F, Puppels G J, Laarse A V D, Princen H M G, Folger O, Buschman H P J, Jukema J W, Havekes L M and Bruschke A V G 1998b Raman spectroscopy provides chemical mappings of atherosclerotic plaques in APOE*3 Leiden transgenic mice *Arteriosclerosis Thrombosis Vasc. Biol.* in review

Römer T J, Brennan J F, Puppels G J, Tuinenburg J, van Duinen S G, van der Laarse A, van der Steen A F W, Bom N A and Bruschke A V G 1998c Intravascular ultrasound combined with Raman spectroscopy to localize and quantify cholesterol and calcium salts in atherosclerotic coronary arteries *Arteriosclerosis Thrombosis Vasc. Biol.* in review

Rozenman Y, Gilon D, Welber S, Sapoznikov D, Lotan C, Mosseri M, Weiss T, Hasin Y and Gotsman M S 1995 Influence of coronary angioplasty on the progression of coronary atherosclerosis *Am. J. Cardiol.* **76** 1126–30

Rubin E and Farber J L 1994 *Pathology* (Philadelphia, PA: Lippincott)

Salenius J P, Brennan J F, Miller A, Wang Y, Aretz T, Sacks B, Dasari R R and Feld M S 1998 Biochemical composition of human peripheral arteries using near infrared Raman spectroscopy *J. Vasc. Surg.* **27** 710–19

Schaeberle M D, Kalasinsky V F, Luke J L, Lewis E N, Levin I W and Treado P J 1996 Raman chemical imaging: histopathology of inclusions in human breast tissue *Anal. Chem.* **66** 1829–33

Schrader B, Dippel B, Fendel S, Keller S, Lochte T, Riedl M, Schulte R and Tatsch E 1997 NIR FT Raman spectroscopy—a new tool in medical diagnostics *J. Mol. Struct.* **408** 23–31

Schweers O, Schonbrunn-Hanebeck D, Marx A and Mandelkow E 1994 Structural studies of tau protein and Alzheimer paired helical filaments show no evidence for β-Structure, *J. Biol. Chem.* **39** 24 290–7

Sebag J, Nie S, Reiser K, Charles M A and Yu N T 1994 Raman spectroscopy of human vitreous in proliferative diabetic retinopathy *Invest. Opthalmol. Vis. Sci.* **35** 2976–80

Sharma S 1996 *Applied Multivariate Techniques* (New York: Wiley)

Shim M G and Wilson B C 1996 The effects of *ex vivo* handling procedures on the near-infrared Raman spectra of normal mammalian tissues *Photochem. Photobiol.* **63** 662–71

——1997 Development of an *in vivo* Raman spectroscopic system for diagnostic applications *J. Raman Spectrosc.* **28** 131–42

Shreve A P, Cherepy N J and Mathies R A 1992 Effective rejection of fluorescence in interference in Raman spectroscopy using a shifted excitation difference technique *Appl. Spectrosc.* **46** 707–11

Siebenga I, Vrensen G F, Otto K, Puppels G J, De Mul F F and Greve J 1992 Aging and changes in protein conformation in the human lens: a Raman microspectroscopy study *Exp. Eye Res.* **54** 759–67

Singer S, Souza K and Thilly W G 1995 Pyruvate utilization, phosphocholine and adenosine triphosphate (ATP) are markers of human breast tumor progression: a ^{31}P- and ^{13}C-nuclear magnetic resonance (NMR) spectroscopy study *Cancer* **55** 5140–5

Sliney D and Wolbarsht M 1980 *Safety with Lasers and Other Optical Sources: a Comprehensive Handbook* (New York: Plenum)

Smeets M H, Vrensen G F J M, Otto K, Puppels G J and Greve J 1993 Local variations in protein-structure in the human eye lens—a Raman microspectroscopic study *Biophys. Biochim. Acta* **1164** 236–42

Spiro T G (ed) 1988 *Biological Applications of Raman Spectroscopy* vol II (New York: Wiley)

Stary H C, Chandler A B, Dinsmore R E, Fuster V, Glagov S, Insull W Jr, Rosenfeld M E, Schwartz C J, Wagner W D and Wissler R W 1995 A definition of advanced types of atherosclerotic lesions and a histological classification of atherosclerosis *Circulation* **92** 1355–74

Sternberg S S 1997 *Histology for Pathologists* (Philadelphia: Lippincott-Raven)

Sureau F, Chinsky L, Amirand C, Ballini J P, Duquesne M, Laigle A, Turpin P Y and Vigny P 1990 An ultraviolet micro-Raman spectrometer—resonance Raman spectroscopy within single living cells *Appl. Spectrosc.* **44** 1047–51

Tarr R V and Steffes P G 1993 Non-invasive blood glucose measurement system and method using stimulated Raman spectroscopy *United States Patent* 5243 983

Toussaint J F, LaMuraglia G M, Southern J F, Fuster V and Kantor H L 1996 Magnetic resonance images lipid, fibrous, calcified, hemorrhagic, and thrombotic components of human atherosclerosis *in vivo Circulation* **94** 932–8

Treado P J, Levin I W and Lewis E N 1994 Indium-antimonide (InSb) focal-plane array (FPA) detection for near-infrared imaging microscopy *Appl. Spectrosc.* **48** 607–15

Turrell G and Corset J (ed) 1996 *Raman Microscopy* (New York: Wiley)

Wang Y and McCreery R L 1989 Evaluation of a diode-laser charge coupled device spectrometer for near-infrared Raman-spectroscopy *Anal. Chem.* **61** 2647–51

Waynant R W and Chenault V M 1998 Overview of non-invasive fluid glucose measurement using optical techniques to maintain glucose control in diabetes mellitus *IEEE LEOS Newsletter* **12** 3–6

Wen Z Q, Overman S A and Thomas G J Jr 1997 Structure and interactions of the single-stranded DNA genome of filamentous virus fd: investigation by ultraviolet resonance Raman spectroscopy *Biochemistry* **36** 7810–20

Wen Z Q and Thomas G J Jr 1998 UV resonance Raman spectroscopy of DNA and protein constituents of viruses: assignments and cross sections for excitations at 257, 244, 238, and 229 nm *Biopolymers* **45** 247–56

Wexler L, Brundage B, Crouse J, Detrano R, Fuster V, Maddahi J, Rumberger J, Stanford W, White R and Taubert K 1996 Coronary artery calcification: pathophysiology, epidemiology, imaging methods, and clinical implications: a statement for health professionals from the American Heart Association *Circulation* **94** 1175–92

Williams A C, Barry B W, Edwards H G and Farwell D W 1993 A critical comparison of some Raman spectroscopic techniques for studies of human stratum corneum *Pharm. Res.* **10** 1642–7

Williamson J M, Bowling R J and McCreery R L 1989 Near-infrared Raman-spectroscopy with a 783-nm diode-laser and CCD array detector *Appl. Spectrosc.* **43** 372–5

Wold S, Esbensen K and Geladi P 1987 Principal component analysis *Chemometrics Intell. Lab. Syst.* **2** 37–52

Wolfarth D L, Han D W, Bushar G and Parks N L 1997 Separation and characterization of polyethylene wear debris from synovial fluid and tissue samples of revised knee replacements *J. Biomed. Mater. Res.* **43** 57–61

Yamagishi M, Ishizuka Y and Seki K 1994 Pathology of olfactory mucosa in patients with Alzheimer's disease *Ann. Otol. Rhinol. Laryngol.* **103** 421–7

Yao Y, Nelson W H, Hargraves P and Zhang J 1997 UV resonance Raman study of demoic acid, a marine neurotoxic amino acid *Appl. Spectrosc.* **51** 785–91

Yazdi Y, Ramanujam N, Lotan R, Mitchell M F, Hittelman W and Richards-Kortum R 1999 Resonance Raman spectroscopy at 257 nm excitation of normal and malignant cultured breast and cervical cells *Appl. Spectrosc.* **53** 82–5

Yeo T-K, Brown L and Dvorak H F 1991 Alterations in proteoglycan synthesis common to healing wounds and tumors *Am. J. Pathol.* **138** 1437–50

Yu N T, DeNagel D C and Kuck J F R 1987 Ocular lenses *Biological Applications of Raman Spectroscopy* ed T G Spiro (New York:Wiley) pp 47–80

Yu N T, DeNagel D C, Pruett P L and Kuck J F R 1985 Disulfide bond formation in the eye lens *Proc. Natl Acad. Sci. USA* **82** 7965–8

Yu N T, Kuck J F R and Askren C C 1979 Red fluorescence in older and brunescent human lenses *Invest. Ophthalmol. Vis. Sci.* **18** 1278–80

——1982 Laser Raman spectroscopy of the lens *in situ*, measured in an anesthetized rabbit *Curr. Eye Res.* **1** 615–8