The Role of Raman Spectroscopy Within Quantitative Metabolomics

Cassio Lima, Howbeer Muhamadali, and Royston Goodacre

Department of Biochemistry and Systems Biology, Institute of Systems, Molecular, and Integrative Biology, University of Liverpool, Liverpool L69 7ZB, United Kingdom; email: roy.goodacre@liverpool.ac.uk

Keywords
Raman spectroscopy, metabolomics, surface-enhanced Raman scattering, SERS, coherent anti-Stokes Raman spectroscopy, CARS, stimulated Raman spectroscopy, SRS, spatially offset Raman spectroscopy, SORS

Abstract
Ninety-four years have passed since the discovery of the Raman effect, and there are currently more than 25 different types of Raman-based techniques. The past two decades have witnessed the blossoming of Raman spectroscopy as a powerful physicochemical technique with broad applications within the life sciences. In this review, we critique the use of Raman spectroscopy as a tool for quantitative metabolomics. We overview recent developments of Raman spectroscopy for identification and quantification of disease biomarkers in liquid biopsies, with a focus on the recent advances within surface-enhanced Raman scattering–based methods. Ultimately, we discuss the applications of imaging modalities based on Raman scattering as label-free methods to study the abundance and distribution of biomolecules in cells and tissues, including mammalian, algal, and bacterial cells.
1. INTRODUCTION

Metabolomics is one of the newest omics platforms and aims to study the metabolome, i.e., the set of metabolites within a biological sample (1). Metabolites are small molecules with low molecular weight (in relation to proteins and nucleic acids) that are involved in biological processes as reactants, intermediates, or products of biochemical reactions (2). The biological sample studied can be an entire organism, a tissue specimen, a cell, an organ, an organelle, or a biofluid. The metabolome is composed of metabolites originating from a number of processes and may include endogenous metabolites; these are naturally produced and consumed within a biological system (e.g., sugars, organic acids, amino acids, nucleic acids, lipids, vitamins) as well as exogenous metabolites that are not synthesized by the organism and therefore are imported from outside (e.g., drugs, environmental contaminants, food additives, toxins, and other xenobiotics) (2). The main focus of metabolomics is to provide a better understanding of biological functions by investigating the link between metabolites (defined as metabolism) and their interactions with other biochemical species. It occupies a unique position compared to the other omics platforms, as the metabolome is the closest link to physiology and reflects all information expressed and modulated by all other major omics approaches (genome, transcriptome, and proteome) (2). Therefore, metabolomics is generally considered as the link between genotype and phenotype because the metabolome is the final response of a biological system to genetic or environmental changes (e.g., diet, age, lifestyle, medications, and diseases) and in humans can also be affected by the micro-organisms that live in unison within what is termed a superorganism (3).

Over the last four decades, extensive research has been conducted in search of genes responsible for many chronic diseases. However, recent studies have revealed that human health relies not only on genes but also on the interaction of multiple genes with environmental factors (4), an area often referred to as the exposome (5, 6). Owing to its close relationship to the phenotype and its unique ability to probe complex biochemistry, metabolomics is a powerful tool for exploring human diseases based on how small molecules are globally affected by physiological and pathological changes. The importance of metabolites and their role on the development of chronic and complex diseases such as cancer (7, 8), diabetes (9), and atherosclerosis (4) have been previously demonstrated by metabolomics approaches. The evaluation of genes and genetic risk scores may be useful for evaluating the diseases that an individual may develop in the future; however, metabolomics has the potential to provide information about the biological disorders currently taking place within an individual (4). The sensitivity of the metabolome to external stimuli has been used not only as a biomarker to report disease status but also as a tool to monitor therapeutic outcomes, pointing out the growing role of pharmacometabolomics in precision medicine (10, 11). The main objective of precision or personalized medicine is to combine the medical treatment of an individual to its profile obtained through advanced omics testing in order to develop specialized treatment and thus improve the therapeutic outcomes (12). Several studies on complex diseases such as cancer have shown that the cancers of two patients are not exactly the same, and hence, different responses to generic treatments such as radiotherapy and chemotherapy may be expected owing to the interpatient variability (10). Providing a treatment modality that best suits the individual response of each patient is a key factor in avoiding the exposure of the individual to ineffective and costly treatments that can cause unnecessary side effects and, most importantly, achieve a favorable prognosis (13). Furthermore, it is well known that drug metabolism can vary according to physiological variables such as ethnicity, age, gender, weight, height, and diet (12). Thus, monitoring the metabolome and the biochemical components obtained by genomics and other omics methodologies can offer major improvements in fighting not only cancer but several other diseases by providing new insights into disease pathophysiology and the mechanisms
responsible by means of differences in drug responses and by customizing drug dosing owing to interpatient variability (4, 10, 14).

The basic methods used in metabolomics research involve advanced analytical techniques combined with computational tools that allow for the characterization of complex mixtures of small molecules within the metabolome (15, 16). It is important to highlight that the metabolome can be wide and diverse regarding the physical and chemical properties of metabolites, which may be found distributed over a wide range of concentrations. Hence, there is no single analytical platform that allows for the identification and quantification of the complete set of metabolites within a biological system and, therefore, combining different analytical techniques is a common strategy in metabolomics investigations (2, 17, 18).

Over the past 15 years, two analytical techniques have emerged as primary platforms in metabolomics research: nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) (8, 12). Both techniques enable the identification and quantification of a large number of metabolites coexisting in complex samples. NMR spectroscopy is highly reproducible and requires minimal sample preparation compared to MS-based methods. However, the sensitivity achieved by NMR spectroscopy is relatively low compared to that obtained by MS platforms owing to the signal overlapping at certain spectral regions (19). Although MS is more sensitive than NMR, it usually involves the coupling of a separation technique such as gas chromatography (GC) and liquid chromatography (LC) because of the complexity of biological mixtures. Gas chromatography-mass spectrometry (GC-MS) is a powerful tool to analyze volatile metabolites with high sensitivity, resolution, and reproducibility. Major drawbacks are that it requires time-consuming sample processing and derivatization steps to improve sample volatility or increase retention on the GC column. Liquid chromatography-mass spectrometry (LC-MS) has become the main technique for metabolite profiling. Sample volatility is not a requirement, so there is no need for derivatization, resulting in a broader range of detectable analytes (18). Chromatography-MS (GC-MS and LC-MS) and NMR spectroscopy are by far the most widely applied platforms for identification and quantification of metabolites. However, other MS-based methods such as capillary electrophoresis-MS and direct infusion (or direct injection) MS can also be used in metabolomics research (18).

Despite the remarkable progress in biomedical fields through MS-based metabolomics, there remain limitations and challenges regarding the translation of such technologies into clinical practice. One major challenge is the data variability obtained by different studies performed on the same or similar samples by independent research groups, which are generally introduced by factors such as sample preparation (quenching, extraction, and derivatization), data acquisition, and data analysis (20, 21). These challenges may be overcome by developing standard methods that are able to measure the absolute concentration of metabolites. However, this approach is not straightforward for MS, as it requires suitable internal standards and calibration methods (22). Until recently, measuring the relative concentration of metabolites has been the main goal of most MS-based metabolomics studies because this approach is sufficient to evaluate the overall changes in the metabolome within a biological system under perturbations. However, given the need for rapid and accurate absolute quantification in clinical settings, there is an increasing drive toward standardized approaches within the metabolomics community, with an emphasis on quantification and automation (21).

Vibrational spectroscopic methods such as infrared (IR) and Raman spectroscopies have also been used as analytical tools in metabolomics investigations. Both methods interrogate the chemical profile of a sample based on its chemical vibrations and provide spectral fingerprints representing a snapshot of the overall biochemistry of the sample at a given time (1, 23, 24). Spectral signatures provided by Raman and IR spectroscopies are obtained through distinct modes of
interaction between the incident radiation and the sample. Raman shift results from inelastic scattering of the incident light due to a change in polarizability of the molecule, whereas an IR spectrum is obtained based on the absorption of the incident radiation due to a change in the dipole moment of vibrating molecules (23). The Raman signal is obtained through a monochromatic laser beam of high intensity (i.e., a laser), whereas IR spectroscopy utilizes a source emitting IR radiation such as a globar, synchrotron radiation, and/or a quantum cascade laser. Owing to their distinct operating modes, both methods have advantages and limitations. One major drawback of IR spectroscopy is the strong absorption of water over the mid-IR region (4,000–400 cm$^{-1}$), which requires the dehydration of samples and specialized sampling such as attenuated total reflectance or digital subtraction of the water signal (23). On the other hand, water molecules are weak scatterers and therefore do not interfere with the Raman signal. However, Raman spectra acquired with a visible laser from biological samples can often be dominated by fluorescence, which is not an issue for IR spectroscopy (23). Vibrational spectroscopic techniques are not as sensitive as MS-based methods, and accurate identification of metabolites in complex biological samples is still a challenge. However, Raman and IR spectroscopies provide information on general groups of biomolecules such as carbohydrates, lipids, proteins, and nucleic acids simultaneously in a rapid, reagent-free, nondestructive, high-throughput, and relatively inexpensive manner compared to other platforms commonly used in metabolomics studies. Furthermore, Raman and IR technologies are available as portable devices (25), which is ideal for point-of-care medicine. Because of its ability to provide biochemical information in a holistic manner, vibrational spectroscopy has been used as a valuable tool for metabolic fingerprinting/footprinting. In light of this, we review the recent developments in Raman spectroscopy as a quantitative method for metabolomics investigation.

2. SPONTANEOUS RAMAN SCATTERING AND RELATED PHENOMENA

The Raman effect was first observed in 1928 by Sir Chandrasekhara Venkata Raman (26), who began experimenting with light scattering in an attempt to understand the blue color of the Mediterranean Sea that he observed on a trip through Europe in 1921 (27). In 1930, he was awarded the Nobel Prize in Physics. During the next 30 years, Raman experiments were carried out on large sample volumes that took up to three months to prepare owing to the long purification process (multiple distillations) to avoid undesirable effects such as fluorescence and stray light from impurities. Spectral acquisition could take days due to the long integration time, and these data were recorded using prism spectrographs, Hg lamps, and photographic plates. The invention of the laser in the 1960s and its implementation as a source of excitation in spectrometers had a major impact on Raman technology, allowing the acquisition of data in a faster and more reproducible manner. This explains the rise in the number of publications observed in the early 1970s (Figure 1d). Similar to other analytical techniques, the development of Raman spectroscopy is directly influenced by advances in technology and has been driven by the need for improvements in sensitivity, speed of collection, spatial resolution, and specificity. Ninety-four years have passed since the discovery of Raman scattering, and there are currently more than 25 different types of Raman-based techniques (28). Figure 1 shows the total number of publications from 1930 to 2020 listed in the Web of Science using names of the main Raman-based methods in the biomedical sciences as keyword search terms.

2.1. Spontaneous Raman Scattering

When electromagnetic radiation interacts with a molecule, the energy of the incident photon can be transferred to the molecule through different modes of interaction such as absorption,
Figure 1
Bar charts showing the total number of publications from 1930 to 2020 in the Web of Science repository using the names of the main Raman-based techniques used in the biomedical sciences as search terms.
emission, or scattering (28, 29). In absorption, the molecule may undergo a process called transition as a consequence of the transference of energy, which is characterized by a change from an initial state to a posterior state. The molecular transition can be classified as electronic, vibrational, or rotational (30). Electronic transitions occur when the molecule absorbs a photon with energy compatible with wavelengths in the ultraviolet–visible (UV-Vis) spectral range. Vibrational transitions require radiation with typical wavenumbers between 400 and 4,000 cm$^{-1}$, i.e., infrared radiation. Rotational transitions occur when radiation with frequencies between 1.6 and 30 GHz is absorbed (microwave) (30). Vibrational transitions occur between different vibrational levels of the same electronic state, while rotational transitions occur between rotational levels of the same vibrational state (31) (Figure 2a). The basic rule for the occurrence of a transition is that the energy of the incident photon matches the energy difference between two states. In the case that the incoming photon does not match any molecular transition, the photon is scattered (28).

Light scattering can be subdivided into elastic or inelastic. The first one does not involve energy exchange (the wavelength remains unchanged after scattering) such as Rayleigh and Mie/Tyndall scattering (28). Inelastic scattering is represented by Brillouin, Compton, and Raman scattering (28). In Raman, the incoming photons go into a virtual state and then are measured as scattered
photon with different energy. Virtual states are created at the time of laser incidence and are not associated to electronic, vibrational, or rotational states of the molecule. If the photon transfers its energy to the molecule, the scattered photon will have lower energy and longer wavelength (Stokes scattering), which means that the initial state is the ground state, and after the scattering process, the final state is an excited vibrational state (Figure 2c). On the other hand, if the molecule is already in an excited vibrational state and transfers energy to the incident photon, the scattered photon will have higher energy and a shorter wavelength (anti-Stokes scattering), which means that the final state is the ground state (28) (Figure 2d). Spontaneous Raman scattering or the Raman effect refers to the inelastic scattering of light, including Stokes and anti-Stokes processes. The energy difference of inelastic scattering corresponds to the energy difference between the excited ($v = 1$) and fundamental vibrational states of mid-IR. At room temperature, most molecules are in the fundamental state; therefore, the Raman effect is dominated by the Stokes process rather than the anti-Stokes phenomenon (29).

### 2.2. Nonlinear Raman Scattering

In 1974, development of the first Raman microscopes was presented at the Fourth International Conference on Raman Spectroscopy (ICORS) held in Brunswick, Maine. However, the concept of combining Raman spectroscopy to a microscope for obtaining molecular information from samples at microscopic level was suggested almost a decade earlier by Delhaye and Migeon in 1966 (29). Since then, Raman microspectrometers have been upgraded, and nowadays it is possible to obtain images through spontaneous Raman scattering. The combination of spectral with spatial information provided by Raman imaging has made it possible to evaluate the distribution of compounds based on their spectral signatures, providing a better understanding of the physicochemical properties of the sample. The main drawback of spontaneous Raman imaging is the long acquisition time per pixel due to the poor scattering cross section of Raman effect (typically only $\sim 10^6$–$10^7$ photons are converted to Raman scattered light), which makes the image acquisition very long and unsuitable for real-time imaging of dynamical processes such as in biological systems (32).

To overcome the long sample analysis times for spontaneous Raman imaging, techniques based on coherent Raman scattering (CRS), such as coherent anti-Stokes Raman spectroscopy (CARS) and stimulated Raman spectroscopy (SRS), have been developed for accelerating the acquisition of Raman images. CRS enables label-free imaging of living cells and tissues by enhancing the weak Raman signal through third-order nonlinear optical processes (33). In CRS, two high-power pulsed lasers (picosecond or femtosecond lasers) are used to generate CARS/SRS signals; one of the lasers is known as the pump beam ($\omega_{\text{pump}}$) and the other as the Stokes ($\omega_{\text{Stokes}}$) beam (28). In CARS, the pump ($\omega_{\text{pump}}$) and Stokes ($\omega_{\text{Stokes}}$) beams are tuned onto resonance with the frequency of a Raman peak. In the following, a probe photon of frequency ($\omega_{pr}$) induces a perturbation for the anti-Stokes scattering process to occur at frequency $\omega_{\text{CARS}} = \omega_{\text{pump}} - \omega_{\text{Stokes}} + \omega_{pr}$. In most CARS systems, the pump and probe photons are provided by the same laser ($\omega_{pr} = \omega_{\text{pump}}$) and thus, $\omega_{\text{CARS}} = 2\omega_{\text{pump}} - \omega_{\text{Stokes}}$ (28, 33) (Figure 3a). The major drawback of CARS is the occurrence of an interfering nonresonance background during the CARS signal generation process, which reduces image contrast and complicates quantification (28). In SRS, photons from $\omega_{\text{pump}}$ and $\omega_{\text{Stokes}}$ beams interact with a molecule, resulting in a transfer of energy from the pump photon to the molecule (stimulated Raman loss), and then the molecule scatters a new photon with frequency ($\omega_{\text{Stokes}}$, stimulated Raman gain) (28, 34) (Figure 3b). The SRS signal is quantified by taking into consideration the amount of energy transferred from the pump beam to the Stokes beam. In contrast to CARS, SRS is free of undesirable background signals and has a linear dependence
Figure 3

(a) Energy transfer in coherent anti-Stokes Raman spectroscopy (CARS) and (b) stimulated Raman spectroscopy (SRS) processes. $\Omega$ represents the energy difference between the excited ($\nu = 1$) and fundamental vibrational states of mid-infrared.

to concentration, which makes SRS a better tool for quantitative chemical imaging and spectral analysis than CARS (35).

2.3. Signal Enhancement Through Surface Plasmon and Resonance Raman Spectroscopy

CARS and SRS microscopies have emerged as powerful techniques for label-free imaging in biological systems; however, the spatial resolution achieved in both methods is restricted by the diffraction limit of light, which is $\sim 1 \mu m$, according to Abbe's criterion (28). To achieve submicron spatial resolution, Raman spectroscopy was combined to scanning near-field optical microscopy technology and gave rise to tip-enhanced Raman spectroscopy (TERS) nanoimaging (28). In TERS, a sharp metal or metal-coated nanotip is positioned a few nanometers above the sample, and the Raman excitation light is focused onto the surface of the tip, enabling nanoscale chemical imaging of the sample surface with a spatial resolution of 10–30 nm (28). Besides the improvement on spatial resolution, the electromagnetic field of the Raman excitation light is also enhanced by localized surface plasmon resonance (LSPR) on the tip's apex. Hence, the number of Raman photons being generated is higher compared to spontaneous Raman scattering. Surface-enhanced Raman scattering/spectroscopy (SERS) and resonance Raman spectroscopy (RRS) can also be used to increase the sensitivity while maintaining the specificity of Raman scattering. In SERS, the sample is adsorbed on a solid substrate with metallic nanostructures on the surface or in a colloidal suspension of metallic nanoparticles such as gold or silver, leading to amplification of the electric field of the incident light via LSPR (similar to TERS) (36). The SERS signal is also amplified by chemical effect due to the transference of charge between the metallic structure and the analyte, which results in an enhancement factor higher than in TERS. It also enables the quantitative determination and identification of molecules at ultralow concentrations and quenching of the fluorescence background (37). In RRS, the Raman signal is improved by frequency coincidence (or resonance) of the excitation laser to the energy required for an electronic transition of a compound (Figure 2b), which is sensitive and highly specific to obtaining information from chromophoric species, e.g., with excitation in the deep UV; aromatic amino acids and nucleic acids are then selectively enhanced (38). Both RRS and SERS techniques can be combined
(surface-enhanced resonance Raman spectroscopy, SERRS), resulting in an amplification factor of up to ten orders of magnitude as compared to spontaneous Raman spectroscopy (39).

3. APPLICATIONS OF RAMAN SPECTROSCOPY IN METABOLOMICS STUDIES

Raman spectroscopy is a powerful analytical technique to study the chemical composition of materials and has been extensively used in metabolomics investigations. In the first part of this section, we present recent results reported by various studies where Raman signatures are collected in single-point mode; i.e., a single Raman spectrum is collected from the material under analysis. Next, we discuss the applications of imaging techniques based on Raman scattering as label-free methods to study the distribution of biomolecules.

3.1. Raman and Metabolomics: Single-Point Mode

Biological information can be extracted from Raman spectra through quantitative and/or qualitative analysis, which are decided by the user based on the spectral features and final goal of the experiment. Figure 4 shows situations that may occur in a Raman experiment: appearance/disappearance of peaks (Figure 4a), changes in peak intensities (Figure 4b), shifts in peak position (Figure 4d), and no clear spectral changes as judged by the naked eye (Figure 4e).

Within the context of qualitative analysis, information can be retrieved from situations depicted in Figure 4a-d. Appearance/disappearance of Raman peaks is a good alternative to analyze the presence and/or absence of molecules in a biological system. Alonso-Pernas et al. (44) evaluated the accumulation of poly-β-hydroxybutyrate (PHB) in bacterial cells by monitoring vibrations from PHB that are not naturally observed in the spectrum of bacteria. In another study, Raman was used for real-time monitoring of the differentiation of neural cells (PC12 cells, embryonic stem cells, and adult stem cells), in which the appearance and disappearance of several peaks were observed during cell differentiation (45). SERS has also been used to diagnose infections caused by pathogens, including bacterial meningitis in cerebrospinal fluid samples, by monitoring the appearance of peaks associated to neopterin [695 cm\(^{-1}\)] (40) (Figure 4a), as well as red blood cells infected with \textit{Plasmodium falciparum} through the analysis of a new peak at 1,599 cm\(^{-1}\) (46). The spectral contrast achieved by the appearance/disappearance of Raman peaks enables the dynamics of chemical species in biological samples to be monitored and plays an important role in studies aiming to image the spatial distribution of target molecules. Nevertheless, peaks from different molecules may overlap in the spectrum of a biological complex mixture, resulting in similar endogenous background and nonspectral contrast (47). In such cases, new chemical bonds can be introduced to circumvent the endogenous cellular background. Incorporation of heavy atoms into cells is one of the most common biochemical tricks used to introduce new chemical bonds and therefore achieve spectral contrast (48). \(^2\)H, \(^13\)C, and \(^15\)N are the most common stable isotope-labeled atoms used in metabolomics investigations to track activity and function of molecules in living organisms and have been exploited in studies involving NMR (8, 49) and MS (49) as well as Raman (42, 50, 51) and IR (50) spectroscopies. The combination of vibrational spectroscopic methods and stable isotope probing enables the direct measurement of unlabeled and labeled molecules, such as carbohydrates, proteins, nucleic acids, and lipids, by monitoring the position of peaks in the spectra and how these peaks shift when heavy atoms are incorporated into their molecular structure (42, 52) (spectral shifts are represented in Figure 4d). The cellular uptake of \(^2\)H results in the appearance of a peak due to carbon-deuterium (C–D) vibration peaking between 2,040 and 2,300 cm\(^{-1}\) in the Raman spectrum, which corresponds to a shift in
Figure 4

Common situations obtained during a Raman experiment. (a) Surface-enhanced Raman spectroscopy (SERS) spectra of healthy cerebrospinal fluid samples and those infected by *Neisseria meningitidis*, in which it is possible to observe the appearance of peaks related to neopterin. Panel adapted from Reference 40 under the terms of the Creative Commons Attribution (CC BY) License, http://creativecommons.org/licenses/by/4.0. (b) Changes in peak intensities of SERS spectra for increasing concentrations of miRNA-203. (c) The calibration curve obtained by plotting the $I_{1499}/I_{783}$ ratio against the logarithmic concentration of miRNA-203. Panels b-e adapted with permission from Reference 41; copyright 2017 American Chemical Society. (d) Shifts in peak position in SERS spectra of *Escherichia coli* cells cultivated on unlabeled (dashed line), $^{13}$C-labeled (red line), $^{15}$N-labeled (blue line), and $^{13}$C-$^{15}$N-labeled (green line) substrates. Panel adapted with permission from Reference 42; copyright 2017 Royal Society of Chemistry. (e) Raman spectra of 11 *Campylobacter* strains and principal component–discriminant function (DF) analysis scores in panel f showing satisfactory discrimination. Panels e,f adapted from Reference 43 under the terms of the Creative Commons Attribution Noncommercial-Share Alike 3.0 Unported License (CC BY-NC-SA) License, https://creativecommons.org/licenses/by-nc-sa/3.0/. Other abbreviation: PBS, phosphate-buffered saline.
C–H vibrations due to the substitution of hydrogen atoms by deuterium (48, 52). The spectral region 1,800–2,600 cm\(^{-1}\) is known as the silent zone because there are generally no detectable Raman peaks in cells grown with natural elements (except for molecules containing alkynes and nitriles, but these are biologically rare) and, thus, the appearance of a peak in this region provides good spectral contrast (47). Spectral changes due to incorporation of heavy isotopes are always shifted toward lower wavenumbers (so-called redshift) regardless of the compound used to label the cells and occur owing to an increase in the reduced mass that leads to a decrease in the vibrational frequencies (wavenumbers) of chemical bonds (42). Shifts resulting from the incorporation of \(^{13}\)C and \(^{15}\)N atoms are mainly observed in the low wavenumber region of Raman spectra, and therefore the appearance of peaks in the silent region is not observed, as is seen from C–D in Raman-deuterium isotope probing (50). With a wide variety of possible isotopic substrates, Raman-stable isotope probing has been used to probe cell activity and the metabolism of a wide range of living organisms, including bacteria (42, 48, 51, 53–56), algae (57), Caenorhabditis elegans (58), human cell lines (59), and human tissues (58).

Raman signals are linearly correlated to the number of molecules under investigation and can thus provide quantitative information (60). In such cases, a calibration curve is constructed based on the area or height of a peak plotted against the concentration of the analyte, as shown in Figure 4b,c (41). Once calibrated, the method can then be used to predict the concentration of the analyte from unknown samples (36). Quantification analysis can be performed when combined with univariate or multivariate statistical methods. For peaks that are not overlapped with other spectral features such as the situation depicted in Figure 4b (41), univariate analysis may be sufficient. In contrast, when the analyte is within a mixture with a complex background, or when information from more than one analyte is needed as input, multivariate analysis applied to the whole spectrum or a subset of the spectral features is recommended (60). Constructing a calibration curve requires spectra with well-defined peaks; thus, signatures with unclear peaks are poor candidates as inputs for calibration. This may be an issue for quantifying molecules within biological samples, as the concentration of the analytes is usually ultralow. In such cases, other techniques based on Raman scattering with higher sensitivity, such as SERS, offer more advantages.

Different analytical techniques, methods, and devices based on SERS have been developed for identification and/or quantification of analytes within biological samples (36, 37, 39, 60). The idea of a portable and cost-effective SERS device that is able to identify disease biomarkers has attracted increasing attention over the last 20 years, especially in the field of point-of-care diagnosis (36). However, some analytes do not possess the affinity to plasmonic nanoparticle surfaces and therefore it is common to conjugate metallic nanoparticles to ligands to enhance the chemical affinity of molecules and to recognize specific analytes. For example, antibodies are commonly employed in SERS platforms aiming to bind to and thus identify proteins, whereas oligonucleotide strands are used to detect selectively nucleic acids (36). The strategy employed to identify a specific analyte plays an important role in the sensitivity of the SERS platform. Zhang et al. (61) were able to detect dopamine at picomolar concentrations by first immobilizing dopamine molecules on gold thin film surfaces, followed by introducing 3-mercaptopropylboronic acid (3-MPBA)-functionalized silver nanoparticles to generate hot spots. In another study aiming to detect the same molecule, Ranc et al. (62) achieved a better limit of detection (picomolar) using a method based on a nanocomposite composed of magnetite and silver nanoparticles with a modified surface [iron nitritoacetic acid (Fe-NTA)] to selectively identify the molecules of dopamine. Several studies have focused on the development of SERS-based methods as a tool for liquid biopsy, which comprises the identification of disease biomarkers and the efficiency of a drug treatment in biofluids such as cerebrospinal fluid, saliva, urine, blood serum, and plasma (36). Table 1 lists some of the molecules targeted by studies using SERS in liquid biopsy research, including small molecules, proteins, nucleic acids, and drugs.
Table 1  Analytes from human biofluids that were quantified through surface-enhanced Raman scattering (SERS) platforms

<table>
<thead>
<tr>
<th>Biofluid</th>
<th>Category</th>
<th>Analyte</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cerebrospinal fluid</td>
<td>Small molecules</td>
<td>Dopamine (61, 62, 103), neopterin (40), melatonin (103), serotonin (103), glutamate (103), GABA (103), norepinephrine (103), epinephrine (103)</td>
</tr>
<tr>
<td></td>
<td>Proteins</td>
<td>Tau protein (104, 105), Aβ1–42 oligomers (105)</td>
</tr>
<tr>
<td></td>
<td>Nucleic acids</td>
<td>miRNAs (106, 107)</td>
</tr>
<tr>
<td>Saliva</td>
<td>Small molecules</td>
<td>Sialic acid (108), nitrite (109), pyocyanin (110), thiocyanate (111)</td>
</tr>
<tr>
<td></td>
<td>Proteins</td>
<td>Urease (112), MnSOD (113)</td>
</tr>
<tr>
<td></td>
<td>Nucleic acids</td>
<td>S100P mRNA (114)</td>
</tr>
<tr>
<td></td>
<td>Drugs</td>
<td>Codeine (115), fentanyl (115), JWH-018 (116), THC (114), buprenorphine (117), cocaine (118)</td>
</tr>
<tr>
<td>Urine</td>
<td>Small molecules</td>
<td>Uric acid (119), 5β-pregnan-3α,20α-diol-3α-glucuronide (120), tetra hydrocortisone, creatinine (121), urea (121), adenosine triphosphate (122), melatonin (103), serotonin (103), glutamate (103), GABA (103), norepinephrine (103), epinephrine (103), thiocyanate (123), tramadol (124)</td>
</tr>
<tr>
<td></td>
<td>Proteins</td>
<td>Albumin (125)</td>
</tr>
<tr>
<td></td>
<td>Drugs</td>
<td>Propranolol (126), methamphetamine (127, 128), cocaine (129), methcathinone (128), methotrexate (130), MDMA (128)</td>
</tr>
<tr>
<td>Blood serum/plasma</td>
<td>Small molecules</td>
<td>Adenosine triphosphate (122, 131), thiocyanate (123)</td>
</tr>
<tr>
<td></td>
<td>Proteins</td>
<td>α-Fetoprotein (132, 133), p53 (134) and p53&lt;sub&gt;R175H&lt;/sub&gt; (134), TNF-α (135), thrombin (136), MnSOD (113)</td>
</tr>
<tr>
<td></td>
<td>Nucleic acids</td>
<td>miRNA (132, 137–139)</td>
</tr>
<tr>
<td></td>
<td>Drugs</td>
<td>Codeine (115, 140), propranolol (126), 6-thioguanine (141), methotrexate (142), estradiol (143), THC (114), fentanyl (115), imatinib (144), doxorubicin (145)</td>
</tr>
</tbody>
</table>

Abbreviations: GABA, gamma aminobutyric acid; MDMA, 3,4-methylenedioxymethamphetamine; MnSOD, manganese super oxide dismutase; TNF-α, tumor necrosis factor-alpha.

SERS is not the only Raman method employed for quantification of molecules with biological interest. Although RRS has also been used for this purpose (63), it is more recommended for quantifying chromophoric species, while SERS can be applied in identifying a wide variety of molecules. Furthermore, SERS technology can be translated into portable devices for point-of-care medicine and is therefore by far the preferred method used for quantitative analysis.

SERS technology has also attracted attention in microbiology as a tool to study microbial cells, with one of the earlier studies showing that this technique was reproducible enough to type different bacteria associated with urinary tract infection (64). However, spectral signatures of different microbes can be similar and sometimes indistinguishable (43) (Figure 4e) to the naked eye, as the example shown in Reference 65, which compared SERS spectra from *Bacillus subtilis* to those from *Escherichia coli*. In such cases, multivariate statistical methods may be required to discriminate signatures from different microorganisms (66). Dina et al. (67) used SERS combined with principal component analysis (PCA) to detect two Gram-positive (*Enterococcus lactis* CE13 and CE39) and two Gram-negative (*E. coli* Rosetta (DE3) pLysS and *E. coli* XL1Blue) species of bacteria at the single-cell level. PCA is the most common technique employed for this purpose and has been combined with SERS as an easy, rapid, and accurate method for identification of microbial infections in clinical biochemistry (68), quantification of bacterial cells (42, 69), characterization of biofilms (70), and detection of biological warfare markers (71). When paired with stable isotope labeling, SERS is a powerful tool for functional analysis of microorganisms (48). For an excellent recent review on where SERS is used within microbiology, the reader is directed elsewhere (38).
3.2. Raman and Metabolomics: Imaging Mode

Visualizing and quantifying the spatial distribution of molecules in biological systems play an important role in understanding biological processes (35). Fluorescence-based methods have been extensively used for this purpose because they allow for the visualization of molecules down to the single-molecule level. However, such methods require endogenous (fluorescent proteins) or exogenous (organic dyes or quantum dots) markers to label the molecules of interest, which may directly interfere with their biological function or indirectly through undesired side effects such as photobleaching and/or phototoxicity (35). Furthermore, fluorescent labels may not work for small molecules due to the comparable size (or even bigger) of the marker compared to the molecule of interest. Thus, label-free chemical imaging obtained through Raman techniques is particularly attractive in its ability to provide information on the spatial distribution of molecules based on its chemical vibrations (32).

Images obtained through spontaneous Raman scattering, commonly referred to as Raman microspectroscopy, can be collected as chemical maps or hyperspectral images (72). In a chemical map, each pixel is a number representing the intensity of a single Raman peak and provides information on the spatial distribution of the chemical species related to the peak. However, owing to the inherent overlapping peaks in biological samples, the molecular specificity achieved by chemical maps is generally low (72). In a hyperspectral image, each pixel contains a Raman spectrum containing multiple Raman shifts, which can be processed through several different methods to obtain spatial information. Chemical maps can be generated from hyperspectral images by selecting the intensity or area of a specific Raman peak and the ratio between peak intensities/areas through univariate approaches. Shao et al. (73) evaluated the intracellular lipid content in the algae Scenedesmus obliquus by generating chemical maps from hyperspectral images using the intensity of a peak related to lipids (1,445 cm\(^{-1}\)). In another study evaluating the profiles of storage lipids in Fistulifera solaris, Hosokawa et al. (74) obtained good correlation between algal oil content measured by GC-MS and integrated Raman intensity at 1,445 cm\(^{-1}\), indicating the potential of Raman techniques as a tool for rapid in vivo quantification of oil in oleaginous microalgae (Figure 5g–j).

Extensive efforts have been devoted to extract morphological information from Raman hyperspectral images through multivariate methods such as clustering algorithms. In such cases, the whole spectrum or a subset of the spectral features within each pixel is used as an input to the clustering algorithm, which partitions the data set into subgroups (clusters) of spectra with similar spectral features and generates a false-color map that will ultimately provide information about the structures and morphology of the biological system. Anna et al. (75) were able to discriminate human brain tumors from normal structures by reconstructing Raman images through \(k\)-means clustering analysis. For other examples of Raman spectroscopy associated to clustering techniques to segment hyperspectral images collected from soft tissue, the reader is directed elsewhere (33, 76). Limited work has been conducted on extracting quantitative information from hyperspectral images (77). In these cases, the Raman spectrum within each pixel is first decomposed into spectral signatures of biochemical components using sophisticated factor-based analysis such as multivariate curve resolution-alternating least squares (MCR-ALS) and spectral unmixing techniques, including vertex component analysis or the N-FINDR algorithm (77). This is possible based on the assumption that the overall tissue spectra are the linear combination of the spectra belonging to each biochemical component. Thus, the algorithm assumes the system is closed and estimates the contribution of each component to the whole spectrum to evaluate the distribution and abundance of each component in the image. Subsequently, a false-color map displaying the abundance and spatial distribution of each identified component is obtained (Figure 5a–f). You et al. (78) used vertex component analysis (VCA) to extract Raman signatures from collagen, elastin, actin,
cholesterol ester, cholesterol, triglycerides, \( \beta \)-carotene, whitlockite, and apatite from hyperspectral images to evaluate the role of each component in medial aortic calcification and atherosclerosis in human aorta. A study by Albro et al. (79) employed MCR to decompose the Raman spectrum within each pixel into signatures of extracellular matrix constituents (glycosaminoglycans, collagen, and water) to assess their distribution in cartilage tissue. Liu et al. (80) applied non-negative matrix factorization to evaluate the distribution of the signatures of \( E. \ coli \) cells, protein, and polyhydroxybutyrate within a biofilm matrix. Other examples include the development of a
classification model for nonalcoholic steatohepatitis (81), analysis of mesenchymal stem cells cultured under biological conditions (77), and volumetric imaging of three-dimensional cell cultures (82) (Figure 5a–f).

Although Raman microspectroscopy has been successfully applied to visualize the spatial distribution of molecules in biological systems, the method is unsuitable for real-time imaging owing to the long acquisition time per pixel (28, 34). Here, CRS imaging, including CARS and SRS, provides significant advantages through its ability to collect images at video-rate acquisition speed (32). CRS images are obtained by tuning the pump and Stokes beams onto resonance with the frequency of a specific Raman peak (illustrated in Figure 3), resulting in a chemical map displaying the distribution of that specific chemical species. A common strategy is tuning the pump beam to multiple wavelengths and collecting a series of chemical maps that will ultimately provide information to reconstruct CARS/SRS spectra based on the intensity of pixels (83–85). SRS is free of nonresonant background, and its signal intensity is linearly proportional to molecular concentration (35); therefore, SRS images can be directly used to reconstruct spectra and perform quantitative analysis (85). In the case of CARS, additional preprocessing is needed due to the nonresonance background (83, 86). CRS microscopy has emerged as a powerful platform for cellular and tissue imaging and has provided important insights into the metabolism and dynamics of lipids, proteins, nucleic acids, and small molecules in mammalian cells (32, 87–90) (Figure 5k–n), tissues (91–95) (Figure 5o,p), microorganisms (96–98) (Figure 5q–w), and simple organisms such as Caenorhabditis elegans (58, 99) by targeting endogenous chemical bonds. One of the main drawbacks of CRS microscopy is that a variety of molecules share similar chemical bonds, which leads to the lack of spectral contrast and affects the required specificity to visualize some molecules. To circumvent this issue, a variety of special probes including stable isotope (2H, 13C, 15N) and triple-bond tags (alkyne and nitriles) have been developed and employed to further increase the sensitivity and specificity of CRS imaging (47, 52, 100, 101).

4. FUTURE PERSPECTIVES

The last 20 years have witnessed the blossoming of Raman spectroscopy as a powerful technique with broad applications in life sciences. There is no doubt that SERS has been the main Raman-based method for quantitative metabolomics and will continue to receive attention in the future owing to its ability to quantify/identify molecules at ultralow concentration in complex biological mixtures and the potential to do this as portable devices for point-of-care diagnoses.
We believe that further improvements regarding sensitivity, selectivity, and reproducibility will contribute to the translation of SERS to the clinical environment as a tool to quantify disease biomarkers from different human biofluids, including blood plasma, serum, cerebrospinal fluid, urine, and saliva. Recent advances have shown that SERS is no longer an emerging technology, and it will find more applications in biological and medical fields, especially as a point-of-care technology and in personalized medicine. More recently, SERS has been combined with spatially offset Raman spectroscopy (SORS) to provide information from molecules at subsurface layers (102). We believe this relatively young method will continue to expand according to advances in technology and analytical algorithms and will play an important role in the detection of disease biomarkers in vivo with no need for biopsy.

Despite the astounding progress in Raman microscopy in the past 20 years, there remain many challenges in the adoption of the technique for imaging chemical species in the biomedical sciences. More standardized preprocessing and data analysis techniques are needed to improve the reproducibility of results obtained by Raman images. Furthermore, in order to gain acceptance in the biomedical sciences, future developments should include correlation of Raman images with molecular-based characterization approaches, such as immunofluorescence and biochemical assays. We believe that SRS microscopy will receive more attention in the future because of its improved ability to provide direct quantitative information and real-time imaging compared to Raman microspectroscopy and CARS. Advances in this technology will hopefully increase the quality of spectra obtained through SRS images, which although significantly faster, is currently inferior compared to the spectrum obtained via spontaneous Raman spectroscopy. We anticipate that in the near future, hyperspectral Raman images will be decomposed into several images showing the spatial distribution and abundance of biochemical species in tissue and cells, which will play important roles in biology and medicine, ranging from three-dimensional mapping of small molecules in cells to studying diseases in tissue (computational histology). For this, the analysis of Raman images will rely on deep learning approaches and multivariate statistical methods that can decompose the overall spectrum within each pixel into the signatures of different molecules. Ultimately, the development of new tags/markers will improve the specificity in detecting specific molecules and decrease the similar backgrounds obtained from biological samples due to the similar chemical species in their composition.

**DISCLOSURE STATEMENT**

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

**ACKNOWLEDGMENTS**

We are very grateful to the University of Liverpool for financial support.

**LITERATURE CITED**


Contents

AI in Measurement Science
Chao Liu and Jiasbu Sun ................................................................. 1

Recent Progress in the Analytical Chemistry of Champagne and Sparkling Wines
Gérard Liger-Belair and Clara Cilindre ........................................... 21

3D Printed Electrochemical Sensors
Aya Abdalla and Bhavik Anil Patel ................................................. 47

Bipolar (Bio)electroanalysis
Laurent Bouffier, Dodzi Zigah, Neso Sojic, and Alexander Kuhn ............. 65

In Situ X-Ray Techniques for Electrochemical Interfaces
Bruna F. Baggio and Yvonne Grunder ............................................. 87

Electrochemical Affinity Assays/Sensors: Brief History and Current Status
Kenneth R. Wehmeyer, Ryan J. White, Peter T. Kissinger, and William R. Heineman ................................................................. 109

Active Flow Control and Dynamic Analysis in Droplet Microfluidics
Nan Shi, Md Mobibullah, and Christopher J. Easley ............................ 133

Environmental Toxicology Assays Using Organ-on-Chip
Patarajarin Akarapipad, Kattika Kaarj, Yan Liang, and Jeong-Yeol Yoon ........ 155

Recent Advances in Microfluidically Spun Microfibers for Tissue Engineering and Drug Delivery Applications
Joseph Scott Magnani, Reza Montazami, and Nicole N. Hasbemi .............. 185

Analytical Technologies for Liquid Biopsy of Subcellular Materials
Camila D.M. Campos, Katie Childers, Sachindra S.T. Gamage, Harshani Wijeratne, Zheng Zhao, and Steven A. Soper .............................. 207
Aqueous Two-Phase Systems and Microfluidics for Microscale Assays and Analytical Measurements
Tasdiq Ahmed, Cameron Yamanishi, Taisuke Kojima, and Shuichi Takayama .......... 231

New Advances and Applications in Field-Flow Fractionation
Christine L. Plavchak, William C. Smith, Carmen R.M. Bria, and S. Kim Ratnathanawanwongs Williams ............................................. 257

Biochemical Sensing with Nanoplasmonic Architectures: We Know How but Do We Know Why?
Andreas Dahlin ............................................................................. 281

Protein Dynamics by Two-Dimensional Infrared Spectroscopy
Goran W. Tumbic, Md Yeatbad Hossan, and Megan C. Thielges ................. 299

The Role of Raman Spectroscopy Within Quantitative Metabolomics
Cassio Lima, Howbeer Muhamadali, and Royston Goodacre ................... 323

Noncontact Nanoscale Imaging of Cells
David Klenerman, Yuri Korchev, Pavel Novak, and Andrew Shevchuk ................ 347

Glycan Labeling and Analysis in Cells and In Vivo
Bo Cheng, Qi Tang, Che Zhang, and Xing Chen .................................... 363

Developments and Ongoing Challenges for Analysis of Surface-Bound Proteins
Tobias Weidner and David G. Castner .............................................. 389

Real-Time Visualization and Monitoring of Physiological Dynamics by Aggregation-Induced Emission Luminogens (AIEgens)
Xuewen He, Jacky W.Y. Lam, Ryan T.K. Kwok, and Ben Zhong Tang .............. 413

Clinical Chemistry for Developing Countries: Mass Spectrometry
Suji Lee, Kavyasree Chintalapudi, and Abrahm K. Badu-Tawiah .................. 437

Current Challenges and Recent Developments in Mass Spectrometry–Based Metabolomics
Stephanie L. Collins, Imboi Koo, Jeffrey M. Peters, Philip B. Smith, and Andrew D. Patterson ....................................................... 467

Environmental Aspects of Oxide Nanoparticles: Probing Oxide Nanoparticle Surface Processes Under Different Environmental Conditions
Izaac Sit, Haibin Wu, and Vicki H. Grassian ...................................... 489

Errata

An online log of corrections to Annual Review of Analytical Chemistry articles may be found at http://www.annualreviews.org/errata/anchem