Vibrational spectroscopy has been an integral tool for the identification and characterisation of drugs. When one thinks of vibrational spectroscopy, typically infrared (IR) techniques come to mind, not Raman spectroscopy. However, over the past 20 years a renaissance of the Raman technique has occurred, mainly through instrumentation development. These developments have led to unique applications in the pharmaceutical and forensic industries in which drug identification and characterisation is necessary. This chapter presents the theory, instrumentation, sampling techniques and applications of Raman spectroscopy as applied to drugs.

Sommaire

- 1 Introduction and theory
- 2 Instrumentation
  - 2.1 Dispersive spectrometers
- 3 Interferometric spectrometers
  - 3.1 Microscopy
  - 3.2 Fibre optics
- 4 Data processing and presentation of results
- 5 System suitability tests
- 6 Sample preparation and sample presentation
- 7 Interpretation of spectra
- 8 Qualitative analysis
- 9 Quantitative analysis
  - 9.1 Chemometrics
  - 9.2 Partial least square
  - 9.3 Discriminant analysis
- 10 Collections of data

Introduction and theory

Raman spectroscopy is a form of vibrational spectroscopy that has widespread use in pharmaceutical investigations. Applications include chemical structure elucidation, routine chemical identification and solid–state characterisation, such as polymorphism. Raman spectroscopy is also applicable to drug product characterisation, including solid form analysis of the drug incorporated into the formulation, contaminant analysis, drug–excipient interaction and problem solving. A distinct advantage of pharmaceutical analysis by Raman spectroscopy is the ease of the technique and its broad range of applicability. Analysis can be performed on virtually any type of

sample, such as single crystals, bulk materials, slurries, creams, particulates, films, solutions (aqueous and organic), oils, gas–phase samples and on–process streams (the latter through the use of fibre–optic probes). Additionally, Raman spectroscopy is typically non–destructive in nature, and so the material can be recovered for further characterisation. Since Raman spectroscopy measures the vibrational motions associated with a molecule, it is complementary to IR spectroscopy as well as to other characterisation techniques. Finally, under proper sampling conditions, Raman spectroscopy is a quantitative technique.

When a compound is irradiated with monochromatic radiation, the radiation is transmitted, absorbed or scattered by the molecule. Of the scattered radiation, a majority of the photons are scattered at the same frequency as the incident radiation frequency. This form of scattering has been termed *elastic* or *Rayleigh scattering*. If the scattered radiation is passed into a spectrometer, a strong Rayleigh line is detected at the unmodified frequency of radiation used to excite the sample. Additionally, a very small proportion of photons (about one per million) are scattered at frequencies arrayed above and below the frequency of the Rayleigh line. The *differences* between the incident frequency of radiation and arrayed frequencies correspond to the frequency of molecular vibrations present in the molecules of the sample. These wavelength–shifted frequencies are termed *inelastic scattering*, and a collection of these wavelength–shifted frequencies is termed a Raman spectrum. For example, a Raman line at ± 2980 cm^{-1} may be obtained on either side of the Rayleigh line and thus the sample possesses a vibrational mode at this frequency. The frequencies of molecular vibrations are typically 10^{12} to 10^{14} Hz. A more convenient unit, which is proportional to frequency, is wavenumber (cm^{-1}) since fundamental vibrational modes lie between 3600 and 50 cm^{-1}.

As shown in Fig.1, when a molecule is irradiated with monochromatic radiation, a number of different transitions may occur. If the radiation is of sufficient energy, an absorption process may occur that represents an electronic transition ($S \rightarrow S_1$, UV/VIS spectroscopy). If a slightly less energetic source is used, the molecule is promoted to a *virtual state*. The virtual state represents a distortion of the electron distribution of a covalent bond within the molecule. After promotion to the virtual state, the molecule immediately relaxes back to the original ground electronic state by emitting a photon. If the molecule relaxes back to the original vibrational state, the emitted photon is of the same frequency as the incident radiation. This represents Rayleigh or elastic scattering. If the molecule relaxes back to a higher vibrational energy state, the emitted photon represents less energy than the incident radiation. The inelastically scattered photon has a longer wavelength (lower frequency) than the incident radiation. This energy transition represents Stokes Raman scattering. Conversely, if the molecule relaxes back to a lower vibrational energy state, the emitted photon represents greater energy than the incident radiation. The inelastically scattered photon has a shorter wavelength (higher frequency) than the incident radiation. This energy transition represents anti-Stokes Raman scattering. Generally, the anti-Stokes lines are less intense than the Stokes lines because these transitions arise from higher vibrational energy levels that contain fewer molecules, as described by the Boltzmann distribution. Hence, the Stokes portion of the spectrum is generally used.

Figure 1. Jablonski energy level diagram illustrating possible transitions. Solid lines represent absorption processes and dotted lines represent scattering processes. A, electronic transition with non–radiative decay (heat, zig–zag line) or radiative decay (fluorescence, thick line); B, Rayleigh scattering; C, Stokes Raman transition; D, anti-Stokes Raman transition. S0 is the singlet ground state, S1 the lowest singlet excited state and v represents vibrational energy levels within each electronic state.

A Raman spectrum is normally represented as a plot of Raman scattering intensity (ordinate) versus wavelength (abscissa). Normally, the abscissa of the spectrum is labelled as wavenumber shift or Raman shift (cm^{-1}) and the negative sign (for Stokes shift) is omitted (Fig.2). The wavenumber or Raman shift represents the shift in frequency of a photon from the exciting wavelength.

Figure 2. FT–Raman spectrum of paracetamol.

Pharmaceutically relevant molecules are typically covalently bound organic molecules. The chemical bonds within these molecules comprise an electron cloud. In a Raman experiment, the electromagnetic radiation incident on the molecule consists of oscillating electric and magnetic fields. Interaction of the electromagnetic radiation with the chemical bond causes the electron cloud to oscillate. The oscillation, in turn, causes a photon to be emitted, which is called scattering. In Rayleigh scattering, the energy from the incident electromagnetic radiation causes electron oscillation and the emitted photon is observed at the same frequency. In Raman scattering, an additional energy transition occurs. The polarisability of the electron cloud may change as the position of the atoms that make up the chemical bond change. In other words, a vibrational mode may cause atoms to be displaced, which in turn affects the polarisability of the electron cloud of the chemical bond. In this case, the incident radiation causes the electron cloud to oscillate, but the electron cloud oscillation is also affected by the change in polarisability caused by a change in the position of the atoms during the molecular vibration. Oscillation causes a photon to be emitted, but the frequency is perturbed by the change in polarisability of the chemical bond. For a vibrational mode to be Raman active, a change in polarisability must take place during the vibration. This is termed the Raman selection rule. Since the polarisability of a chemical bond is dependent upon the atoms making up that bond, as well as atoms in close proximity, Raman spectroscopy is a probe into the chemical or physical structure of a molecule. This is why Raman spectroscopy is an important tool for pharmaceutical analysis.

Instrumentation

Although the Raman effect was discovered in 1928, the first commercial Raman instruments did not start to appear until the early 1950s. These instruments did not use laser sources, but used elemental sources and arc lamps. In 1962, laser sources started to become available for Raman instruments and the first commercial laser Raman instruments appeared in 1964 to 1965. The first commercial Fourier transform (FT)–Raman instruments were available from 1988, and by 1989 FT–Raman microscopy was possible.

Dispersive spectrometers

The basic configuration and components of a dispersive spectrometer are shown in Fig 3. The source of monochromatic radiation is a laser. Typically, helium–cadmium (325, 354 or 442 nm), air–cooled argon–ion (488 or 514 nm), doubled continuous–wave neodymium–yttrium aluminium garnet (Nd:YAG or Nd:$Y_3Al_5O_{12}$) (532 nm), helium–neon (633 nm), or stabilised diode (785 nm) lasers are used for dispersive Raman spectrometers. The stability of the emitted radiation from a laser is one of the key attributes of a good spectrometer. Frequency stabilisation of the laser under standard laboratory conditions (slight temperature fluctuations, vibrational effects, etc.) is required. Laser lifetimes and cost are also considerations in the choice of laser to use. One additional consideration for laser selection in dispersive Raman systems is the generation of fluorescence. As stated previously, the Raman signal is fairly weak. For many organic systems, fluorescence may occur depending upon the laser used and, instead of promoting the molecule of interest to a virtual state, an electronic transition occurs with subsequent radiative decay (fluorescence, Fig 1.). The fluorescence background signal can be so intense as to mask the Raman scattered photons. Fluorescence is wavelength dependent, so a sample that fluoresces with one laser source may not with another. If fluorescence does not pose a problem, lower frequency lasers can be used (532, 514 nm) for enhanced sensitivity, as the efficiency of Raman scatter is proportional to $1/(\lambda)^4$, where λ is the wavelength. If fluorescence is a problem when using these high–energy sources, then lower energy sources, such as those used in FT-based Raman spectroscopy, can be used to minimise the fluorescence affects (*vide infra*).

In a dispersive Raman spectrometer, the sample is positioned in the laser beam and the scattering radiation collected in either a 180° backscattering or a 90° right–angle scattering configuration. Subsequently, a laser–line rejection filter is put in place to filter out the Rayleigh scattering. Finally, a detector is positioned in the spectrometer. For dispersive systems, typically a charge–coupled device (CCD) is utilised. Silicon CCD detectors are normally used for Raman spectrometers in which visible wavelength lasers are used. Previously, photomultiplier tubes (PMTs) were used for detection, but since the advent of CCDs and their inherently better performance, PMTs are no longer normally used. Unfortunately, CCD detectors are sensitive to cosmic rays which add artefacts to the spectra. These artefacts can be removed via software methods. All commercial spectrometers are controlled digitally by a computer system.

Interferometric spectrometers

Fig 4 displays the configuration of a FT-based Raman spectrometer. Advantages of a FT–Raman spectrometer are wavelength accuracy and the use of a near-IR laser, which typically eliminates fluorescence. In the FT-based system, a neodymium–yttrium orthovanadate ($Nd:YVO_4$) laser (1064 nm) is used to irradiate the sample. Analogous to the dispersive system, the sample is positioned in the laser beam and the scattering radiation collected in either a 180° backscattering or a 90° right–angle scattering configuration. The scattered photons are then passed into an interferometer with laser line filtering. Detectors of the scattered photons from systems that utilise lasers emitting light with wavelengths greater than 1000 nm are of the single–element type, either high purity p–type germanium (Ge) or indium–gallium–arsenic (InGaAs) detectors. These detectors are noisier than CCDs or PMTs, but do exhibit high quantum efficiencies. By cooling the Ge detector to 77° K, the frequency response is extended to 3400 cm^{-1}. Unfortunately, Ge detectors are subject to interference by cosmic rays so artefacts in their output may be generated. However the number of interferograms containing artefacts from cosmic rays can be reduced by utilising software methods.

By utilising the longer wavelength $Nd:YVO_4$ laser in a FT–Raman spectrometer, fluorescence is minimised. The reason for fluorescence minimisation is that the longer wavelength excitation poses a smaller chance of inducing an electronic transition with subsequent fluorescence as the relaxation mechanism.

Microscopy

Raman spectra can be acquired on small amounts of material through the use of a Raman microprobe. Utilising the microscope, the Raman scattered photons are collected in a 180° backscattering configuration that allows the operator to view the sample optically, focus the incident radiation and subsequently collect the Raman spectrum. Most commercial Raman microscope systems utilise confocal microscopy to increase axial resolution (*z*-axis). Confocal points are defined as the point source, the in–focus sample location and the focused image of the sample point. Axial resolution, defined as the distance away from the focal plane in which the Raman intensity from the sample decreases to 50% of the in–focus intensity, can be approximated from the numerical aperture (NA) used in the microscope. When utilising a 0.95 NA objective on a confocal microscope system, the axial resolution is proportional to the square of the NA, in this case 0.9025 µm.

An additional advantage of Raman microscopy is spatial resolution as opposed to axial resolution. The spatial resolution (*xy*-plane) is dependent upon the NA of the collecting objective and the wavelength of the laser radiation. Larger NA values and shorter wavelengths provide higher spatial resolution, often down to 1 µm. Since a high intensity of monochromatic radiation from the laser is focused upon a small amount of sample, sample degradation by the laser must be monitored. Otherwise, the Raman microprobe is ideal for investigating polymorphism (single crystals), particulate contamination and small amounts of samples. Using an apparatus

similar to those used for IR microspectroscopy, variable temperature studies can be performed with a Raman microprobe.

Fibre optics

Fibre optics have been used in Raman spectroscopy since the early 1980s. Solids and liquids can be analysed with an arrangement of optical fibres on the end of a probe. Today, much of the research in the use of fibre optics in FT–Raman spectroscopy centres around fibre and fibre–bundle design. The number, type and arrangement of the fibres in a fibre bundle are all factors that are varied to produce fibre bundles for different applications. Fibre systems include single fibre (in which the laser excitation and collected scattered radiation travel along the same fibre), and multifibre [in which laser excitation is transmitted along one (or multiple) fibres, and the scattered radiation is transmitted to the detector along different fibres]. The arrangements of the fibres in a multifibre system can also vary. Two examples are an arrangement in which one excitation fibre is surrounded by several collection fibres or one in which several excitation and collection fibres are randomly mixed in a bundle. The greatest single advantage of the use of fibre optics in Raman spectroscopy is the ability to sample remotely. No longer does the spectrometer have to be brought to the sample or *vice versa*. Fibre optics can link the spectrometer to the sample, typically in distances of tens of metres. Common applications include monitoring process streams or hazardous reactions.

Data processing and presentation of results

All modern–day dispersive and FT-based Raman spectrometers are controlled through a digital computer that handles instrument control, data collection, data processing and presentation of the spectral results. Utilisation of a FT–Raman spectrometer requires one extra step of data processing that is not needed for dispersive–based Raman spectrometers. In the FT system, the original data are collected on a time scale and subsequently Fourier transformed to obtain a frequency domain spectrum. After this point, data processing for both systems is analogous.

Since the Raman spectrum of a particular sample is represented digitally, various additional processing applications can be performed. Spectral subtraction is one commonly used data–processing technique, as are spectral smoothing, spectral searching and resolution enhancement. Fig. 5 displays some of these processing techniques on one original data file. Since Raman spectroscopy can be utilised for quantitative measurements, digital representation of the spectrum allows numerous ways to measure the analytical response of an analyte’s signal and relate it to concentration. Electronic integration of the peak area, curve fitting of the Raman spectrum and chemometric approaches can be utilised (see Fig. 5).

Figure 5. Examples of digitally processed spectral files. A, original Raman spectrum; B, spectral result after calculating the second derivative of the original Raman spectrum; C, spectral result after resolution enhancement by Fourier deconvolution of the original Raman spectrum; D, spectral result after Fourier smoothing the original Raman spectrum by 80%.

System suitability tests

In a current good manufacturing practices environment, all laboratory instrumentation must be inspected, cleaned and maintained adequately. Additionally, instruments used for the generation, measurement or assessment of data must be tested, calibrated and/or standardised adequately. Based upon these regulatory agency requirements, the accuracy of Raman spectrometers needs to be assessed before usage. Typically, regulatory agencies require the use of a recognised standard material, distributed by an official agency, for the calibration testing of any instrument. The American Society for Testing and Materials (ASTM) has published a guide for Raman shift standards for spectrometer calibration (ASTM 1996).The spectroscopy community has recognised cyclohexane and sulfur as calibration standards for the frequency scale (cm^{-1}) of Raman spectrometers. The Raman spectrum of sulfur displays vibration bands at 153, 219 and 472 cm^{-1}, whereas the spectrum of cyclohexane displays bands at 802, 1267, 1444, 2853, 2924 and 2938 cm^{-1} (Fig .6). In many pharmaceutical spectroscopy laboratories, the spectral features of these two compounds are used to calibrate Raman instruments and determine system suitability. Also, analogous to the lack of an officially distributed frequency calibration standard for Raman spectroscopy, an intensity standard does not exist (at the date of publication of this book). For dispersive spectrometers, atomic emission lamps can be used for wavelength verification.

Figure 6. FT–Raman spectra of wavelength calibration standards cyclohexane and sulfur.

Sample preparation and sample presentation

Sampling techniques for Raman spectroscopy are relatively simple, since the only requirements are that the monochromatic laser beam irradiates the sample of interest and the scattered radiation is focused upon the detector.

Raman spectroscopy may be performed on very small samples (e.g. a few nanograms). Powders do not need pressing into disks or diluting with KBr, as in IR spectroscopy; the material just needs to be irradiated by the laser beam. Solid samples are often examined in gold-coated or glass sample holders that generally require about 25 to

50 mg of material. Typically, liquid samples are analysed in quartz or glass cuvettes, which may have mirrored rear surfaces to improve the signal intensity. Glass is a very weak Raman scatterer and so many samples (liquid and solid) simply can be analysed in a bottle or, for example, in a nuclear magnetic resonance (NMR) tube, although fluorescence from some glasses can be problematic. Water is a good solvent for Raman studies, since the Raman spectrum of water is essentially one broad, weak band at 3500 cm^{-1}. One sample type that may pose a problem is darkly coloured material. Often, these samples absorb excessive heat and burn, which causes sample and spectral degradation. Amorphous materials also have a tendency to absorb heat in the laser beam. To avoid sample burning, it may be necessary to dissipate the heat. This can be accomplished by reducing the laser power or by using an accessory that spins the sample, and so avoids irradiation of a single point in the sample. Further reduction of the laser power can also be accomplished through a neutral density filter. Sometimes, the sample can be diluted in KBr to help reduce sample burning.

The complete Stokes Raman spectrum, which covers shifts in the range 100 to 3500 cm^{-1}, can be obtained and the intensity of Raman scattering is directly proportional to the concentration of the scattering species, an important factor for quantitative analysis. However, the Raman effect is relatively weak and hence the material needs to be present at a concentration of at least about 0.1–0.5% for accurate assessments, whereas IR can be used to detect materials down to a concentration of approximately 0.01%. Fluorescence can also be problematic in Raman studies, but typically arises from additives in the glass sample tubes or impurities within the sample of interest. Data massaging techniques can sometimes blank out Raman spectral contributions from fluorescent materials. Photobleaching is another way to suppress fluorescence. This technique involves irradiating the sample for a prolonged period of time (seconds to hours) with the laser. During this time the fluorescence may decrease through the destruction of the fluorescing component from the prolonged exposure to the laser irradiation. The spectrum is then acquired after photobleaching is complete. Of course, it is also possible that the laser radiation may change the component of interest. For example, a solvated crystalline material may desolvate with exposure to laser radiation. It is wise to establish sample integrity by comparing spectra acquired with short acquisition times to those acquired with long acquisition times. Additionally, another technique, such as IR spectroscopy or X-ray powder diffraction, could be used to check for sample integrity after Raman analysis.

Variable temperature studies in Raman spectroscopy provide a wealth of information. A Raman spectrum typically covers a wavelength range that extends beyond the range normally associated with mid-IR spectroscopy (typically 4000 to 400 cm^{-1}), so information about the lattice vibrations of compounds is readily available. By varying the temperature of a sample, the lattice energies of the compound can be changed, which allows the nature of the crystal lattice to be interpreted. In addition, information similar to that obtained in IR variable–temperature studies (crystal form changes and the nature of solvate association) can be obtained with variable–temperature Raman investigations.

Interpretation of spectra

The number of fundamental vibrational modes of a molecule is equal to the number of degrees of vibrational freedom. For a nonlinear molecule of *n* atoms, 3*n* – 6 degrees of vibrational freedom exist. Hence, there are 3*n* – 6 fundamental vibrational modes. Six degrees of freedom are subtracted from a nonlinear molecule since:

- three co–ordinates are required to locate the molecule in space
- an additional three co–ordinates are required to describe the orientation of the molecule based upon the three co–ordinates that define the position of the molecule in space.

For a linear molecule, 3*n* – 5 fundamental vibrational modes are possible, since only two degrees of rotational freedom exist. Thus, in a total vibrational analysis of a molecule by complementary IR and Raman techniques, 3*n* – 6 or 3*n* – 5 vibrational frequencies should be observed. In complex molecules (such as pharmaceuticals), it may not be possible to observe every vibration because of overlap and relative intensity differences.

Regarding spectral interpretation, two of the best textbook sources are *Introduction to Infrared and Raman Spectroscopy* (Colthup et al. 1990) and *The Handbook of Infrared and Raman Characteristic Frequencies of Organic Molecules* (Lin-Vien et al. 1991). Both of these texts contain chapters dedicated to functional groups (e.g. methyl and methylene groups) and chemical compound types (e.g. ethers, alcohols and phenols). By examining peak location and intensity in Raman spectra, functional groups (and chemical compound type) can be determined. Discussed below are some guidelines as to how to interpret a Raman spectrum, but clearly this section serves as general background information only. Any complete spectral interpretations should be made with the guidance of one of the reference texts.

Raman spectroscopy provides information about the molecular bonding of a molecule. Certain functional groups give rise to fundamental vibrational modes. For example, the C=C stretch in ethylenes occurs near 1650 cm^{-1}. The C–N stretch (amide III peak) for primary amides occurs between 1430 and 1390 cm^{-1}. A C–S stretch usually appears strongly in the 735 to 590 cm^{-1} region. Many other functional groups give rise to Raman peaks in specific spectral regions. IR bands also occur in the same spectral regions, but because of selection rules, the band intensities differ, often dramatically. In the most basic terms, a vibrational mode is Raman active when there is a change in polarisability during the vibration. Conversely, a vibrational mode is IR active when there is a change in the molecular dipole moment during the vibration. Hence, vibrational modes that give rise to strong Raman peaks often give weak IR bands and *vice versa*. It is this characteristic that gives the description ‘complementary’ to using IR and Raman spectroscopy together to examine the molecular bonds in a compound. Some of the strongest Raman peaks come from functional groups such as C=C, N=N, S–S, C–H, S–H, C=N, C=S and C–S, which have low polarity and high polarisability. These functional groups tend to occur in rather constant frequency ranges, although some shifts can occur based on other substituents in the molecule.

Vibrational modes can be separated into two classes: those in which the molecular bond is stretching and those in which it is bending. It takes a specific amount of energy for these actions to occur. The amount required is dependant on the atoms involved and the strength of the bond. In the theory section of this chapter, the nature of Raman excitation of a molecule is discussed. The location of the vibrational mode (its peak wavelength position in the spectrum) is related to the frequency of the excitation source (the laser) and the frequency of the scattered light. The equation that defines this relationship is $h\nu = h\nu' + \Delta E_{vib}$, where h is Plank's constant, ν is the excitation frequency, ν' is the scattered light frequency and ΔE_{vib} is the vibrational energy. The vibrational energy is related directly to the strength of the bond and the amount of energy required to make that bond stretch or bend. These actions are elaborated below.

A stretching vibration is the motion a molecular bond undergoes when the two atoms involved in the bond move apart and then contract. The stretch can be a simple contraction–expansion between two atoms, such as the C=C stretch of ethylene. This type of motion occurs as a peak in the Raman spectrum at approximately 1650 cm^{-1}. When three atoms are involved, two types of stretches are possible, symmetric and antisymmetric. Using a methylene group as an example, a symmetric CH_2 stretch occurs when the two hydrogen atoms move apart from the carbon atom at the same time. This type of stretch appears near 2853 cm^{-1}. The antisymmetric CH_2 stretch occurs when one hydrogen atom moves away from the carbon atom while the other hydrogen atom moves closer to the carbon atom. This vibrational mode appears near 2926 cm^{-1}. The symmetric stretches are typically more intense in Raman spectra (greater change in polarisability), whereas the antisymmetric stretches are more intense in IR spectra. When many atoms are involved in the stretching, such as with aromatic rings, the types of stretches become more complicated. For example, a benzene ring monosubstituted with a halogen has 30 vibrational modes. Of these modes, some are strictly stretches, some are bends and some are combinations of both. Perhaps the simplest stretching mode for benzene rings is the ring–breathing mode, in which the 2, 4 and 6 carbons move outwards. This mode occurs as a very strong Raman peak near 1000 cm^{-1} for mono-, *meta*– and 1,3,5–trisubstituted benzenes. Clearly, many types of stretches can occur.

In a bending vibration, the molecular bond bends instead of stretches. Many types of bends can occur: antisymmetric, symmetric, rock, torsion, scissor, wag, twist and rock. Using a CH_3 group as an example, antisymmetric, symmetric, rock and torsion bends are possible (Fig 7). In the antisymmetric bend, two of the CH bonds bend towards each other in a pinching motion, while the third bends outwards and away from the pyramid. This mode occurs between 1470 and 1430 cm^{-1}. In the symmetric bend, the three CH bonds all bend inwards, similar to a grasping type of action. This mode occurs from 1395 to 1365 cm^{-1}. In the CH_3 rock, the three CH bonds all bend in one direction, in a sweeping type of mode. Finally, the CH_3 torsion involves the three CH bonds all bending in a clockwise direction in a twisting motion. The spectral ranges for these last two vibrational modes are highly variable.

Figure 7. Bending vibrational motions associated with CH3 ((+) represents movement above the plane, (-) represents movement below the plane).

Figure 8. Bending vibrational motions associated with CH2 ((+) represents movement above the plane, (-) represents movement below the plane).

The CH_2 group can bend in slightly different ways (Fig 8). A CH_2 scissor bend occurs when the two hydrogen atoms move towards each other in a scissoring motion. For a CH_2 wag, the two hydrogen atoms are bent towards the carbon atom. If the two hydrogen atoms alternately twist around one another, it is a CH_2 twist. Finally, if the two hydrogen atoms move back and forth, in line with the carbon atom, it is a CH_2 rock. As seen with the CH_3 modes, all of these different types of motions appear in different portions of the spectrum because they all require a different and unique amount of energy to occur. The scissor mode appears near 1465 cm^{-1}. The rocking, wagging and twisting modes are more complicated, falling in the range 1422 to 719 cm^{-1}. Depending on the substituents in the molecule, these modes can each be narrowed down to tighter ranges. It is this dependence of certain modes on substituents that gives rise to correlation tables, discussed next.

Although the general location for a certain type of mode, for example a C=C stretch, can be listed, the exact location of the peak varies slightly with the type of molecular substitution present. For example, a vinyl C=C stretch (monoalkyl) occurs from 1650 to 1638 cm^{-1}. A vinylidine C=C stretch (1,1–dialkyl) occurs from 1660 to 1640 cm^{-1}. A *cis*-dialkyl–substituted C=C stretch occurs from 1662 to 1631 cm^{-1}, whereas a *trans*-dialkyl–substituted C=C stretch occurs from 1676 to 1665 cm^{-1}. Finally, trialkyl- and tetraalkyl–substituted C=C stretches appear from 1680 to 1665 cm^{-1}. Some modes are very substituent sensitive, whereas others appear relatively consistently near a certain wavenumber.

Over the years, many types of compounds have been studied in great detail. When a spectrum of a single compound is interpreted such that every peak or band is assigned to a type of motion, we have a vibrational assignment. Through the examination of many vibrational assignments, it has been possible to draw correlations between types of vibrational modes and types of substituents. The culmination of this work is detailed in correlation tables that an investigator can use to aid in spectral interpretation (Dollish et al. 1974; Colthup et al. 1990; Lin-Vien et al. 1991).

The most efficient way to use correlation tables is to look for certain peaks that will quickly narrow down the type

of compound. For example, if peaks are found at 3000 cm^{-1} or slightly higher, the compound is aromatic or olefinic. However, if the CH_2 stretches appear below 3000 cm^{-1}, the compound is aliphatic. Once that determination is made, the correlation tables can be consulted to look for other confirmatory modes, such as CH wags for aromatic compounds. This type of process is continued until a compound class can either be verified or discounted.

For complex molecules, the structure may not be able to be determined from the Raman spectrum alone, or even if the IR spectrum is available as well. Other techniques such as NMR and mass spectrometry provide important information to aid the structure elucidation.

Qualitative analysis

Chemical identity testing of compounds is one role of the pharmaceutical spectroscopy laboratory. Testing can be accomplished with methods that utilise Raman spectroscopy. A FT–Raman method has been developed to identify the two active components (tegafur and uracil) in formulated capsules (Petty et al. 1996). The Raman spectrum of the formulated product displayed a spectral region in which Raman bands unique to uracil and tegafur were observed. The presence of these bands allowed the analyst to confirm that both components were present in the formulated product. In part because of the ease of use and chemical specificity inherent with Raman spectroscopy, it has become an essential chemical and physical identification tool for the pharmaceutical spectroscopist.

Figure 9. FT –Raman spectra of the polymorphs of carbamazepine (top, polymorph I, bottom, polymorph III).

Polymorphism (the ability of a molecule to crystallise in different three–dimensional structures) is a very important aspect of the drug–development process. Raman spectroscopy is now being used for the qualitative and quantitative characterisation of polymorphic compounds of pharmaceutical interest. For the sake of brevity, the term polymorphs encompasses polymorphs, pseudopolymorphs, hydrates and solvates throughout the remainder of this chapter. The term polymorphs is more commonly used and is defined as the collection of different crystal structures that can exist for a chemical entity. The terms hydrates and solvates, sometimes referred to as pseudopolymorphs, refer to the collection of different crystal structures that result from various degrees of solvation. These materials differ chemically only by the degree of solvent incorporation (monohydrate, dihydrate) into the crystalline lattice, but physically display different crystal structures. Since solid–state vibrational spectroscopy (Raman and IR) can be used to probe the nature of polymorphism at the molecular level, these methods are particularly useful in instances where full crystallographic characterisation of polymorphism may not be possible. Figure 9, displays the Raman spectra of different polymorphs of carbamazepine. An added advantage of using Raman spectroscopy for polymorphic investigations is the ease of the technique and the ability to

measure low frequency vibrations (500 to 50 cm^{-1}) attributed to lattice modes. In many cases, two different solid–state forms of a pharmaceutical entity display spectral differences in the low–frequency region of the Raman spectrum. These spectral differences are also noted in the spectral comparison of crystalline versus amorphous material, a significant advantage for the amorphous–crystalline characterisation of the drug. In addition, no sample preparation is required, which provides a significant advantage over other techniques.

Raman microspectroscopy is well suited for in–situ analysis of contaminants found in pharmaceutical processes. The non–destructive nature of the analysis means that further experiments, such as energy dispersive X-ray analysis or IR microspectroscopy, may be performed on the same sample. A consideration for contaminant analysis by Raman spectroscopy is the axial and spatial resolution of the technique as compared to that of IR microspectroscopy. In general, IR microspectroscopy is diffraction limited to investigating samples typically larger in size than 5 µm. As discussed earlier, with a 0.95 NA objective, 1 µm spatial resolution can be achieved with a Raman microscope, which enables the analysis of very small amounts of contaminant.

Real–time monitoring of pharmaceutically relevant processes is an exciting new application for Raman spectroscopy. Real–time monitoring by Raman spectroscopy has been utilised to examine synthetic organic reaction schemes to investigate kinetics, as well as to identify non–isolated reaction intermediates. Distinct advantages for Raman spectroscopy in this area are:

- ability to work with aqueous–based systems with little spectral interference from water
- utilisation of a fibre–optic probe for direct and/or remote sampling
- collection of the Raman spectrum directly through the glass vessel with little or no spectral interference
- ability to analyse the spectrum quantitatively.

Other recent applications of real–time process monitoring by Raman spectroscopy include polymorphic interconversion under slurry conditions and crystallisation monitoring (Findlay and Bugay 1998).

Another exciting new application of Raman spectroscopy is chemical imaging. By incorporating a programmable, *xyz*-movement stage into a Raman microscope, it is now possible to generate a chemical image of a two–dimensional area of a sample. By using a mapping stage, a sample such as a microtomed tablet can be moved in the *x* and *y* directions, obtaining spectra at each step. If the sample requires re–focussing at different locations, the *z* direction can be automated as well. The distance the stage moves in the *x* or *y* direction is called the step size and usually can be as small as 1 µm.

There are three types of maps: point, line and area (Fig 10). A point map provides several different areas of a sample to be analysed consecutively, but the spectra are not related to each other spatially. A point map can be considered as a type of auto–sampler. An example of a point–map application is a 96–well plate that contains

combinatorial beads in each well. The second map type (the line map) defines a series of spectra obtained along one dimension. In line maps, chemical changes that occur along this dimension are investigated. An example of a line–map sample is a cross–sectioned pharmaceutical beadlet that has several different layers exposed. One of the more practical reasons why line maps are so popular is that they can provide detailed information regarding the chemical changes that occur across a sample without the need to collect as many spectra as is required for area maps. The final map type (the area map) defines a series of spectra collected in two dimensions (i.e. over an entire region). This type of map provides a Raman image that can be compared directly to the visual image, and often allows non–visible (to the eye) features to be identified. A common area–map sample is a tablet. The tablet can be mapped and the various ingredients can be monitored as to content uniformity. A series of images that demonstrate the dispersion of several excipients in a tablet is displayed in Fig 11. This particular area map was obtained with a step size of 4 µm, a sampling spot size of approximately 1 µm and a total sampling area of 87 µm × 52 µm. The time required to collect this map was approximately 26 h, but not all Raman mapping experiments require this amount of time. For example, if a map is to search for a particular component, the step size need only be on the order of the particle size of that substance. A good approach is to first obtain a larger area map on the sample with larger step sizes and shorter sampling times per point. Once an area of interest is defined by analysing the data from the first map, a smaller, higher resolution map can be defined. Also, area maps and imaging have experienced a recent surge in popularity, mainly because of the use of CCD array cameras and liquid crystal tuneable filter (LCTF) technology, respectively. These detectors greatly reduce the amount of time required to collect an area map (Zugates and Treado 1999). In one case, the time required to collect an area map decreased by approximately 4 h upon switching from a CCD camera (>5 h experiment) to LCTF technology (<1 h).

Figure 10. Examples of defined point, line and area maps.

Figure 11. Peak area profiles (images) representing: A, mannitol; B, aspartame; C, cellulose; D, magnesium stearate; E, corn starch; F, monoammonium glycyrrhizinate.

Once line- or area–mapping experiments have been performed, profiles are created that enable certain spectral features to be monitored spatially on the sample. The data sets generated by mapping can consist of hundreds or even thousands of spectra, so to view all these spectra concurrently is unrealistic. Profiles can aid in this task by reducing the data set into a more easily viewed format (an image). A profile is a representation of map data in which a measurement of spectral intensity or some other characteristic is shown for each sample point. A profile verifies the presence, location and extent of a defined spectroscopic feature in the sample. There are many types of profiles. Some profiles, such as chemigram or component profiles, compare an entire reference spectrum to every spectrum in the map, with the resultant image showing spectral similarity across the mapping region. Other images can be created based on profiling a specific peak area, peak height, peak–area ratio or peak–height ratio. Profiles can also be performed based on a group of peaks specific for certain functional groups (e.g. alkanes) or

even on a quantitative method.

Literature references of other mapping applications relative to the pharmaceutical industry include some of the following: solid dispersions of ibuprofen in polyvinylpyrrolidone (Breitenback et al. 1999), crystal formation in hormone replacement therapy patches (Armstrong et al. 1996), particle size analysis in mixtures (Theophilus and Lancaster 2000) and pharmaceutical matrix determination of dosage formulations (Clarke et al. 2000).

Quantitative analysis

The ability to perform quantitative analysis by Raman spectroscopy is a significant advantage of the technique. Through mathematical treatments, it has been shown that the Raman scattering intensity is proportional to the number of molecules being irradiated. The intensity of scattered radiation is also proportional to the intensity of the incident radiation and the fourth power of the difference in frequencies between the laser frequency and the molecular vibrational frequency. Thus, increased Raman scattering intensity, and potentially lower limits of detection, can be achieved by increasing the intensity of the laser radiation and/or increasing the frequency of the laser irradiation. This quantitative relationship between Raman scattered intensity and concentration can be expressed as:

where I_R is the measured Raman intensity (photons per second), I_L is the laser intensity (photons per second), σ is the absolute Raman cross–section (cm^2 per molecule), K is the measurement parameters, P is the sample path length (cm) and C is the concentration (molecules per cm^3). The constant K represents measurement parameters such as utilising the same spectrometer (collection optics efficiency), sample positioning and overall efficiency of the Raman spectrometer.

In the past, quantitative analysis was not often performed with Raman spectroscopy because of some problems inherent to dispersive systems. When FT–Raman spectrometers became popular, the feasibility of quantitative applications greatly improved (Walder and Smith 1991). Some of the FT advantages that lend themselves well to quantitative applications are Jacquinot's advantage, Fellgett's advantage and Connes' advantage. Jacquinot's advantage, also known as the throughput advantage, depends on the large entrance aperture of the interferometer, which makes the sampling geometry less sensitive to absolute repositioning of the sample cell. Fellgett's advantage, otherwise known as the multiplex advantage, concerns the increased speed at which spectra can be collected, since all wavelengths of the spectrum are collected simultaneously. This rapid collection implies that FT–Raman spectroscopy may be less sensitive to instrument drift. Finally, Connes' advantage, also called the

precision advantage, arises because there is only one moving part in a FT system (one mirror in the interferometer). A FT system is therefore very rugged, which allows excellent day–to–day wavelength precision.

Although FT–Raman spectroscopy is more applicable to quantitative applications than dispersive Raman spectroscopy, some issues are still of concern. These are mainly with the optimisation of sampling conditions. FT–Raman spectroscopy typically samples a relatively small area (e.g. a 1 to 2 mm spot) of the total sample. As such, it is important that any spectra collected be truly representative of the bulk sample. For solution studies, homogeneity of the multicomponent samples presented for quantitative analysis is not an issue. Contrarily, solid–phase analysis can present significant inhomogeneity issues for quantitative analysis. One way to improve sample uniformity is to prepare all solid mixtures by geometric mixing. Slurry mixing is another good technique if the sample can withstand contact with the slurry solvent. For spectral acquisition, sample–spinning accessories should be used to spin the sample cup that contains the mixture during spectral acquisition. Using this type of accessory results in the collection of a more representative sample spectrum (a ring of data is collected, instead of just at a single point). To minimise effectively any sample heterogeneity concerns, different rings of data can be collected for each sample cup. The individual spectra from each sampled area can then be co–averaged into groups to create representative spectra for the sample in the cup.

Another factor to consider when using Raman spectroscopy quantitatively is that the power output of the Raman laser can vary from day to day, and thereby affect the intensity of spectral peaks. It is advisable to normalise any data before using them in a quantitative manner. One possible method for normalising spectra is to ratio the spectral response of the analyte against a peak response for a non–changing component (e.g. an excipient in a drug product or an internal standard). Alternatively, a ratio can be measured using a peak response of a component that changes in the opposite direction to that of the component being monitored. For example, in a quantitative method used to monitor the amount of crystalline drug substance in the presence of the amorphous form, a crystalline peak response (e.g. peak area or peak height) can be normalised by dividing it by an amorphous peak response.

When developing a quantitative method, it is very important to follow regulatory agency guidelines. Important quantitative issues include:

- system suitability – an overall test of system function
- specificity – the ability of Raman spectroscopy to differentiate the analyte from the matrix
- working range – the concentration range over which the method is validated
- linearity – demonstration of a direct relationship between a measured analytical response and concentration over the working range of the method
- precision – the repeatability with which a number can be represented
- accuracy – degree of conformity of a measurement to a standard or true value
- limit of detection – lowest concentration at which an analyte can be detected
- minimum quantifiable limit – lowest concentration at which an analyte can be quantified with acceptable accuracy and precision
- robustness – demonstration of the reliability of an analysis with respect to deliberate variations in method parameters.

One literature example of quantitative Raman analysis addresses the question of amorphous versus crystalline content of indomethacin samples (Taylor and Zografi 1998). The article highlights the quantitative nature of Raman spectroscopy, the need to produce homogeneous calibration and/or validation samples and difficulties associated with collecting a Raman spectrum that is truly representative of the concentration. A linear correlation curve was constructed in which low concentrations of both amorphous and crystalline material could be detected and predicted in mixtures. The authors felt that the largest source of error in the measurements arose from inhomogeneous mixing of the amorphous and crystalline components in the blends. For solid–state analysis, this conclusion illustrates the need for a sampling device that collects a truly representative Raman spectrum of the sample, in this case a mixture. A recent paper in Applied Spectroscopy is an excellent resource addressing quantitative Raman spectroscopy, and the reader is referred to this for additional information (Pelletier 2003).

Chemometrics

In the above literature example of quantitative analysis, a univariate approach (a single–peak response) was used to create the calibration curve or predictive model. Chemometrics represents a multivariate approach to creating a predictive model for quantitative analysis. Chemometrics may be defined as the ‘use of statistical and mathematical techniques to make either quantitative or qualitative measurements on chemical data.’

A chemometric approach is useful when there are very few spectroscopic differences between the compounds in a mixture, which is often the case for polymorphic studies. Partial least squares (PLS) and discriminant analysis are examples of two types of chemometric approach that can be taken for quantitative and semi–quantitative work. These methods are able to identify one or more regions of complex overlapping spectra and monitor their differences by looking at peak position, peak shape, first derivative and second derivative spectra. They represent

the differences between the calibration spectra as linear combinations (factors). These factors account for the differences between the spectra; the first factor represents the strongest differences and subsequent factors represent more subtle differences. Eventually, a point is reached at which the factor simply represents changes in random noise between the spectra.

Partial least square

In the article ‘Quantitative analysis of vitamin A using FT Raman spectroscopy’, the authors used PLS to obtain a method quantified to a concentration of 0.005% w/w (Hancewicz and Petty 1995). One of the concerns they addressed is that Raman spectra need to be standardised to account for changes in the overall intensity of the spectrum. Possible sources of errors include changes in sample positioning, laser power fluctuations and sample wall–thickness variations. As mentioned above, performing peak response ratios with an internal or external standard could be used to normalise the data, but these authors took another approach. They used multiplicative scatter correction (MSC) to compensate for the random multiplicative and additive contribution to the spectra caused by sampling. MSC is traditionally used for diffuse reflectance, but can be used on any data collected with unknown or changing path lengths.

Discriminant analysis

Discriminant analysis is a chemometric technique that allows the statement to be made that an unknown sample most resembles a certain class, where each class represents sample ‘X’, ‘Y’ or ‘Z’. For example, ‘X’ could be 0% crystalline material in amorphous material, ‘Y’ could be 5% crystalline material and ‘Z’ could be 15% crystalline material. In this example, a discriminant analysis method is a pattern–recognition model that can be semi–quantitative.

There are three main parts to developing a discriminant analysis method. First, decisions must be made regarding the number of spectra (or standards) needed for each class. Once all the spectra have been collected, some are used to create a calibration (training) model. This training model is ‘tweaked’ with regard to the spectral regions examined, the number of factors (or principal component scores) used, baseline treatments, etc. Once the training model is defined, such that it correctly predicts all of the spectra within it (i.e. it is calibrated), the model is tested by internal cross–validation. This validation consists of omitting one standard at a time from the data set, using the depleted set to obtain a new calibration model and applying this new model to the omitted standard. This process is repeated for each standard. If each spectrum is predicted accurately, then the method is validated. The internal cross–validation also helps to determine if the model is over–fitted.

Overall, chemometrics is an excellent approach to quantitative method development by enhancing method

accuracy and precision, and by enabling lower limits of detection and quantification. The reader is referred to Further reading for further publications regarding chemometrics.

Collections of data

By far the quickest method to identify unknown materials is to search the spectrum against spectral libraries. Often an answer can be found within seconds. However, care must be taken when performing spectral searching.

Computer search programs determine the difference between a sample spectrum and the reference spectra in a library. Several possible algorithms are used to compare spectra, some based on intensity and others based on peak position. For example, difference algorithms accentuate peak intensity over location. The *absolute difference* algorithm puts more weight on the small differences between the unknown spectrum and library spectra, which means that impurities will have a larger effect on the search results. The *squared difference* algorithm emphasises the large peaks in the unknown spectrum, so it is useful when searching a noisy spectrum. *Derivative* algorithms accentuate peak position over intensity. The *absolute derivative* algorithm gives small peaks and peak shifts an increased effect on the search results. This algorithm removes any differences between the unknown and library spectra caused by an offset in the unknown spectrum and so is useful for the analysis of spectra with distorted baselines. The *squared derivative* algorithm emphasises large peaks as well as peak shape and works best for noisy spectra with distorted baselines. Perhaps the most versatile search algorithm is *correlation*, which balances the contributions of both intensity and peak position. This algorithm normally gives the best results and is recommended for most applications.

When a search is performed, a hit list is produced that ranks the reference spectra in order of match quality. Some programs assign a value of 100 to a perfect match, others use zero. In either case, a good hit would obviously be one that is closest to perfection. In the best–case scenario, the top hit (if 100 is perfect) is above 90 and all other hits are significantly lower. In cases of spectral mixtures, however, the best hit may not even be 50. In such cases, the best reference spectra are compared to the sample spectrum to determine if they could represent a portion of the sample. If so, a spectral subtraction can be performed to remove the reference component from the sample spectrum and the resultant spectral subtraction can be searched again to look for additional components. Often, if strong peaks are being subtracted from one another, regions of over- or under–subtraction will occur, producing derivative–shaped peaks in the subtraction spectrum. In these cases, it may be advantageous to ‘blank’ these regions before searching for lesser components. The process of spectral searching and subtracting can be repeated until the signal–to–noise of the subtraction spectrum yields unusable results and no more components can be identified.

Several commercial libraries are useful for drug analysis (see Spectral Library Collections). Sigma-Aldrich has a

comprehensive library containing over 14 000 FT–Raman spectra. Thermo Nicolet markets two useful libraries. One is a Raman forensic library containing 175 spectra of drug compounds, excipients, precursors and metabolites. The other is a pharmaceutical excipients library that contains 300 reference spectra. For this second library, there is also a matched set of corresponding FT–IR spectra. Galactic Industries and Bio-Rad are two other popular suppliers of spectral libraries. Of course, sometimes the most useful libraries are ones that analysts create themselves.

A final note on the use of spectral libraries: the results from a hit list should always be verified by:

comparing the reference spectrum to the sample spectrum
ensuring that the reference spectrum is named accurately in the library (no library is completely perfect)
determining that the suggested hit is a logical component to expect in that particular sample.