ADVANCED RAMAN, SERS AND ROA STUDIES OF BIOMEDICAL AND PHARMACEUTICAL COMPOUNDS IN SOLUTION

A thesis submitted to The University of Manchester for the degree of

PhD

in the Faculty of Life Sciences

2012

CLARE LEVENE

Contents

38,450 words

List of Figures

List of Tables

Abbreviations

ASR	Atomic scale roughness
CNS	Central nervous system
CT	Charge transfer
DAAP	Di-amino acid peptide
DFT	Density functional theory
EM	Electromagnetic mechanism
EMSC	Extended multiplicative scatter correction
FFD	Fractional factorial design
FWHM	Full width half maximum
GA	Genetic algorithm
HPLC	High-performance liquid chromatography
IR	Infrared
LOD	Limit of detection
MOEA	Multiobjective evolutionary algorithm
MOP	Multiobjective problem
MRS	Magnetic Resonance spectroscopy
MS	Multiple sclerosis
NMR	Nuclear Magnetic Resonance
PED	Potential energy distribution
PESA	Pareto envelope-based selection algorithm
ROA	Raman optical activity
RHMFW	Ratio between half max and full width at half max
SERS	Surface-enhanced Raman scattering
SP	Surface plasmon

Abstract

The primary purpose of this study was to investigate the combination of experimental and computational methods in the search for reproducible colloidal surface-enhanced Raman scattering of pharmaceutical compounds. In the search for optimal experimental conditions for colloidal surface-enhance Raman scattering, the amphipathic β-blocker propranolol was used as the target molecule. Fractional factorial designs of experiments were performed and a multiobjective evolutionary algorithm was used to find acceptable solutions, from the results, that were Pareto ranked. The multiobjective evolutionary algorithm suggested solutions outside of the fractional factorial design and the experiments were then performed in the laboratory. The results observed from the suggested solutions agreed with the solutions that were found on the Pareto front. One of the experimental conditions observed on the Pareto front was then used to determine the practical limit of detection of propranolol. The experimental conditions that were chosen for the limit of detection took into account reproducibility and enhancement, the two most important parameters for analytical detection using surface-enhanced Raman scattering. The principal conclusion to this study was that the combination of computational and experimental methods can reduce the need for experiments by > 96% and then selecting solutions from the Pareto front improved limit of detection by a factor of 24.5 when it was compared to the previously reported limit of detection for propranolol. Using the same experimental conditions that were used for the limit of detection, these experiments were extended to plasma spiked with propranolol in order to test detection of this pharmaceutical in biofluids. Concentrations of propranolol were prepared using plasma as the solvent and measured for detection using colloidal surface-enhanced Raman scattering. Detection was determined as <130 ng/mL, within physiological concentrations, previously achieved using separation techniques.

The second part of this thesis also involved a combination of experimental and computational methods. Raman optical activity was utilized to investigate secondary structure of amino acids and diamino acid peptides in combination with density functional theory calculations. Amino acids are important biological molecules that have vital functions in the biological system. They have been recognized as neurotransmitters and implicated in neurodegenerative diseases. Raman and Raman optical activity experimental results were compared to determine site-specific acetylation, marker bands for constitutional isomers and identification of functional groups that interact with the solvent. The experimental spectra were then compared to those from the density functional theory calculations. The results indicated that; constitutional isomers cannot be distinguished from the Raman spectra but can be distinguished from the Raman optical activity spectra, site-specific acetylation can be identified from the Raman spectra, however, Raman optical activity provides more structural information in relation to acetylation. When the results were compared to the density functional theory calculations for the diamino acid peptides the results agreed reasonably well, however, agreement was not as good for the monoamino acids because diamino acid peptides support fewer conformations due to the peptide bond whereas monoamino acids can adopt a far greater number of conformations.

Combined computational and experimental techniques have developed the ability to detect and characterize biomedical compounds, a significant move in the advancement of Raman spectroscopies.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

I. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it ("the Copyright") and she has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

II. Copies of this thesis, either in full or in abstracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of such copies made.

III. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

IV. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property Rights and/or Reproductions described in it, may take place is available from the Vice-president and the Dean of the Faculty of Life Sciences.

Acknowledgement

My first thanks go to my supervisors, Ewan Blanch and Royston Goodacre, whose support, advice and open doors (of which I have utilized greatly) have been of great value during my studies. I thank Elon Correa and Myra Kinalwa for their time, effort, enthusiasm and patience during the development of the multiobjective evolutionary algorithm and throughout my PhD. I would like to thank all the members of our group, past and present, whose advice was often required and has been greatly valued. I would like to thank Babur Chowdhry from the University of Greenwich and Trevor Dines from the University of Dundee for supplying the diamino acid peptides and density functional theory calculations, respectively.

I would also like to thank the numerous people I have met at conferences whose suggestions, advice and experience have been a great help.

I thank the RS·C and EPSRC for their generous funding of my PhD.

My final thanks go to my much suffering husband, who proof read lots of my work even though he had no idea what it meant and my daughters who have supported me throughout the highs and lows (of which there have been many) of my PhD. I could not have done it without their love and support and I hope I have made them proud.

Chapter 1. General Introduction

1.1. Determining Molecular Structure

A variety of techniques have been applied to determine the three dimensional (3D) structures of molecules. Typically, X-ray diffraction (Crick 1957) or Nuclear Magnetic Resonance (NMR) spectroscopy (Krushelnitsky & Reichert 2005) are used. X-ray diffraction has been the most widely used and is generally the highest resolution data but it requires the molecule to be in a crystal lattice form (Takahashi & Nakatani 1995). A beam of X-rays collide with the crystal lattice where the beam of electromagnetic radiation is then scattered by the electrons within the crystal to generate a distinctive pattern of reflections which produce a 3D image of the density of the electrons which is then observed using an X-ray detector thus, allowing the arrangement of atoms and its chemical bonds to be determined (Takahashi & Nakatani 1995). However, even when X-ray diffraction can be obtained there are some disadvantages to using this technique, as molecules such as peptides may be forced into conformations by crystal forces which deviate from those adopted in solution and the dynamic average may not be represented fully as X-ray diffraction provides a static picture of the investigated molecule (Schweitzer-Stenner, 2006). In addition, X-ray diffraction is not always suitable techniques for the determination of the 3D structure of a molecule as numerous molecules do not easily yield diffractable crystals and the preparation of crystals for analysis can be quite complicated. NMR spectroscopy can be utilised to probe conformational changes adopted in solution which occur on a picosecond timescale. NMR spectroscopy measures the spin properties of atoms such as ^{1}H and ^{13}C isotopes in molecules. NMR active nuclei, which have odd spin-quantum numbers, absorb electromagnetic

radiation when placed in a magnetic field and the resonant frequency is proportional to the strength of the magnetic field. Detailed spectra are then interpreted to determine the 3D structure of the molecule (Krushelnitsky & Reichert 2005). However, NMR spectra can be difficult to resolve as this technique encounters problems for large molecules because coupling constants cannot be used to derive a clear picture of co-existing conformations (Schweitzer-Stenner, 2006). Also NMR works on timescale which is slower than true molecular vibrations so these get averaged out in a NMR spectrum. Therefore complementary techniques for structure analysis are required and with the recent advances in the field of vibrational spectroscopy there is the potential to provide a more detailed structural picture for which NMR spectroscopy and X-ray diffraction cannot.

Vibrational spectroscopic techniques, such as infrared (IR) (Subirade *et al*. 1996; Martin *et al*. 2003; Arkin 2006) and Raman (Shashilov *et al*.2006; Ellepola *et al*. 2006), symbolize one of the most valuable tools utilised for acquiring information concerning the structure of molecules from their vibrational structure. Vibrational spectroscopies are employed in the analytical sciences for a variety of applications including, detection and identification of counterfeit medicines (de Veij *et* al. 2007), identification of substances (Smith & Dent 2005), early detection of precancerous cells (Kast *et al*. 2008), bioterrorism (Smith & Dent 2005) and biomolecular studies (Abdali *et al*. 2007; Jarvis & Goodacre 2004; Hobro *et al*. 2008) IR and Raman spectroscopies provide information on the vibrational bands found in the fingerprint region of the spectrum. IR spectroscopy involves collecting absorption information from the infra-red region of the electromagnetic spectrum and analyzing it. The frequencies at which there are absorptions of IR radiation can be correlated directly

to bonds within the molecule in question. For a molecule to absorb IR radiation a minimum of one vibrational motion must alter the net dipole moment of the molecule under analysis. Raman spectroscopies are similarly non-invasive techniques which measure molecular vibrations giving holistic fingerprints of analytes under investigation (Goodacre 2003; Smith & Dent 2005). Raman spectroscopy was first suggested by Smekal in 1923 (Smekal 1923) and then observed experimentally in 1928 by Raman and Krishnan (Raman & Krishnan 1928). Where IR involves absorption of radiation and requires a change in the dipole moment, Raman involves the measure of radiation that is scattered by each vibrational mode undergoing a change in polarizability.

1.2 Raman Theory

Raman spectroscopy measures the inelastic scattering of radiation, with scattered photons from a molecule having an altered frequency to the incident light (Colthup 1964). When monochromatic radiation of a single frequency is used to irradiate a sample, photons may be absorbed by the sample and immediately emitted as scattered radiation (Speight 2005). The dominant scattered radiation has the same frequency and wavelength as the incident radiation with no energy change occurring, and is known as 'Rayleigh scattering' or 'elastic scattering'. There is however, a small amount of the scattered radiation, approximately 1 in 10^8 photons, which are scattered at optical frequencies shifted from the incident light by the energies of molecular vibrations (Aroca 2006a; Kuldelski 2008; Smith & Dent 2005). This shift in vibrational energy may be to a higher level or lower level than for the incident

photon and is referred to as ‘anti-Stokes’ and ‘Stokes’, respectively, or ‘inelastic scattering’ (see Fig. 1.1). Stokes and anti-Stokes scattering make up the Raman Effect, where the wavelengths and intensities of the inelastic scattered light are measured. Most molecules, when at room temperature, are at their lowest energy level, referred to as the ‘ground state’ (Brown 1998; Smith & Dent 2005).

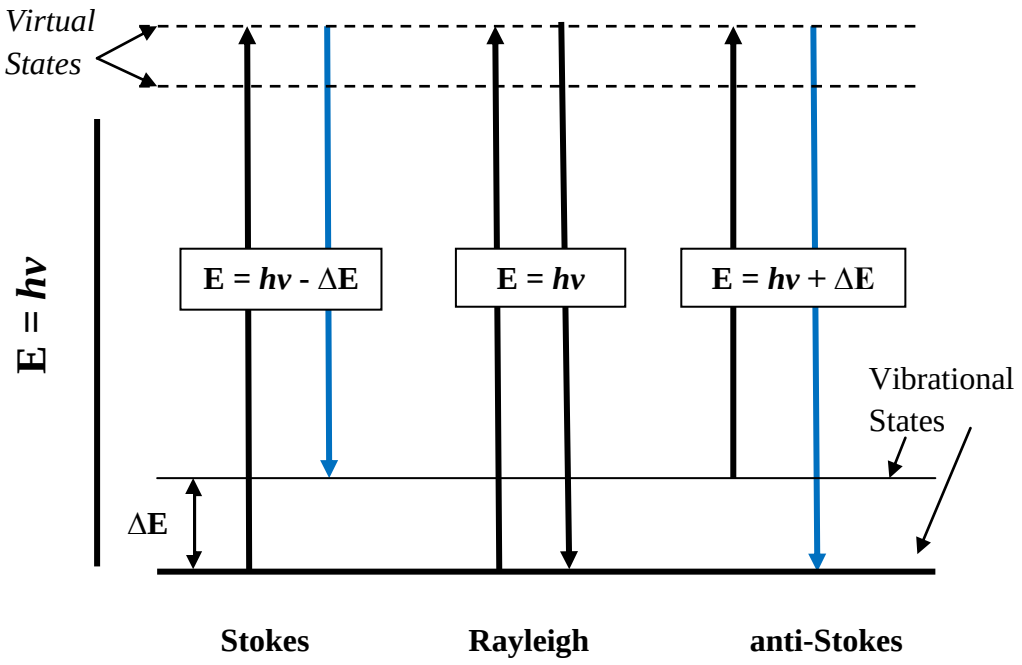


Fig. 1.1. –*The Raman Effect*. When a molecule interacts with an incident photon the increase in energy is equal to the energy of the photon ($h\nu$) where h is Plank’s constant and ν is the frequency of the radiation. In Raleigh scattering the released energy is equal to the incident photon, however, in Stokes scattering the released energy is less than the incident photon and in anti-Stokes scattering the released energy is greater than the incident photon. Adapted from Skoog *et al*. 1998.

Light scattering occurs when an oscillating dipole (light) interacts with the molecule and distorts the electron cloud surrounding the nuclei, which results in polarization (Brown 1998; Long 2002; Smith & Dent 2005; Speight 2005). This interaction promotes the sample to a higher energy state or ‘virtual state’. The virtual state is

considered an unstable complex and the energy is released immediately as scattered radiation. Furthermore, the intensity ratio of Stokes and anti-Stokes scattering is dependent upon the number of molecules present in both the ground and excited vibrational levels. Given that most molecules exist in the ground vibrational state (~ 90%) before interaction with the laser, the probability of Stokes scattering is greater than anti-Stokes scattering (Brown 1998; Long 2002; Smith & Dent 2005).

1.3 Surface-enhanced Raman scattering

Raman scattering intensities are typically weak with approximately only 1 in 10^6 to 10^8 photons being scattered thus, further enhancement is required. A technique that is widely used to increase the Raman signal by factors of 10^6 or more, is surface-enhanced Raman spectroscopy (SERS) (Moskovits, 1985; Rupérez & Laserna, 1996; Garcia-Vidal & Pendry 1996; Kniepp *et al*. 2002). SERS is a spectroscopic technique that proposes significant advances for discriminative and sensitive detection of molecules adsorbed or close to nanoscale roughened metal surfaces (Dieringer *et al*. 2005; Smith & Dent 2005). This phenomenon was first observed in 1974 by Fleischmann *et al*. (Fleischmann *et al*. 1974) for pyridine molecules adsorbed on silver electrodes. Since then SERS has been reported for a wide range of molecules on varied roughened metal surfaces with in excess of 5000 primary research articles published (Aroca 2006a; Dieringer *et al*. 2005). Nanostructured metal surfaces can be in the form of colloids, nanoshells, roughened electrodes, Langmuir-Blodgett films or silver coated beads. Typically the metals used for SERS are silver, gold and copper. The exact nature of SERS enhancement is not fully understood, however, most authors accept that are two main contributing factors,

namely the electromagnetic mechanism (EM) and the chemical mechanism or charge transfer (CT) (Moskovits 1985; Brolo 1997; Smith & Dent 2005; Aroca 2006b).

1.3.1 Electromagnetic Enhancement

The EM is considered to be the more dominant mechanism, however, the true nature for the theory of SERS is still the subject of active research (Brolo 1997; Smith & Dent 2005). The enhancement process through the EM mechanism is primarily due to surface plasmon (SP) resonance which is a result of electromagnetic radiation interaction with the metal surface which causes collective oscillation of the electrons localised on the metal surface (Rupérez & Laserna, 1996; Moskovits, 1985; Garcia-Vidal & Pendry 1996; Kniepp *et al*. 2002). As a result the surface plasmon resonance effect creates a large field-induced polarisation at the metal particle surface which increases the Raman signal. Aggregated or clustered particles have markedly different EM properties compared to isolated particles and therefore enhancement levels are approximately six orders of magnitude more than SERS enhancement on single isolated metal particles (Metiu & Das 1984; Moskovits 1985; Moskovits 2005). The greatest enhancements occur between particles and are often referred to as ‘hot spots’ which is likely due to coupling between electrons of adjacent particles which results in a new plasmon resonance generating extensive electric fields at the point of contact between the two particles (Xu & Käll 2000; Xu *et al.* 2003). A roughened metal surface is necessary for absorption and scattering to occur as smooth surfaces confine surface plasmons resulting in no excitation by the incident light or momentum mismatch (Brolo 1997). Thus, SERS is observed for molecules perpendicular to the metal surface and the magnitude of SERS enhancement

decreases as the high energy field decays rapidly with increasing distance from the surface as a function of $1/r^3$, where r is the distance from the metal surface (Moskovits 1985; Brolo 1997; Smith & Dent 2005; Smith 2008). Therefore, any molecule that is within close proximity of the metal surface will experience an increase in signal intensity compared to molecules that are not in close proximity of the metal surface.

The EM theories do not explain all the phenomena observed in SERS as they predict a SERS spectrum to be a uniform enhancement of a Raman spectrum for any Raman active molecule in close proximity to an appropriate metal surface, however, experimental studies have shown that intense bands observed in Raman spectra can appear as weak bands in the SERS spectra. Results have also shown that molecules with identical Raman cross-sections such as CO_2 and N_2 exhibit distinctly different SERS enhancement factors (Moskovits 1985) and also in the case of methanol which has a strong Raman signature but displays no further enhancement under SERS conditions (Kneipp *et al.* 2002), therefore, other mechanisms should be considered when explaining SERS enhancements and, as suggested by Otto, 'without the electromagnetic mechanism there would be no signal. But the chemical mechanism determines what is observed' (Otto 2005).

1.3.2 Chemical Effect or Charge Transfer (CT)

For charge transfer to occur the metal surface and analyte must form a bond which produces a surface complex comprising of analyte molecule and metal surface atoms allowing an adsorbate-transfer complex which results in an increase in the Raman

cross-section of the adsorbed molecule (Brolo 1997; Smith & Dent 2005). Polarizability of the molecule is increased significantly due to direct interaction between an electron from the metal surface and an incident photon. New electronic states are formed when an electron below the metal's Fermi level is promoted to a new excited state which creates an electron-hole and enhancement results from a difference in energy between the metal's Fermi level and the frontier molecular orbital as excitation is transferred from the Fermi level of the metal to the molecule and vice versa (Fukui *et al.* 1952; Brolo 1997; Smith & Dent 2005; Aroca 2006b).

The charge transfer mechanism may also result from another common approach referred to as the 'adatom-charge transfer complex model'. Raman enhancement occurs when the analyte molecules in the first monolayer are in direct contact with an atomic scale roughness (ASR) (Otto *et al.* 1980; Billmann *et al*. 1980). This mechanism is dependent upon size and roughness of the metal as areas of the metal that has a protuberance larger than 1000 Å and below 20 Å in mean diameter are ineffective, as large protuberances would suffer from multipolar resonances which are higher than the excited dipolar and smaller protuberances are ineffective for the EM mechanism to occur (Moskovits 1985; Monreal *et al*. 1987). At the ASR site electron-photon interactions are greatly increased creating more electron-hole pairs adding to the momentum available which results in an increased Raman cross-section of the adsorbed Raman active molecule.

Both the EM and chemical mechanisms play an important role in SERS enhancement. The chemical effect is considered the weaker contributor (Otto 1980; Moskovits 1985; Brolo 1997), however, to fully understand the chemical

contributions it would be useful to carry out density function theory calculations on the structure of both the analyte and colloid metal. Both mechanisms are difficult to separate experimentally and the contribution of each component is still debated (Le Ru & Etchegoin 2009) although it is considered that the EM mechanism yields the enhancement while the chemical effect changes the nature and identity of the adsorbate (Moskovits 1985). It is also believed that modification of the polarizability, i.e. the chemical effect, could originate from the electromagnetic mechanism (Le Ru & Etchegoin 2009). The SERS process is illustrated in Fig. 1.2.

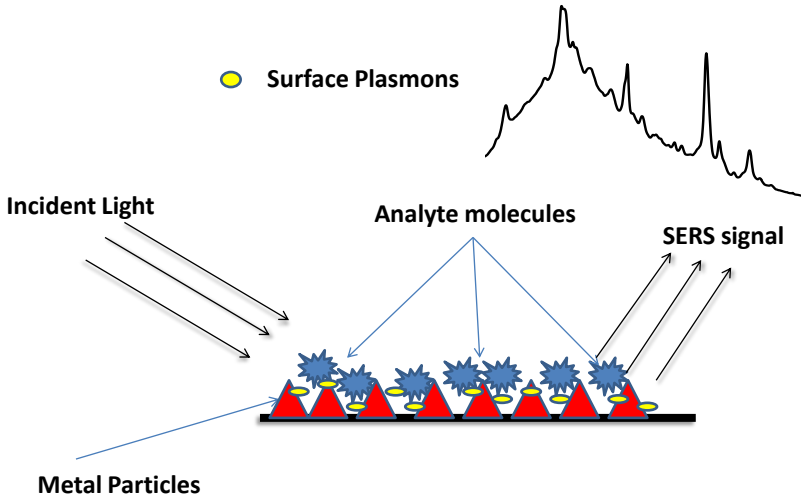


Fig. 1.2 *The SERS process* – Incident light is focussed on the metal substrate which excites the metal surface forming surface plasmons. The surface plasmon resonance effect creates a large field-induced polarisation at the metal particle creating an enhanced Raman signal.

1.3.3 Colloids

In general to generate SERS enhancement a roughened metal surface is needed and the stability of SERS substrate is paramount. Many different surfaces have been reported for SERS enhancement and these include; nanotags which are attached to biological molecules, electrode surfaces, nanoparticles, island films, colloids, core-shell particles, triangular prisms, dumbbells, hollow cubes and cold deposited films. SERS has also been measured using fine metal tips such as AFM or STM tips (Moskovits 1985; Natan 2006; Anderson 2000; Stockle 2000). In this thesis we are using metal colloids.

Colloids are homogeneous or heterogeneous particles of dispersed matter held in a solution that vary in size and shape, ranging from 10 – 10,000 Å (0.001 - 1μm) and are found in many forms; emulsion, gel, sol (solids dispersed in liquids), solid aerosol liquid, aerosol or foam. The particles are stabilised in the solution by chemical or electrochemical means to prevent the particles aggregating and precipitating out of the solution (Fleer & Lyklema 1975; Herne *et al.* 1991; Aroca 2006b). Colloids can be prepared using many different methods; however, the most common methods are laser ablation, chemical reduction and photoreduction (Aroca 2006). Chemical reduction is the most universally used method where a metal salt is reduced by a chemical agent to produce colloidal suspensions of nanoparticles (Lee & Meisel 1983; Leopold & Lendl 2003).

The size and shape (rod, star, spherical, disc) of the particles and the dielectric constant of the metal are essential physical properties of colloids as collective

resonance of electrons can be generated when coupled with the incident light which increases SERS (Homola *et al*. 1999). Further important parameters of colloids include; motion, electrical charge, optical and adsorption, electrostatic interactions, entropic forces and steric forces which contribute to the stability of the colloid (Herne *et al* 1991).

The surface charge of the colloidal particle influences the distribution of nearby ions in the polar medium, whereby opposite charged ions are attracted and opposite charged ions are repelled. This is referred to as electric double layer and is concerned with the distribution of ions and the electric potential which occurs in the locality of the charged surface. This contributes to the experimental observations and helps to understand the electrokinetic property and stability of the charged colloidal system (Shaw 1992). Energy is radiated from the surface plasmons and when colloidal particles have similar electromagnetic resonance the surface plasmons are excited and have the symmetry of a time varying dipole. The emitting molecular dipole creates resonances to any order and results in enhancement. The electromagnetic energy can be concentrated where particles are observed in close proximity to each other which results in a larger electromagnetic field creating significant enhancement. These areas are referred to 'hot-spots' (Kneipp *et al* 2002). The size and shape of the particles contribute to the hot-spot because an increase in the local optical field intensity occurs in the nanoscale gaps between the colloidal particles which rely on the field intensity of the inter-particle distance and shape (Lee & Meisel 1982; Leopold & Lendl 2003; Creighton *et al* 1976).

A useful technique used to determine whether a colloid has precipitated out of solution and no longer viable for analysis is the Tyndall effect. The Tyndall effect, also referred to as the optical effect or turbidity, is the capability of materials to scatter light. Colloidal particles are larger than atoms or molecules and can reflect scattered light producing a milky appearance in a sol, in solutions there is no obvious reflection (Radeva 1997; Shaw 1982). Another method that determines whether the particles have precipitated out of the solution is the motion effect. Particles in a colloidal system constantly move in a zigzag motion through the dispersing medium, referred to as Brownian movement (Ramirez *et al*. 1999). The Brownian movement explains the force of gravity that acts on the colloids and helps provide stability to the colloid by preventing the particles from settling out of the solution. This random motion reflects light whereas solutions do not reflect the light (Ramirez *et al*. 1999). Colloids are further stabilized when physical or chemical adsorption occurs between the metal surface and the adsorbate. Physical adsorption involves the substrate attaching itself onto the metal surface for a short period and chemical adsorption involves the formation of a bond between the metal surface and the substrate. This theory stabilizes the colloid by confining the analyte molecule in close proximity to the metal surface (Fleer & Lyklema 1975). It can also be noted that the pH of the colloid can produce different SERS response as the analyte interacts with the charge at the double layer surface of the metal particle.

The size and shape of the metal particles within the colloid can be measure using a Zetasizer nano instrument using a process referred to as Dynamic Light Scattering (DLS). DLS is used to determine the size distribution profile of small particles in suspension. When a laser hits the metal particles the light will scatter in all directions

and a time-dependent fluctuation is observed in the scattering intensity where the fluctuations are due to Brownian motion. These fluctuations are then correlated to the size of the particles. Minimum fluctuations suggest large particles and similarly if there are many fluctuations then the particles are small in size as smaller particles move faster in Brownian motion.

In order to optimize widely applicable SERS systems in this project, sols (See Fig. 1.3) were used as they are inexpensive, easy to produce and label free. The preparation methods used will include reduction of $AgNO_3$ using citrate (Lee & Meisel 1982), hydroxylamine (Leopold & Lendl 2003), borohydride ions (Creighton 1976) and reduction of $AuCl_4$ using borohydride (Lee & Meisel 1982) and citrate ions (Grabar K.C. *et al*. 1995).

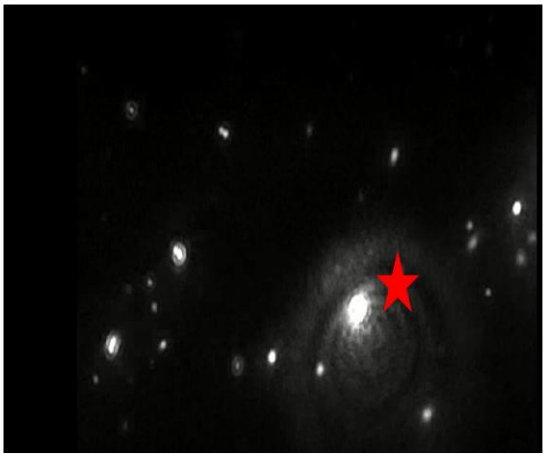


Figure 1.3 *Borohydride reduced silver colloid image.* The brighter particles are larger particles that are observed in the colloidal solution; the particle marked by the red star is extremely large and is at risk of precipitating out of the solution. The colloids were measured using Nanoparticle Tracking Analysis (NTA) Version 2.0 Test Version Build 0252 (using a laser illuminated microscopical technique, Brownian motion of nanoparticles are analysed in real-time by a CCD camera, each particle being simultaneously but separately visualised and tracked by a dedicated particle tracking image analysis programme), the results observed that ~90 % of the metal particles were at least 41 nm in size and there were 3.70 x 10^8 particles/mL (NanoSight UK).

1.4 Raman optical activity (ROA)

Determining conformational preferences of biological molecules in aqueous solution is essential for understanding their structure and folding (Grdadlonik 2008). With this in mind Raman optical activity (ROA) has proven a useful spectroscopic technique which is highly sensitive to the stereochemistry of biological and chemical molecules (Atkins & Barron 1969; Barron *et al*. 1973). ROA analyzes chiral molecules which rotate plane polarized light (Rice *et al*. 1993). Achiral molecules do not give rise to ROA signals as the difference between scattered left- and right-circularly polarized light is equal to zero.

Raman optical activity (ROA) measures vibrational optical activity by means of a small difference in the intensity of Raman scattering from chiral molecules in right- and left-circularly polarized incident light (Atkins & Barron 1969; Barron and Buckingham 1971; Barron *et al*. 1973). Chiral molecules are defined as molecules that are not superimposable on their mirror image. The difference between the right- and left-circularly polarized components of the scattered light is defined as the circularly intensity difference (CID) (Barron *et al*. 2004) and is expressed as;

$$\Delta = (I^R - I^L) / (I^R + I^L) \qquad (1)$$

where I^R and I^L are the scattered intensities in the right- and left-circularly polarized incident light. Raman scattering is equivalent to (I^R + I^L) and is the basis of normalization for the intensity of ROA scattering, (I^R - I^L) (Barron *et al*. 2007).

In Raman spectroscopy, when considering the interaction of electromagnetic radiation with matter we only need take into account the induced electric dipole (μ_α), however for optical activity we also have to consider the magnetic dipole (m_α) and the electric quadrupole moments ($\Theta_{\alpha\beta}$), which have been expressed using Cartesian tensor notation (Nafie 1997) which is a mathematical unit describing 3D Euclidean space which is determined by three orthogonal co-ordinates, *x*, *y* and *z*. The distribution of the charge of the molecule in 3D space is defined by the charge, *e*, and the distance from the origin, *r*, by;

$$\mu_\alpha = \sum_i e_i \, r_{i\alpha}, \tag{2}$$

$$m_\alpha = \sum_i \frac{e_i}{2m_i} \varepsilon_{\alpha\beta\gamma} r_{i\beta} p_{i\gamma}, \tag{3}$$

$$\Theta_{\alpha\beta} = \frac{1}{2} \sum_i e_i \left(3 r_{i\alpha} r_{i\beta} - r^2_{\delta a\beta}\right), \tag{4}$$

where particle *i* at position r_i has a charge e_i, with mass m_i and linear momentum p_i. $\varepsilon_{\alpha\beta\gamma}$ is the third-rank unit tensor, α and β are transition tensors and the Kronecker delta, $\delta_{\alpha\beta}$, is a function of two variables which is equal to 1 if they are equal and 0 otherwise (Barron *et al*. 2004).

The real oscillating electric dipole, magnetic dipole and electric quadrupole moments induced in the molecule by the real part of the electric vector (E) and associated with the magnetic vector (*B*) and electric field gradient ($\nabla_\alpha E_\beta$) in the far 'from' resonance approximation (Barron 2004; Barron *et al.* 2004) are;

$$\mu_\alpha = \alpha_{\alpha\beta}E_\beta + \frac{1}{\omega}G'_{\alpha\beta}B_\beta + \frac{1}{3}A_{\alpha\beta\gamma}\nabla_\beta E_\gamma + \cdots, \quad (5)$$

$$m_\alpha = \frac{1}{\omega}G'_{\alpha\beta}E_\beta + \cdots \quad (6)$$

$$\Theta_{\alpha\beta} = \mathrm{A}_{\gamma\alpha\beta}E_\gamma + \cdots \quad (7)$$

As ROA measurements are described in terms of CID, the CID can also be expressed in terms of the electric dipole-electric dipole molecular polarizability tensor, $\alpha_{\alpha\beta}$, and the electric dipole-magnetic and electric dipole-electric quadrupole optical activity tensors are $G'_{\alpha\beta}$ and $\mathrm{A}_{\gamma\alpha\beta}$, respectively (Barron *et al*. 2006). The CID for forward (0^o) and backward (180^o) scattering for an isotropic chiral molecule for incident wavelengths much greater than the molecular dimensions (Blanch *et al*. 2003; Barron 2004) are;

$$\Delta(0^o) = \frac{4\left[45_\alpha G'+\beta(G')^2-\beta(A)^2\right]}{c[45\alpha^2+7\beta(\alpha)^2]} \quad (8)$$

and

$$\Delta(180^o) = \frac{24\left[\beta(G')^2+\frac{1}{3}\beta(A)^2\right]}{c[45\alpha^2+7\beta(\alpha)^2]} \quad (9)$$

where, the isotropic invariants of the polarizability tensor are expressed as;

$$\alpha = \frac{1}{3}\alpha_{\alpha\alpha} \tag{10}$$

and

$$G' = \frac{1}{3}G'_{\alpha\alpha} \tag{11}$$

and, the anisotropic invariants are expressed as;

$$\beta(\alpha)^2 = \frac{1}{2}\left(3\alpha_{\alpha\beta}\alpha_{\alpha\beta} - \alpha_{\alpha\alpha}\alpha_{\beta\beta}\right) \tag{12}$$

$$\beta(G')^2 = \frac{1}{2}\left(3\alpha_{\alpha\beta}G'_{\alpha\beta} - \alpha_{\alpha\alpha}G'_{\beta\beta}\right) \tag{13}$$

$$\beta(A)^2 = \frac{1}{2}\omega\alpha_{\alpha\beta}\varepsilon_{\alpha\gamma\delta}A_{\gamma\delta\beta} \tag{14}$$

These equations apply specifically to Rayleigh scattering, for Raman scattering the same basic CID expressions apply, however the molecular property tensors would be replaced by corresponding vibrational Raman transition tensors. A simple bond polarizability theory model (Barron 2004) has shown that for a molecule composed entirely of axially-symmetric bonds, for which $\beta(G')^2 = \beta(A)^2$ and $\alpha G' = 0$, ROA is generated exclusively by anisotropic scattering, therefore the CID expressions for forward and backward scattering (equations 8 and 9) can be reduced to;

$$\Delta(0^o) = 0 \tag{15}$$

$$\Delta(180^o) = \frac{32\beta(G')^2}{c[45\alpha^2+7\beta(\alpha)^2]} \tag{16}$$

which denotes that, unlike conventional Raman scattering intensities which are identical in the forward and backward scattering directions, ROA is zero in the forward direction and maximized in the backward direction. Therefore, the ROA signal is enhanced by backscattering relative to the background Raman intensity and is an important approach for ROA studies of biomolecules in aqueous solutions due to the weak signal and high backgrounds (Blanch *et al*. 2003; Barron *et al.* 2004; Barron 2004).

The original design of ROA instruments employed the incident circular polarization (ICP) approach whereby the polarization state of the incident light between the right- and left-circularly polarization is modulated simultaneously with Raman scattering. The Raman intensities are subtracted computationally, left from right, which gives rise to the ROA acquisition (Hug 2001; Hecht *et al*. 1999). An alternative instrument design takes advantage of the scattered circular polarization (SCP) whereby the polarization of the incident light is fixed and the difference in intensity between the right- and left-circularly polarized components of the scattered light is measured (Nafie & Freedman 1993; Hug 2001). There are two further ROA instrument variations where components of the left- and right-circularly polarization states of the incident and scattered light are recorded in phase, dual circular polarization one (DCP_I), and out of phase, dual circular polarization two (DCP_{II}). These variants are not widely employed (Hug 2001; Nafie *et al*. 1995). For this thesis Raman and ROA spectra were collected using a BioTools ChiralRAMAN spectrometer (BioTools Inc., Jupiter, FL, USA) operated via Critical Link software. The instrument has a design based on the backscattering geometry using a Nd/VO_4 laser with an excitation wavelength of 532 nm and spectral resolution of ~7 cm^{-1}.

ROA is an excellent technique for studying polypeptide and protein structure in aqueous solutions and affords added information regarding the stereochemistry of the functional groups that give rise to molecular vibrations (Barron *et al* 2004; Barron 2004; Blanch *et al*. 2003). However, ROA necessitates relatively high concentration samples and longer acquisition time than conventional Raman spectroscopy as the intensity measurements are comparatively small, three to four orders of magnitude weaker, than related Raman bands (Barron *et al*. 2004).

1.5 Raman, SERS and ROA of biomolecules and pharmaceuticals

Although each molecule displays a Raman spectrum, not all vibrational modes are Raman active and the likelihood of a vibrational mode displaying a Raman cross-section is dependent upon the bond being polarized. This is an important advantage over IR, which relies on a fixed dipole moment, as the OH bond in water molecules has weak polarizability and displays a weak band in the Raman spectrum whereas a phenyl ring (C_6H_5), for example, which is ubiquitous in organic chemistry (March 1992), is highly polarizable and display a strong Raman cross-section. In IR spectroscopy the OH bond of water molecules have an intense absorbance in the mid-IR and the phenyl ring is weak. With this consideration, for biological samples where water is prevalent, IR spectroscopy can be difficult to interpret. The exact frequency and intensity of vibrational bands for IR and Raman spectroscopies are dependent upon interactions with other vibrational bands which effectively presents a unique fingerprint of the analyte under analysis (LaPlant 2010). When assigning the bands in the fingerprint region model peptides of known structures are widely used (Wallach *et al*. 1970; Lis *et al*. 1976; Sengupta & Krimm 1985) with the

fingerprint region for both IR and Raman spectroscopy allowing identification of pharmaceuticals (Hausman 2005; Fevotte 2007) and biological materials (Jarvis *et al*. 2007; Carey & Dong 2004). Raman spectroscopy is non-destructive and can be used for a variety of molecules with very little sample preparation and the lasers used in Raman can generate a signal through water, glass or any transparent material which gives Raman spectroscopy a large advantage over IR spectroscopy which absorbs most of the light through only the top microns of a sample allowing only the surface of the sample to be analysed. However, Raman spectroscopy does have some disadvantages, the biggest one being the probability of a Raman photon being generated is typically one out of a million incident photons and the Raman effect is weak. Fluorescence, from impurities in the sample and intrinsic sources, is another disadvantage as it competes with the Raman effect and can overwhelm the signal, however, advances in laser technology enables the use of different excitation wavelengths of which 785 nm can significantly depress fluorescence.

SERS as an analytical technique has been gaining popularity in physical and analytical chemistry and more recently in the biomedical field (Vo-Dinh *et al*. 2002; Yonzon *et al*. 2004). SERS was first observed in 1974 for pyridine molecules adsorbed on silver electrodes (Fleischmann *et* al. 1974) SERS can overcome the weak Raman effect as an increase in the Raman cross-section is observed in the presence of an electromagnetic field where molecules are in close proximity to metal nanostructures, typically gold or silver (Otto 1984; Campion 1998; Moskovits 1985; Persson 1981). SERS is a highly sensitive technique that can probe low concentrations, is non-destructive and requires short exposure times making SERS an ideal bioanalytical tool (Xu *et al*. 1999; Cao 2002).

Moskovits determined that small amino acids, such as alanine, interact with the metal surface through the amine and carbonyl terminals (Moskovits 1985). SERS does present some limitations with the main one being its limited reproducibility. Reproducibility from lab to lab is a difficulty that can hinder the advancement of SERS for colloidal systems. Sample preparation can differ from person to person and a slight difference can express significant variations on the spectrum analysed. More control is needed in the experimental set up for SERS and thus, an optimization process is the way forward.

ROA is an ideal technique for studying the structure and behaviour of both chemical and biological molecules (Barron *et al.* 2000; Barron *et al*. 2002; Blanch 2003) due to ROA's enhanced sensitivity to chirality which provides detailed information about the stereochemistry of bioactive molecules. A Raman spectrum can be difficult to interpret due to contributions from the backbone and side chains of proteins producing broad overlapping bands, whereas ROA normal vibrational modes are more easily interpreted as ROA is more sensitive to the rigid and chiral fragments of a molecule (Barron *et al*. 2002) which makes ROA an ideal complementary technique to conventional Raman spectroscopy as almost all biological compounds are chiral it is logical to investigate them using both techniques. ROA can also determine marker bands for isomers which are difficult to distinguish in conventional Raman spectra, whereas ROA is sensitive to absolute stereochemistry (Barron *et al* 2000). ROA does have its limitations in as such as it is a weak effect with typical Δ-values being 10^{-3} – 10^{-5} which requires extreme care with measurements (Barron *et al*. 2004), its data acquisition times are typically between hours and days and the concentration required for measurements are usually 10 –

100 mg/mL which is higher than the concentrations required for conventional Raman spectroscopy.

1.6 Machine Learning

Machine learning refers to a scientific discipline that is concerned with the development of genetic algorithms (GAs) where computers will evolve behaviours based upon empirical data to acquire increasingly better solutions to any complex problem put forward by the analyst. A computer programme is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P improves with E (Mitchell 1997). The motivation of machine learning methods, such as GAs is to make intelligent decisions efficiently in order to optimize variable of importance to the model. Evolutionary computational-based algorithms are particularly popular optimization methods (Corne *et al*. 1999; Michalewicz & Fogel 2000).

1.6.1 Genetic Algorithms (GAs)

Heuristic evolution computation methods explore search space, where exhaustive search is impractical, to improve the performance of techniques (Handl *et al*. 2007). The concept of GAs, which were developed by John Holland in the 1970s, are heuristic techniques that have been demonstrated to solve computational problems by finding a range of acceptable solutions to the problem presented (Holland 1975; Bäck *et al*. 1997; Mitchell 2002; Handl *et al*. 2007). Jarvis and Goodacre used GA

optimization for pre-processing and variable selection of spectroscopic data, demonstrating that rapid search of optimal or near optimal solutions simultaneously is achievable (Jarvis & Goodacre 2005). GAs, using a method analogous to principles of Darwinian Theory, uses natural selection to separate out individuals (individuals represent the parameters of a problem to be optimized) that are unsuitable within their environment. The fitter individuals have a greater opportunity to pass their genes onto the future generation through the process of survival of the fittest where individuals are selected for their fitness, then crossed to produce a new generation of individuals. During the process of evolution random changes can occur which may provide additional advantages in the challenge for survival. As the process is repeated the solutions become fitter until the point where the GA will eventually converge, signifying that an optimal or near-optimal solution has been achieved (Holland 1975; Goodacre & Kell 2003; Konak *et al.* 2006). The overall process of a GA is summarized in Fig. 1.4.

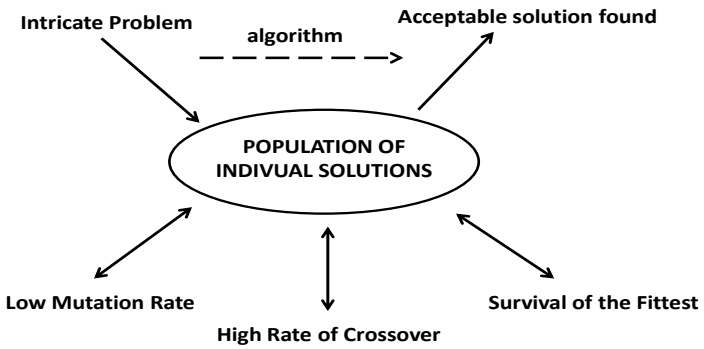


Fig. 1.4 The approach taken by a GA algorithm when faced with an intricate problem that is difficult to solve (adapted from Goodacre & Kell 2003).

For this research the measure of fitness was the number of reproducible SERS peaks generated from an analyte, however, we also had to take into account the intensity of the bands observed. Thus, as there were two objectives to be satisfied we were faced with optimizing a multi-objective problem (MOP) that required selection based on the Pareto optimal front. We therefore used a multiobjective evolutionary algorithm (MOEA) in this thesis towards optimization of SERS for propranolol. A MOEA takes into account more than one objective to be optimized and finds solutions based on the Pareto front.

1.6.2 Multi-objective optimization problem (MOP) and Pareto optimum

The aims of this project were to optimize experimental conditions for colloidal SERS. Given that the objectives under consideration could conflict with each other if, optimizing a particular solution with respect to a single objective could result in unacceptable results with respect to the other objective an acceptable compromise is to investigate a set of solutions which satisfy all objectives without any solution being dominated by another solution. MOP is concerned with decision vectors that satisfy constraints and optimize a vector function whose elements represent the objective functions. A solution is achieved when each objective is optimized to the extent that is acceptable to the decision maker and without other objectives suffering as a result if further optimization were to take place (Coello Coello *et al.* 2002; Osyczka 1985).

The general MOP $\vec{f}(x)$ can be expressed as (Coello Coello *et al.* 2002):

Find the vector; $\vec{x}^* = [x_{1,}^* x_{2,}^* \ldots\ldots, x_n^*]^T$ which will satisfy the m inequality constraints:

$$g_i(\vec{x}) \geq 0 \qquad i = 1,2, \ldots\ldots, m \tag{17}$$

which will satisfy the p equality restraints

$$h_i(\vec{x}) = 0 \qquad i = 1,2, \ldots\ldots, p \tag{18}$$

and will optimize the vector function

$$\vec{f}(\vec{x}) = [f_1(\vec{x}), f_2(\vec{x}), \ldots\ldots, f_k(\vec{x})] \tag{19}$$

In MOPs the aim is to find acceptable 'trade-offs' between solutions rather than finding one ultimate solution, therefore no solution will be selected if there is another solution that is more dominant, a notion defined by the Pareto optimum which was originally proposed by Francis Edgeworth in 1881 and later generalized by Vilfredo Pareto in 1896 (Coello Coello *et al.* 2002; O'Hagan *et al*. 2005).

The Pareto optimum is a vector of decision variables $\vec{x}^* \in \mathcal{F}$ which is expressed as Pareto optimal if there is no other $\vec{x} \in \mathcal{F}$ such that $f_i(\vec{x}) \leq f_i(\vec{x}^*)$ for all $= 1, \ldots\ldots, k$

$f_j(\vec{x}) < f_j(\vec{x}^*)$ for at least one *j*. Here, *F* symbolizes the plausible section of the problem where the constraints are satisfied (Coello Coello et al 2002).

According to this definition, the $\vec{x}^*$ is Pareto optimal when no viable vector of decision variables $\vec{x} \in \mathcal{F}$ exists that would lessen some principal factor without causing a concurrent improvement in at least one other principal factor. This notion yields a set of solutions referred to as the Pareto optimal set where the $\vec{x}^*$ are non-dominated (Coello Coello *et al.* 2002). Thus, if considering two factors such as reproducibility and intensity, the product should not be dominated by either factor.

The Pareto optimal set (*P**) for a given $\vec{f}(x)$ is defined as (Coello Coello *et al.* 2002):

$$P^* = \{x \in \mathcal{F} | \neg \exists x' \in \mathcal{F}\ \vec{f}(x') \preceq \vec{f}(x)\} \tag{20}$$

The plot of objective functions whose non-dominated vectors are in the Pareto optimal set is referred to as the Pareto front ($P\mathcal{F}^*$) and is expressed as (Coello Coello *et al.* 2002):

$$P\mathcal{F}^* = \left\{\vec{u} = \vec{f} = \left(f_1(x), \ldots\ldots, f_k(x)\right) \mid x \in P^*\right\} \tag{21}$$

GA will search the data space to determine which non-dominated vectors lie on the Pareto front and select these parameters as an objective for producing high performance colloidal systems for this project (Castillo 2006).

1.7 References

Albrecht M.G. & Creighton J.A. (1977) *J. Am Chem. Soc.* **99**: 5215-5217

Anderson M. S. (2000) *Appl. Phys. Lett.* **76:** 3130

Arkin I.T. (2006) *Curr. Op. Chem.Biol*. **10**(5): 394-401

Aroca R. (2006a) SERS publication database. www.spectroscopynow.com

Aroca R. (2006b) *Surface-enhanced Vibrational Spectroscopy*. John Wiley & Sons: Chichester

Atkins P.W. & Barron L.D. (1969) *Molecular Physics* **16**: 452-456

Bäck, T. *et al*. (1997) *Handbook of Evolutionary Computation.* IOP Publishing Oxford University Press: Oxford

Barron L.D., Bogaard M.P. & Buckingham A.D. (1973) *J. Am. Chem.* **95**: 603-605

Barron, L.D., Blanch E.W. & Hecht L. (2002) *Adv. Protein Chem.* **62**: 51-90

Barron L.D., Hecht L., Blanch E.W. & Bell A.F. (2000) *Prog. Biophys. Mol. Bio.* **73**: 1-49

Barron L.D. (2004) *Molecular Light Scattering and Optical Activity.* 2nd Edition. Cambridge University Press: Cambridge

Barron L.D., Hecht L., McColl I.H. & Blanch E.W. (2004) *Mol. Phys.* **102**: 731-744

Barron L.D., Zhu F., Hecht L., Tranter G.E. & Isaacs N.W. (2007) *J. Mol. Struc.* **834**: 7-16

Barron L.D., Zhu F. & Hecht L. (2006) *Vib. Spectrosc.* **42**(1): 15-24

Billmann J., Kovacs G. & Otto A. (1980) *Surf. Sci.* **92:** 153-173

Blanch E.W., Hecht L. & Barron L.D. (2003) *Methods* **29**: 196-209

Brolo A.G., Irish D.E. & Smith B.D. (1997) *J. Mol. Struct*. **405:** 29-44

Campion A. & Kambhampati P. (1998) *Chem. Soc. Rev* **27**: 241-250

Cao Y.W.C., Jin R.C. & Mirkin C.A. (2002) *Science* **297**: 1536-1540

Carey P.R. & Dong J. (2004) *Biochemistry* **43**: 8885-8893

Castillo F., Kordon A., Smits G., Christenson B. & Dickerson D. (2006) *GECC* **2**: 1613-1620

Coello Coello C.A., Van Veldhuizen D.A. & Lamont G.B. (2002) *Evolutionary Algorithms for Solving Multi-Objective Problems*, Kluwer Academic Publishers: Boston

Colthup N.B. (1964) *Introduction to Infrared and Raman Spectroscopy* Academic Press Inc.: London

Corne D., Dorigo M. & Glover F. (Ed.) (1999) *New Ideas in Optimization* McGraw Hill: London

Creighton J.A., Blatchford C.G., Albrecht M.G. & J. Chem. (1976) *Soc. Faraday Trans.* **75**: 790-798

Crick F.H.C., (1957) *Methods in Enzymology* **4**: 127-146

Dieringer J.A., McFarland A.D., Shah N.C., Stuart D.A., Whitney A.V., Yonzon C.R., Young, M.A., Zhang X. & Van Duyne R.P., (2005) *Faraday Discussions* **132:** 9–26.

Ellepola S.W., Choi S.M., Phillips D.L. & Ma C.Y. (2006) *Journal of Cereal Sci.* **43**(1): 85-93

Fleer G.J. & Lyklema J. (1975) *Coll. Interface. Sci.* **55:** 228-238

Fleischmann M., Hendra P.J. & McQuillan A.J. (1974) *Chem. Phys. Letts.* **26**: 163-166

Fukui K., Yoneza T. & Shingu H. (1952) *J. Chem. Phys.* **20:** 722-725

Garcia-Vidal F. J. & Pendry J. B. (1996) *Phys. Rev. Letts.***77**(6): 1163-1166

Grdadolnik J., Grdadolnik S.G. & Avbelj F. (2008) *J. Phys. Chem.* **112**: 2712-2718

Goodacre R. & Kell D.B. (2003) *Metabolic Profiling: Its Role in Biomarker Discover and Gene Function: Evolutionary Computation for the Interpretation of Metabolomic Data* ed. by Harrigan G.C. & Goodacre R. Kluwer Academic Publishers: Boston

Hamola J., Yee S.S. & Gauglitz G. (1999) *Sens. Actuators* B. **54**: 3-15

Handl J., Kell D.B. & Knowles J. (2007) *Computational Biology and Bioinformatics* **4**(2): 279-292

Hausman D.S., Cambron R.T. & Sakr A. (2005) *Int. J. Pharm* **298**: 80-90

Hecht L., Barron L.D., Blanch E.W., Bell A.F. & Day L.A. (1999) *J. Raman Spectrosc.* **30**: 815-825

Herne T.M., Ahern A. & Garrell R.L. (1991) *J. Am. Chem. Soc.* **113**: 846-854

Holland, J. (1975) *Adaptation in Natural and Artificial Systems*. Ann Arbour: University of Michigan Press

Hug W. (2001) *Chem. Phys.* **264**: 53-69

Jarvis R. & Goodacre R. (2005) *Bioinformatics.* **21:** 860-868

Jarvis R.M., Blanch E.W., Alexander A.P., Screen J. & Goodacre R. (2007) *Anal. Biochem.* **132**: 1053-1060

Jeanmarie D.L. & Van Duyne J. (1977) J. *Electroanal. Chem.* **84**: 1-20

Kneipp K., Kneipp H., Itzkan I., Dasari R.R., & Feld M.S. (2002) *J. Phys. Condens.* **14**: R597-624

Krushelnitsky A. & Reichert D. (2005) *Progress in Nuclear Magnetic Resonance Spec.* **47**(1-2): 1-25

LaPlant F. (2010) *Emerging Raman Applications and Techniques in Biomedical and Pharamceutical Fields: Lasers, Spectrographs and Detectors* Ed. by Matousek P. & Morris M.D. Springer Verlag: Berlin

Lee P.C. & Meisel D. (1982) *J. Phys. Chem*. **86**: 3391-3395

Leopold N. & Lendl B. (2003) *J. Phys. Chem.* B **107**: 5723-5727

Le Ru E.C. & Etchegoin P.G. (2009) *Principles of Surface-Enhanced Raman Spectroscopy: And Related Plasmonic Effects.* Elsevier: Amsterdam.

Lis L.J., Kauffman J.W. & Shriver D.F. (1976) *Biochimica et Biophysica Acta (BBA) – Biomembranes* **436**(3): 513-522

March J. (1992) *Advanced Organic Chemistry* 4th Edition. Wiley and Sons, New York

Martin I., Goormaghtigh E. & Ruysschaert J.M. (2003) *Biochimica et Biophysica Acta (BBA) – Biomembranes* **1614**(1): 97-103

Metiu H. & Das P. (1984) *Annual Review Physical Chem*. **35**: 507-536

Michalewicz Z. & Fogel D.B. (2000) *How to Solve It: Modern Heuristics* Springer-Verlag: Heidelberg

Mitchell, T. (1997). *Machine Learning*, McGraw Hill: London

Mitchell, M. (2002) *Introduction to Genetic Algorithms.* 8th Edition. MIT Press: London

Monreal R., Flores F., Gao Y. & López-Rios T. (1987) *Europhys. Lett.* **4**(l)**:** 115-120

Moskovits M. (1985) *Rev. Mod. Phys.* **57**(3): 783-826

Moskovits M. (2005) *J. Raman Spectrosc*. **36**: 485–496

Nafie L.A. (1997) *Annu. Rev. Phys. Chem.* **48**:357–86

Nafie L.A. & Freedman T.B. (1993) *Methods in Enzymol.* **226**: 470-482

Nafie L.A., Yu G. & Freedman T.B. (1995) *Vib. Spectrosc.* **8**(2): 231-239

NanoSight Ltd Minton Park, London Road, Amesbury, SP4 7RT, UK www.nanosight.com

O'Hagan, S., Dunn W.B., Brown M., Knowles J.D. & Kell D.B. (2005) *Anal. Chem.***77:** 290-303

Osyczka A. (1985) *Multicriteria optimization for engineering design. In Design Optimization*. J. S. Gero, Ed. Academic Press, Inc., New York, NY, 193-227

Otto A. *Light Scattering Solids IV Electronic Scattering, Spin Effects, SERS and Morphic Effects* ed. by Cardona M. & Gunthrodt G. (1984) Springer Verlag: Berlin

Otto A. (2005) *Raman Spectrosc*. **36:** 497-509

Otto A., Timper J., Billman J., Kovacs G. & Pockrand I. (1980) *Surf. Sci*. **92:** L55-L57

Persson B.N.J. (1981) *Chem. Phys. Letts.* **82**: 561-565

Radeva T. (1997) *J. Colloid Interface Sci*. **187**: 57-61

Ramirez J.A., Zinchenko A. Loewenberg M. & Davis R.H. (1999) *Chemical Engineering Science* **54**(2): 149-157

Rice K.G., Pengguang W., Brand L. & Lee Y.C. (1993) *Current. Opinion. Structural. Biol*. **3**(5): 669–674

Rupérez A. & Laserna J.J. (1996) *Surface-enhanced Raman spectroscopy, Modern Techniques in Raman Spectroscopy*, John Wiley & Sons: Chichester, UK. 227-264

Sengupta P.K. & Krimm S. (1985) *Spectrochimica Acta Part A: Molecular Spectroscopy* **41**(1-2): 205-207

Shaw D.J. (1992) *Introduction to Colloid and Surface Chemistry.* 4th Edition. Butterworth-Heinemann: Oxford, UK.

Shashilov V.A., Xu M., Ermolenkov V.V. & I. K. Lednev (2006) *J. Quan. Spec. Radiative Trans.* **102**(1): 46-61

Smith E. & Dent G. (2005) *Modern Raman Spectroscopy: A practical Approach.* Jon Wiley & Sons: Chichester UK.

Smith W.E. (2008) *Chem. Soc. Rev*. **37**: 955-964

Stockle R. M., Yung Doug S., Deckert V. & Zenobi R. (2000) *Chem. Phys. Lett.* **318:** 131

Subirade M., Salesse C., Marion D. & Pezelot M. (1995) *Biophys. J.* **69**(3): 974-988

Takahashi T. & Nakatani S. (1995) *Surface Sci.* **326**(3): 347-360

Vo-Dinh L.R., Allain D.L. & Stokes J. (2002) *J. Raman Spectrosc.* **33**: 511-516

Wallach D.F.H, Graham J.M. & Oseroff A.R. (1970) *FEBS Letts.* **7**(4): 330-334

Xu H.X., Bjerneld E.J., Kall M. & Borjesson L. (1999) *Phys. Rev Letts*. **83**: 4357-4360

Xu H.X. & Kall M. (2003) *J. Phys. Chem.* **4**: 1001-1005

Xu H.X., Aizpurua J., Kall M. & Apell P. (2000) *Phys. Rev. B.* **62**: 4318-4325

Yonzon C.R., Jeoung E., Zou S., Schatz G.C., Mrksich M. & Van Duyne R.P. (2004) *Anal. Chem.* **76**: 78-85

Chapter 2. Enhancing surface-enhanced Raman scattering (SERS) detection of Propranolol with multiobjective evolutionary optimization

This work is due to be submitted (May 2012) to the Journal of Analytical Chemistry.

Personal contribution to the paper

I was responsible for the experimental work and the writing of the paper. I contributed to the development of the multiobjective evolutionary algorithm (MOEA) which was provided by Dr. Elon Correa. I determined the experimental parameters for the experimental design and an in-house script provided by Dr. Elon Correa, specifically ANOVA was used to determine the full factorial landscape. I performed all sample preparation for, and measurement of surfaced-enhanced Raman scattering (SERS). I conducted analysis of the results in OriginPro 8 and converted the data in excel for input into the MOEA. I analyzed the results observed from the MOEA and performed the further suggested experiments in the laboratory. I performed all sample preparation for, and measurement of SERS for the practical limit of detection (LOD). The discussion and conclusions are drawn from the results, a review of the literature and discussions with Dr Ewan Blanch, Dr Elon Correa and Professor Roy Goodacre.

Chapter 2

Enhancing surface-enhanced Raman scattering (SERS) detection of Propranolol with multiobjective evolutionary optimization

Clare Levene[1], Elon Correa[2], Ewan W. Blanch[1] and Royston Goodacre[2*]

[1]Faculty of Life Sciences and [2]School of Chemistry, University of Manchester, MIB, 131 Princes Street, Manchester, M1 7DN

**Correspondence to Royston Goodacre, School of Chemistry, University of Manchester, MIB, 131 Princes Street, Manchester, M1 7DN Roy.Goodacre@manchester.ac.uk*

2.1 Abstract

Colloidal surface enhanced Raman scattering (SERS) is a complex technique where interaction between parameters such as colloid type, its concentration and aggregating agent, is poorly understood. As a result SERS has so far achieved limited reproducibility. With the development of a multiobjective evolutionary algorithm based on Pareto optimality we tested a combination of five different colloids, six different aggregating agents, a wide range of concentrations for both and three laser excitation wavelengths towards the optimization of experimental conditions for SERS, with the β-adrenergic blocker drug propranolol as the target analyte. The aim of the research was to improve enhancement and reproducibility of the SERS spectra. The objective functions chosen as being suitable for this multiobjective problem were the ratio between the half maximum and full width at the half maximum intensity for enhancement and correlation coefficient for reproducibility. To analyze a full search of all the experimental conditions 7785 experiments would have to be performed empirically; however, we demonstrate that the search for acceptable experimental conditions of SERS can be achieved using only 4 % of these possible experiments. The multiobjective evolutionary algorithm identified several experimental conditions for each objective which allowed a limit of detection of 2.4 ng/mL propranolol to be achieved, which is significantly lower (>24.5 times) than previous SERS studies aimed at detecting this β- blocker

Keywords: MOEA, SERS, fitness function, FWHM, propranolol

2.2 Introduction

Surface-enhanced Raman scattering (SERS) is a spectroscopic technique proposing significant advances for discriminative and sensitive detection of molecules adsorbed on, or in very close proximity to, nanoscale roughened metal surfaces which results in a large enhancement of the Raman signal by typically 10^4 - 10^6 (Jeanmarie and Van Duyne 1977, Brolo *et al*. 1997, Dieringer, *et al*. 2005, Smith and Dent 2005). The Raman modes that are perpendicular to the metal surface give rise to the largest enhancement which raises the potential of determining the orientation of the molecule and which functional groups are involved in adsorption (Penn *et al*. 2003). However, colloidal SERS, which is widely used across a variety of disciplines, faces a major challenge to find the optimal experimental conditions to realise the largest reproducible enhancement. Reproducibility is the degree of agreement between measurements conducted in different laboratories by different people and a less than 20% variation is considered acceptable (Natan 2006). Inconsistencies occur in SERS spectra due to the unpredicatable nature of colloidal behaviour which can result from variability in particle size and morphology and even small differences in independent sample preparation (Kerker 1984). Furthermore, localized differences can also occur in the SERS spectra influenced by the choice of metal in the colloid preparation, the selection of aggregating agent, the relative concentrations of the colloid and aggregating agent in the final mixture and the choice of laser excitation wavelength. The primary interest for improving colloidal SERS is the development of a more uniform substrate (Jackson and Halas 2004, Sabur and Gogotsi 2008), however, the importance lies more in the experimental set up rather than the reliance on a single specific colloid which could still lead to sub-optimal enhancements. Developing the

experimental conditions to determine the best substrate for an individual analyte and the ability to be able to predict the analytical performance of any substrate is the way forward in improving the reliability of SERS (Natan 2006). Studies have been carried out which have investigated the effects of experimental set up on the SERS response (Abdali 2007, Yaffe and Blanch 2007), however, more attention is required to address the reproducible optimization of experimental conditions as there has not been sufficient work done to tackle this problem. Jarvis *et al.* carried out studies on the optimization of SERS using L-cysteine as the target analyte, three different silver colloids and six different aggregating agents. Using PESA-II, an evolutionary algorithm, these authors concluded that optimizing the experimental conditions of SERS is possible with the minimum amount of experiments by implementing a computational approach (Jarvis *et al*. 2010). Moving forward from these studies, further parameters need to be included in the optimization process such as metal type, aggregating agent, concentration of colloid and aggregating agent in the final solution as well as the laser excitation wavelength. As the number of parameters increases so does the number of experiments that would need to be performed in order to effectively optimize colloidal systems by empirical means with a multiobjective problem (MOP), with the objectives being reproducibility and enhancement.

2.2.1 Multiobjective Evolutionary Algorithm

We have developed a 'Multiobjective Evolutionary Algorithm' (MOEA) based on the Pareto front (Coello Coello *et al.*, 2002; O'Hagan *et al*. 2005) providing fast and effective approximations in the major challenge of achieving improvement in

spectral reproducibility and enhancement factors. We have applied the MOEA to optimize the experimental parameters of colloidal SERS for propranolol, a lipophilic β-blocker. MOEAs have been used in previous applications towards optimization, see (Jarvis and Goodacre 2004; Gronwald *et al*., 2008; Small *et al*. 2011). The concept underlying this approach is that the MOEA will approximate the best available values of the objective functions; in the case of the desired experimental outcome for SERS the objective is to find a feasible solution that maximises both reproducibility and enhancement of the SERS signal. In this work the evolutionary algorithm used is a multiobjective genetic algorithm (Goldberg, 1989). The MOEA mimics the process of natural evolution through an iterative process which evolves towards better solutions where the experimental data are evaluated for the objectives given. The evolution process starts from a population of randomly generated individuals for which the fitness of each individual is evaluated; several individuals are then stochastically selected from the current population (with the fittest individuals having higher probability of being selected) and through the processes of crossover, mutation and survival of the fittest a new population is formed from the fittest individuals, determined from a fitness function that assesses the robustness of the model proposed by each individual. A new solution is created during each loop of a genetic algorithm (GA) until a new population has been created. The new population is then used in the next iteration where the GA will either repeat the process until a solution with a predetermined level of quality is found or when the maximum number of generations has been reached (Holland, 1975).

2.3 Materials and Methods

2.3.1 Colloid preparation and Instrumentation

Five different colloids were prepared, citrate reduction of silver and gold colloid colloid, hydroxylamine reduction of silver colloid, borohydride reduction of silver and gold colloid, all methods are available in supplementary information. All of the colloids were measured to determine their average particle size using Nanoparticle Tracking Analysis (NTA) Version 2.0 Test Version Build 0252 (NanoSight Ltd. Amesbury, UK) (see supplementary information).

Four different instruments with three different excitation laser wavelengths; 785 nm, 633 nm and 532 nm were used during the course of the study (see supplementary information).

2.3.2 Experimental Design

The β-blocker propranolol, used to treat hypertension and the only drug proven effective for the prophylaxis of migraines in children (Ludvigsson 1974), has been chosen as the model system for the evolution of new experimental protocols in which to assess the robustness of the MOEA. Interest in the field of bioanalytics, which involves quantitative measurements of drugs and their metabolites and endogenous substances in biological molecules for SERS, has increased over the years (Vankeirsbilck *et al.* 2000; Torreggiani *et al*. 2003; Dong *et al*. 2003; Ryder 2002; Bell 2000; Neville 1994; Yonzon *et al*. 2004; Qian & Nie 2008; Xie 2011). A

stock solution of 8.79 x 10^{-4} M propranolol was prepared, five different colloid types were prepared as described above and six 0.500 M stock solution of the following aggregating agents were prepared: NaCl, KCl, K_2SO_4, Na_2SO_4, KNO_3 and $MgSO_4$. Samples were prepared to the total volume of 1 mL in 1 mL clear glass sampling vials with the concentrations in the final mixture being; colloidal sol 10 – 90 % *v/v* (intervals of 10), aggregating agent solution 0 – 10 % *v/v* (intervals of 1) and propranolol solution 4.39 x 10^{-5} M (5 % *v/v*), H_2O was added to the final solution to ensure the final volume was 1 mL. Empirically, using a full factorial study would take approximately eight months (~50 experiments per day performed in triplicate) to search the true landscape fully (6 salts x 10 (non-zero) concentrations x 5 colloidal sols x 9 concentrations x 3 wavelengths plus adjustments for combinations with no salt and combinations that exceed 1 mL= 7,785 experiments). Therefore, we have refined our study and for the purpose of training the MOEA based upon practical laboratory constraints we have chosen a fractional factorial design (FFD) of experiments (Bailey 2008) to determine which experiments should be physically performed in the lab. Based on this FFD, 315 experiments were selected to be performed, assessed and subsequently serve as a test search space to train the MOEA (see Table S1 in supplementary information for the experimental conditions and Figure S1 for a full summary of the MOEA workflow). We then sorted the top 35 experiments by peak intensity for each instrument (105 experiments) and performed these experiments in triplicate in the laboratory. The MOEA then ran the solutions obtained from the 105 experiments in order to show that the MOEA would find solutions on the Pareto front and allow the trained MOEA to automatically suggest new solutions that were different from the 315 used to train the model. The solutions suggested by the trained MOEA model are designed by the algorithm alone without

any human interference. These proposed solutions were later physically performed in the lab and the results show that they are also in the Pareto front. Therefore, the MOEA can be used to automatically and independently design new and feasible experimental conditions satisfying Pareto optimality or at least acceptable for the decision maker.

2.3.3 Determining the fitness function

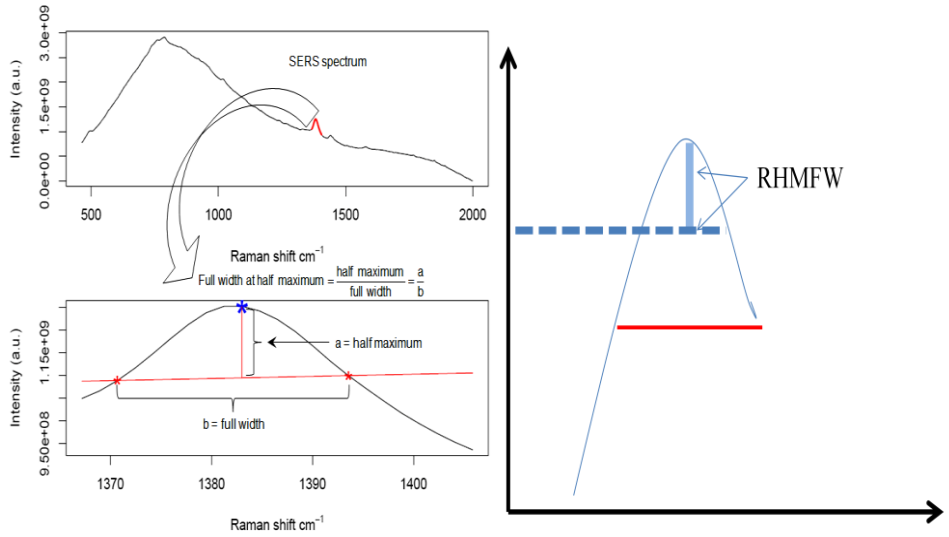
As there are two objectives to be satisfied in this research, selection was based on the Pareto optimality where MOPs are concerned with decision vectors that satisfy constraints and optimize a vector function whose elements represent the multiple (in our case two) objective functions. A solution is achieved when each objective is optimized to the extent that is acceptable to the decision maker and without other objectives suffering as a result if further optimization were to take place (see Fig. S2.1 in supplementary information) (Osyczka 1985; Coello Coello *et al*. 2002). This is termed the Pareto Optimal Front.

There are two main objectives towards the optimization of SERS: (1) spectral intensity to maximum (maximize ratio between the half maximum and full width at the half maximum *RHMFW)*; and (2) reproducibility of peak enhancement to maximum (maximize *Rep* (e_k)) which are discussed below:

1. As a measure of spectral intensity a ratio between the half maximum and full width at half maximum (RHMFW) value was chosen which is expressed mathematically as:

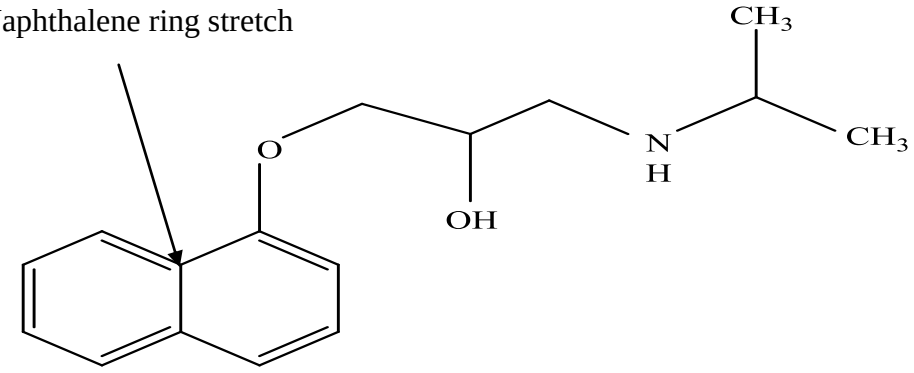
$$\text{RHMFW} = \frac{\text{half maximum}}{\text{full width}} = \frac{a}{b}, \qquad (1)$$

where a and b represent the half maximum and the full width values of the peak respectively, as Fig. 2.1 illustrates



(I)

Naphthalene ring stretch



(II)

Fig. 2.1: (I) RHMFW computation applied to SERS spectra; (II) Molecular structure of propranolol highlighting the structural origin observed of the SERS band that was optimized.

The RHMFW was considered appropriate for this fitness function (spectral intensity) as the objective is to enhance the SERS signal to its maximum, thus, the RHMFW measures the perpendicular distance between the vertex of the curve and the half maximum. The peak used to the computation of RHMFW value was between the spectral range 1367 – 1407 cm^{-1} which covers the spectral region for the well defined naphthalene ring stretch of propranolol (See Fig. 2.1) which is observed at centre ~1375 cm^{-1} (Rupérez and Laserna 1996). In order to take into account scattering and chemical variability within the spectra and compare the results fairly an extended multiplicative scatter correction (EMSC) (Martens *et al*. 2003) process was applied as a filter which removes the variance within the spectra. The EMSC method has originally been developed to reduce the problematic effect of light-scattering, which is due to the fact that small particles scatter light more than larger ones (Næs *et al.* 1990), and we have found this method to be very efficient for removing unavoidable baseline shifts. This type of normalization takes the information registered in the spectra and attempts to separate physical light-scattering effects from the actual light absorbed by molecules (Martens *et al.* 2003).

Our experiments involved three different laser excitation wavelengths (instruments) and each of these instruments is likely to generate different signal intensities even for identical experiments. The effect of the laser excitation wavelength on the analyte under study can vary considerably. Instrument variability differs from instrument to instrument, laser power at the sample has a large effect and accumulations times can be adjusted to take into account this variability. The signal to noise ratio also differs through instrument response and longer accumulation times can reduce noise and increase the signal. In an attempt to make the direct comparison of spectra obtained

by these three instruments as fair as possible, the following normalisation process was applied before the computation of the RHMFW. All spectra coming from instrument 1 are normalised so that the minimum intensity value is equal to 0 and the maximum intensity value is equal to 1. The same process is individually repeated for the other 2 instruments. This [0, 1] normalization ensures that the maximum and minimum intensity values will be similar over all instruments and that sharper peaks will tend to produce better RHMFW values than broader ones.

Other methods to measure spectral intensity, such as scaling to laser power, were considered and thus discounted as being unreliable (see supplementary information).

2. As a measure of reproducibility the average of the Pearson's correlation coefficient (Eq. 2) was implemented across the three replicated spectra for each condition. Pearson correlation (Rep) is an indication of the strength of linear relationship between two variables and can take any value between -1 and 1, thus, larger Rep values such as 0.73 indicate stronger relationships between the variables.

$$Rep\ (e_k) = \frac{1}{3}\left(\frac{\mathrm{cov}(e_{k1}, e_{k2})}{\sqrt{\mathrm{var}(e_{k1})\ \mathrm{var}(e_{k2})}} + \frac{\mathrm{cov}(e_{k1}, e_{k3})}{\sqrt{\mathrm{var}(e_{k1})\ \mathrm{var}(e_{k3})}} + \frac{\mathrm{cov}(e_{k2}, e_{k3})}{\sqrt{\mathrm{var}(e_{k2})\ \mathrm{var}(e_{k3})}} \right), \quad (2)$$

where $Rep(e_k)$ represents the reproducibility fitness function, e_{k1}, e_{k2} and e_{k3} represent replicates 1, 2 and 3 of a particular experiment e_k and var(e_{ki}) represents the variance of experiment e_{ki} (i = 1, 2, 3 and k varies from 1 to the total number of unique combinations of experiments tested).

All calculations were performed in R version 2.13.1 (2011): A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria). The scripts for the MOEA were developed in-house and are available from the authors on request.

2.4 Results and discussion

The objective functions are weighted equally due to the importance of realising that the optimal conditions for one objective may not be the optimal conditions for the other objective. Thus, the MOEA selects non-dominated solutions that sit on the Pareto (optimal) front (see Fig S2.2). By selecting non-dominated solutions the MOEA can explore the Pareto front in considerably less time and experiments than if a full factorial study was undertaken. For this research a full factorial study would have consisted of almost 8×10^3 experiments, and as we shall report we have only had to perform 315experiments in total (plus 8 experiments suggested by the MOEA to validate the model), thus reducing the total number of required experiments by 96 %.

2.4.1 MOEA Parameter Optimization

The purpose of the research was to maximize two objective functions (*viz*. reproducibility and enhancement) for the optimization of SERS for the detection and hence quantification of propranolol. The operating parameters of the MOEA can greatly affect the performance of the algorithm and for this reason in order to develop the MOEA several parameters need to be adjusted in order to demonstrate the optimal ability of the MOEA. To determine the values of the parameters the

MOEA was run several times using the fractional factorial design of experiments (315 solutions) with the target analyte propranolol. The values tested were: initial population, 100 – 1000 (intervals of 50), crossover probability 50 – 100% (intervals of 50) and mutation probability 0 – 0.05 (intervals of 0.01). By comparing the performance of the MOEA with different values for the operating parameters this gives us the ability to estimate the optimal experimental conditions in optimizing SERS for propranolol. Our empirical tests showed that the parameters that produced the best performance for the MOEA were as follows; the initial population was fixed at 100, as the solution necessitates extensive search, the initial population calls for a population large enough to do this, the probability of crossover was fixed at 1 (100%), without the crossover operator the offspring would be replicates of the parents. The mutation rate was fixed at 0.01, the mutation probability optimum value is not as critical when the crossover operator is used and is considered as the secondary operator (Maynard Smith 1978; Ochoa *et al*. 1999). On each of the MOEA runs, particular care has been taken to generate the 100 individuals for the initial population so that none of these 100 individuals was already a Pareto optimal solution. This has been done for two reasons: (1) to force the MOEA to find the Pareto optimal solutions by its own mechanisms alone; and (2) to demonstrate that MOEA will find such solutions. The stopping criterion was either, stop if convergence was reached or after 100 iterations. The overall evolutionary process is summarized in Fig. 2.2.

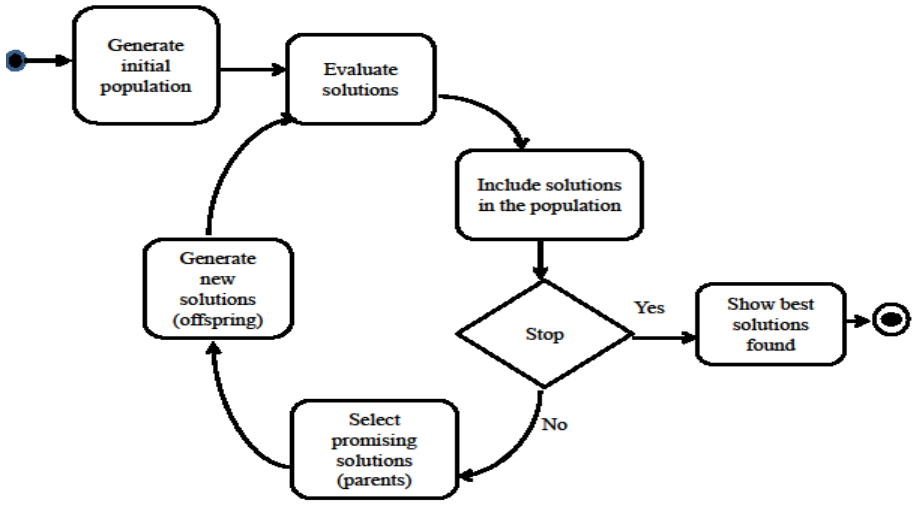


Fig 2.2: Workflow of a genetic algorithm (GA). The procedure is stopped when population converges or when a maximum number of 100 iterations is reached.

Finding the optimal experimental conditions for SERS is a difficult process and can be very dependent upon chance. Published literature has demonstrated that although it has been possible to obtain good quality results these results are not always reproducible from lab to lab. With the MOEA technique we have demonstrated that SERS experimental conditions can be optimized for an analyte in a speedily reliable process, reducing the empirical process by at least 96 %. In reality this figure could be reduced further as the MOEA has already been tuned for the performance of SERS optimization, although the parameters may have to be retuned for different analytes, which reduces the amount of experiments required through an iterative process of experiment, evaluation and evolution. With a MOEA approach it is expected that the optimization process would proceed rapidly as we have already demonstrated a reduction in the number of required experiments.

2.4.2 SERS spectra of propranolol

Triplicate experiments for the top 35 solutions of each instrument were performed (see Table S2.2) from the original 315 experimental conditions. The results were evaluated by the MOEA and Pareto front solutions were observed (Fig. 2.3). The results demonstrate that there is a trade-off when observing superior spectral enhancement at the cost of reduced spectral reproducibility and *vice-versa* (Fig. 2.4). A second generation (with 8 potential SERS solutions) of experiments was then automatically generated by the MOEA without any human interference, these solutions were outside of the FFD. All the suggested solutions were then performed in the laboratory in order to evaluate whether they gave results that were similar to the results observed on the Pareto front (see Fig. S2.4 and Table S2.3) and the results which were analyzed in OriginPro 8, corresponded well to those observed from the Pareto front. With this consideration the results that are presented in the middle of the Pareto front give an acceptable compromise for both fitness functions when determining which experimental conditions work towards the optimization of SERS for propranolol. The experimental conditions,RHMFW and reproducibility values for the results on the Pareto front are shown in Table S2.4 (see supplementary information).

Extensive spectra were collected for propranolol during this study where several different experimental conditions were applied. Observable patterns were noted; hydroxylamine reduction of $AgNO_3$ with a sulphate aggregating agent produced improved results for reproducibility and citrate reduction of $HAuCl_4$ with a sulfate aggregating agent produced improved results for intensity enhancement. However,

citrate reduction of $HAuCl_4$ with a chloride aggregating agent had the tendency to produce results in the middle of the Pareto front which were the acceptable trade-off results for the two fitness functions (see Fig. 2.3). The laser excitation wavelength that produced the best overall results was 785 nm. As to why the different experimental conditions produce different results we cannot say as the chemistry of colloids is extremely complex and thus too difficult to model accurately, moreover the interaction between aggregating agents, metal surfaces and analytes will impact upon the conformation of propranolol with respect to the colloid surface. A further observable trend that was noted during the study is that samples prepared without an aggregating agent can produce results that differ considerably from sample to sample, apart from samples prepared using citrate reduction of $HAuCl_4$ which appear to produce consistent results (see supplementary information Fig. S2.5), with a large enhancement, although the intensity value can vary by up to one order of magnitude.

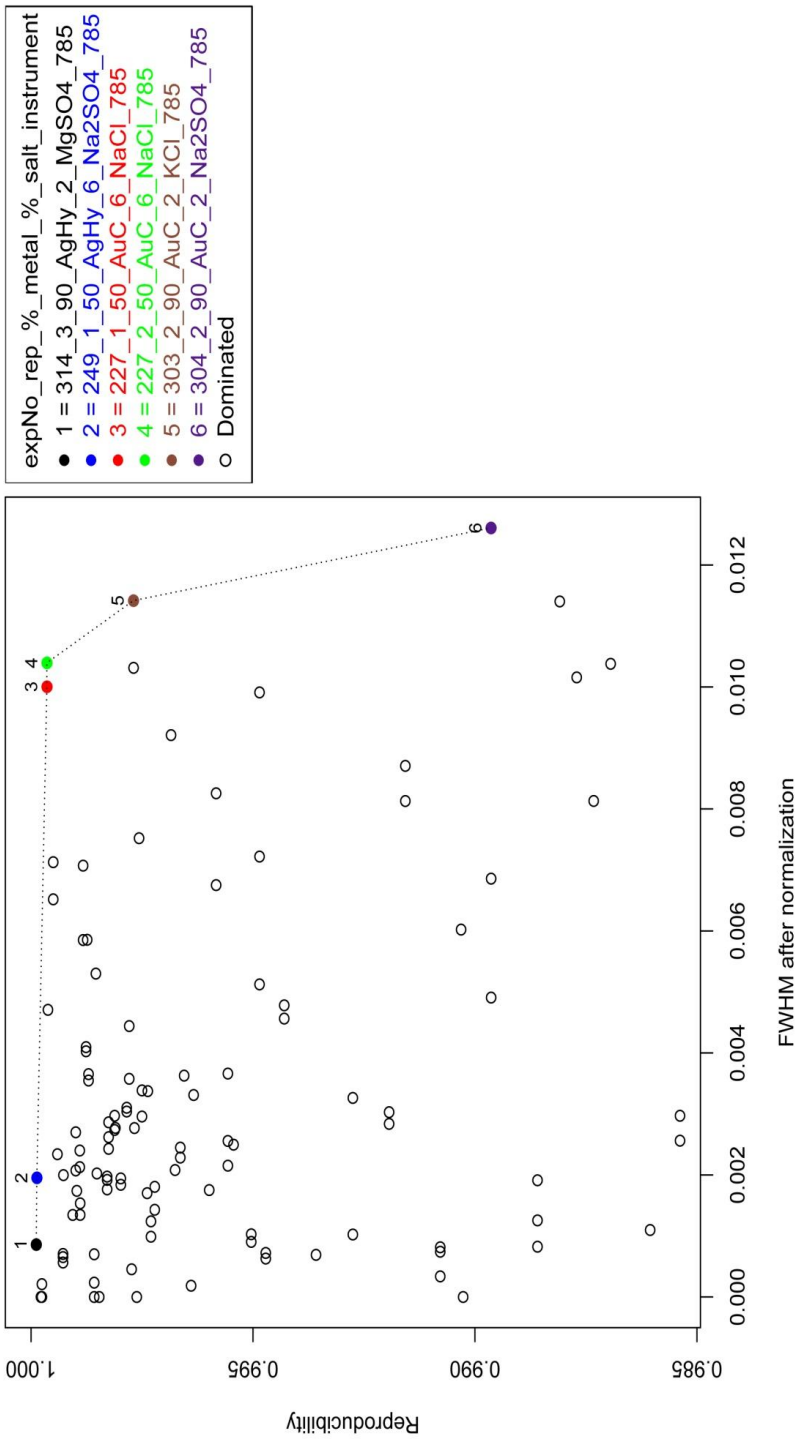


Fig. 2.3: Pareto optimal solutions. The solutions in filled circles are those that are highlighted in the box and represent SERS solutions on the Pareto front.

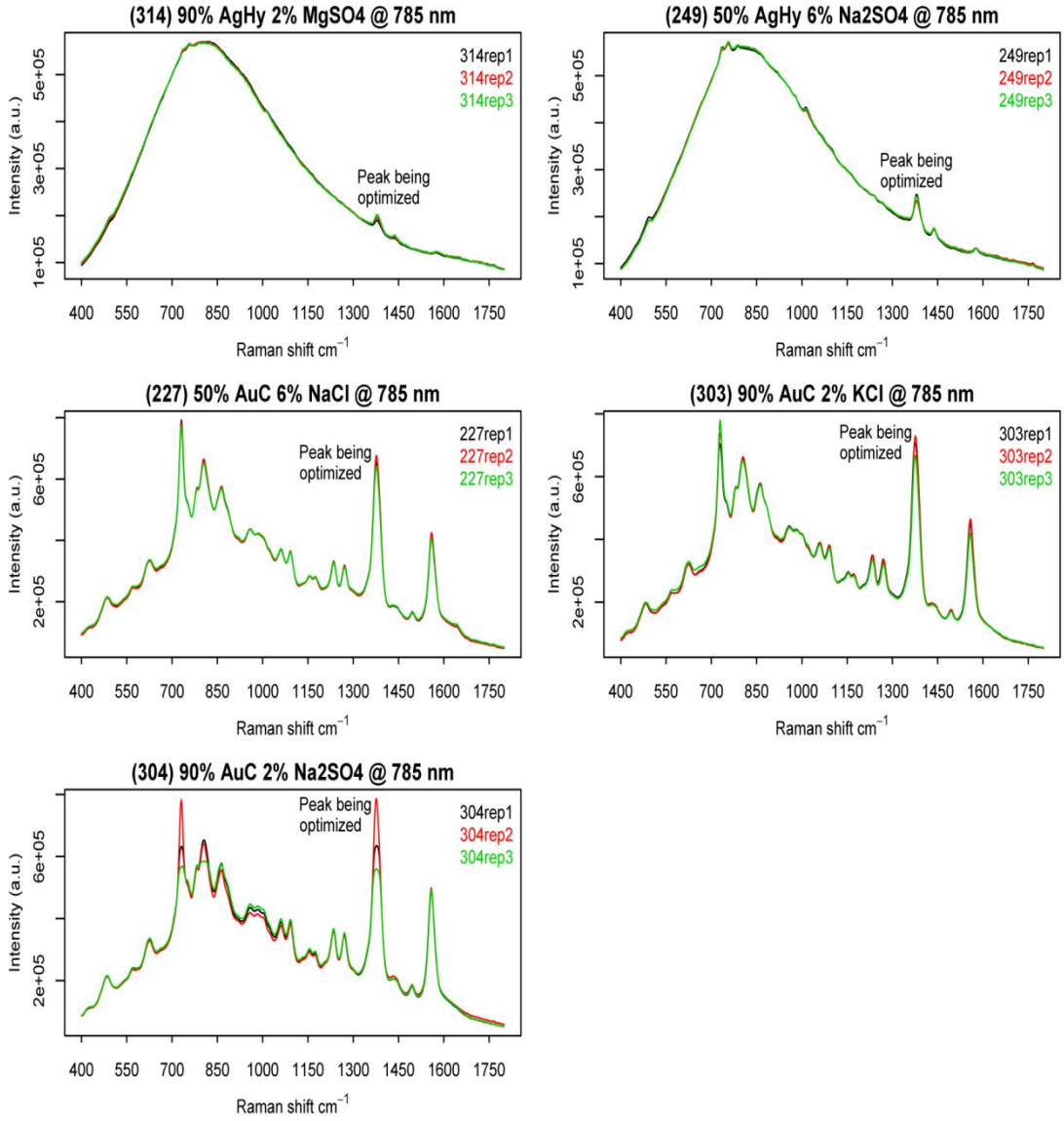


Fig.2.4: Examples of observed results from the Pareto Optimal Front (samples identified in Fig. 3). Experiments 227, 303 and 304 demonstrate the greatly improved enhancement with less reproducibility than experiments 314 and 249. Experiments 227 and 303 sit in the middle of the Pareto front and are acceptable solutions when considering both fitness functions.(The naphthalene peak is observed between the spectral range 1367 – 1407 cm^{-1} and this was the spectral area that we compared only).

2.4.3 Practical limit of detection

A significant application of SERS relates to the practical limit of detection (LOD) and thus the usable quantitative detection of the propranolol analyte. The LOD is defined as the concentration of the analyte that is required to produce an instrument response that is three times as large as the standard deviation of the noise level ($\sigma \geq 3$) (MacDougall & Crummett 1980). The experimental conditions observed on the Pareto front for experiment 303 (90% citrate reduction of $HAuCl_4$ solution, 2% 0.5M KCL solution, 5% 8.79 x 10^{-5} M propranolol solution and 3% dH_2O) were used as a starting concentration for LOD measurements. Ten concentrations were measured between 1300 ng/mL and 0.13 ng/ml (see Fig. 2.5) and these experiments were performed in triplicate. The peak area for the naphthalene peak was measured and the mean area was determined for each concentration. From these values the standard deviation was determined and the LOD was found to be 2.4 ng/mL (7.97 x 10^{-9} M), which is well within physiological concentrations in human blood. Moreover this level is comparable with previous studies which have been shown using gas chromatographic assays and high performance liquid chromatography. (Pine et al. 1975; Majahan et al. 1984; Murillo Pulgarín *et al*. 1998; Mullet *et al*. 2001). Previous studies of propranolol using SERS found the LOD to be 59 ng/mL using citrate reduction of $AgNO_3$ (Pérez *et al*. 1998). The fact that our LOD is 24.5 fold lower than this and within physiologically relevant concentrations demonstrates the importance of optimizing experimental conditions in order to achieve the optimal or near optimal results. Moreover, LOD is strongly dependent on accumulation time and as we performed the experiments over 23s there is the potential to reduce this LOD further.

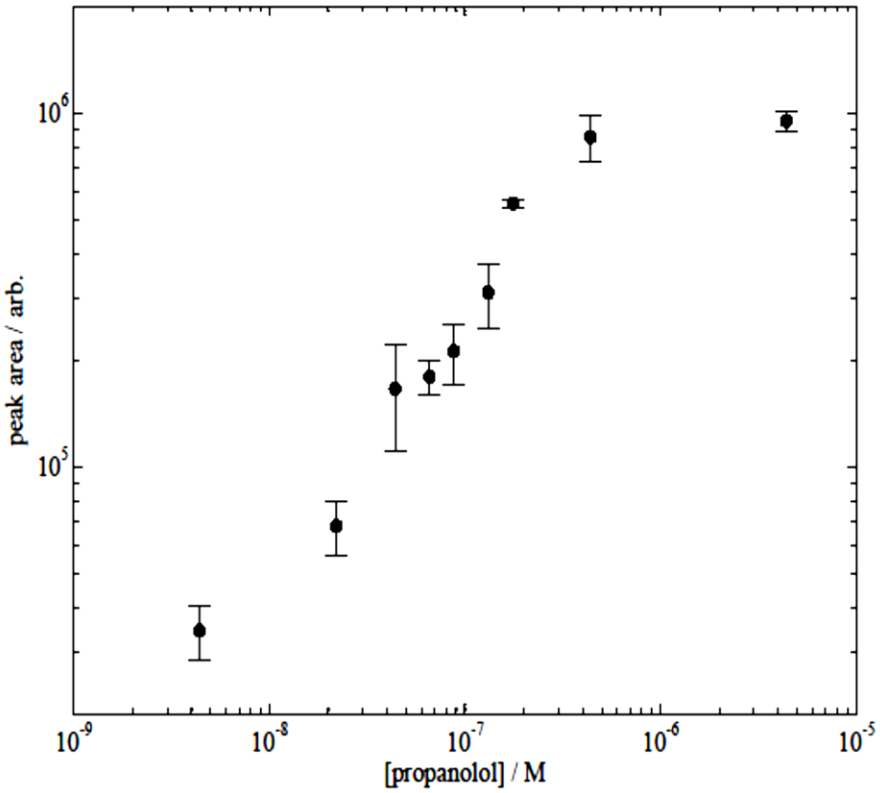


Fig. 2.5: Calibration curve illustrating the LOD experiments

2.5 Conclusion

Currently the best SERS experimental conditions for a particular analyte lack any definite plan and sub-optimal search strategies are used. These typically involved optimizing a single parameter at a time and this is not a very good way to search a multi-dimensional problem space where it is the combination of many parameters that are important. Although these single parameter optimizations may achieve reasonably good results in general these cannot be reproduced by other laboratories. With the use of multiobjective evolutionary algorithms and in particular the Pareto front computer design used here, and by selecting the solutions that have produced

middle range values for both objective functions significant improvements have been made. We have demonstrated a reduction of ~96 % in the number of experiments that would have been required by performing a full factorial search and we have also demonstrated that the solutions on the Pareto front are reproducible and are 24.5 times lower in detection limit than previously published values for this analyte.

We believe that this multiobjective evolutionary algorithm strategy for experimental design is a significant move towards the optimization of colloidal SERS which are still widely used for quantitative studies and can be coupled with relatively low cost portable Raman spectrometers. We have demonstrated the potential of MOEA to a real world problem using a lipophilic analyte with important pharmacological significance and have achieved desirable solutions in a fraction of the time it would have taken using empirical methods. We have also presented a LOD using results observed on the Pareto front and detected concentrations that are well within physiological concentrations. Thus it is our belief that taking this computational approach forward it will be possible to optimize experimental conditions of colloidal SERS for a variety of important analytes.

Acknowledgements

The authors thank the Royal Society of Chemistry and the Engineering and Physical Sciences Research Council for funding CL's Analytical Trust Fund PhD Studentship. EC and RG also acknowledge the EU Commonsense (www.fp7projectcommonsense.eu/) project (Grant 261809) financed by the European Commission under the 7th Framework Programme for Research and Technological Development.

2.6 References

Abdali S. (2007) *J. Phys. Condens. Matter* **19**: 5194-5205

Bailey R. A. (2008) *Design of Comparative Experiments. Cambridge Series in Statistical and Probabilistic Mathematics.* Cambridge University Press, New York

Bell S.E.J., Burns D.T., Dennis A.C. & Speer J.S. (2000) *Analyst* **125**: 541-544

Brolo A.G., Irish D.E. & Smith B.D., (1997) *J. Mol. Struct*. **405:** 29-44

Coello Coello, C.A., Van Veldhuizen D.A. & Lamont G.B. (2002) *Evolutionary Algorithms for Solving Multiobjective Optimization Problems*, Kluwer Academic Publishers: Boston

Creighton J.A., Blatchford C.G., Albrecht M.G. & J. Chem. (1976) *Soc. Faraday Trans.* **75**: 790-798

Dieringer J.A., McFarland A.D., Shah N.C., Stuart D.A., Whitney A.V., Yonzon C.R., Young, M.A., Zhang X. & Van Duyne R.P. (2005) *Faraday Discussion* **132:** 9–26

Dong J., Wan Z., Popov M., Carey P.R. & Weiss M.A. (2003) *J. Mol. Biol.* **330**: 431-442

Gronwald W., Hohm T. & Hoffmann D. (2008) *BMC Bioinformatics* **9**:109-120

Grabar K.C., Freeman R.G., Hommer M.B.& Natan N.J. (1995) *Anal. Chem.* **67**: 735-743

Handl J., Kell D.B. & Knowles J. (2007*) IEEE-ACM Trans. Comput. Biol. Bioinf.* **4**(2): 279-292

Holland J. (1975) *Adaptation in Natural and Artificial Systems*. Ann Arbour: University of Michigan Press

Jackson J.B. & Halas N.J. (2004) *PNAS. USA* **101**: 17930–17935

Jarvis R.M. & Goodacre R. (2004) *Bioinformatics* **21**: 860-868

Jarvis R.M., Rowe W., Yaffe N.R., O'Connor R., Knowles J.D., Blanch E.W., Goodacre R. (2010) *Anal. Bioanal. Chem.* **397**: 1893–1901

Jeanmarie D.L. & Van Duyne J. (1977) *J. Electroanal. Chem.* **84**: 1-20

Ke W., Feng X. & Huang Y. (2011) *J. Appl. Phys*. **109**: 083526-1-083526-6

Kerker M. (1984) *Acc. Chem. Res.* **17**: 271-277

Lee P.C. & Meisel D. *J. Phys. Chem.* (1982) **86**: 3391-3395

Leopold N. & Lendl B. *J. Phys. Chem. B* (2003) **107**: 5723-5727

Ludvigsson J. (1974) *Acta Neurol. Scandinav*. **50**: 109-115

MacDougall D. & Crummett W.B. (1980) *Anal. Chem.* **52**(14): 2242–2249

Majahan P., Grech E. D., Pearson R. M., Ridgway E. J. & Turner P. (1984) *Br. J. Clin. Pharmacol.* **18**: 849-852

Martens H., Nielsen P. J. & Engelsen S. B. (2003) *Anal. Chem.* **75**: 394–404

Maynard Smith J. (1978) *The Evolution of Sex* Cambridge University Press, Cambridge, UK

Mullett W.M., Martin P. & Pawliszyn J. (2001) *Anal. Chem*. **73**: 2383-2389

Murillo Pulgarín J.A., Molina A.A. & López P.F. (1998) *Anal. Chim. Acta* **370**(1): 9-18

Næs T., Isaksson T. and Kowalski B. (1990) *Anal. Chem*. **62**: 664-673

Natan M.J. (2006) *Faraday Discussion.* **132**: 321-328

Neville G.A., Beckstead H.D. & Shurvell H.F. (1994) *J. Pharm. Sci.* **83**: 143-151

Ochoa G., Harvey I. *et al.* (1999) *Proceedings of the Eighth Genetic and Evolutionary Computation Conference (GECCO),* Morgan Kaufmann, Orlando, FL, USA 488–495

O'Hagan, S., Dunn W.B., Brown M., Knowles J.D. & Kell D.B. (2005) *Anal. Chem.* **77:** 290-303

Osyczka A. (1985) *Multicriteria Optimization for Engineering Design. In Design Optimization*. J. S. Gero, Ed. Academic Press, Inc., New York, NY, 193-227

Penn S.G.,.He L & Natan M.J. (2003) *Curr. Opin. Chem. Biol.* **7**: 609-615

Pine M., Favrot L., Smith S., McDonald K & Chidsey C.A. (1975) *Circulation* **52**: 886-893

Pérez R., Rupérez A. & Laserna J.J. (1998) *Anal. Chim. Acta* **376**(2): 255-263

Qian X.M. & Nie S.M. (2008) *Chem. Soc. Rev.* **37**: 912-920

Rupérez, A., Laserna, J.J. (1996) *Anal. Chim. Acta* **335**: 87–94

Ryder A.G. (2002) *J. Forensic Sci.* **47**: 275-284

Sabur A. & Gogotsi M. (2008) *J. Raman Spectrosc*. **39**: 61–67

Small B.G., McColl B.W., Allmendinger R., Pahle J., Lopez-Castejon G., Rothwell N.J., Knowles J., Mendes P., Brough D. & Kell D.B. (2011) *Nature Chem. Biol*. **7**: 902-908

Smith E. & Dent G. (2005) *Modern Raman Spectroscopy: A practical Approach.* Jon Wiley & Sons: Chichester UK

Torreggiani A., Tamba M., Bonora S. & Fini G. (2003) *Biopolymers*. **72**: 290-298

Vankeirsbilck T., Vercauteren A., Baeyens W., Van der Weken G., Verpoort F., Vergote G. & Remon J.P. (2002) *Trends Anal. Chem.* **21**: 869-877

Washer G.A., Brooks T.M.B. & Saulsberry R. (2007) *ICONIC Third International Conference on Electromagnetic Near-Field Characterization and Imaging,* St. Louis, MO, USA

Xie W., Qiu P. & Mao C. (2011) *J. Mater. Chem.* **21**: 5190-5202

Yaffe N.R. Blanch E.W. (2007) *Vib. Spectrosc.* **2**: 196–201

Yonzon C.R., Nilam O.L, Shah C., Dieringer J.A. & Van Duyne R.P. (2004) *Anal. Chem.* **76**: 78-85

2.7 Supplementary Information

Colloid preparation

$AgNO_3$ and $HAuCl_4$ were purchased from Sigma Aldrich UK, all salts were purchased from Avocado Research Chemicals Ltd and propranolol hydrochloride was purchased from Fisher Scientific UK. All reagents were used without further purification.

Five colloids were prepared as follows:

1. Citrate reduction of silver colloid (Lee and Meisel 1982). 90 mg $AgNO_3$ was dissolved in 500 mL of deionised H_2O (dH_2O). The solution was then heated until boiling, stirring continuously. Following this 10 mL of a 1% sodium citrate solution was added dropwise and stirred vigorously. The solution was then left to boil for a further 60 min and a colour change from clear to grey-green was noted. The solution was then stored in a dark cupboard until use.

2. Hydroxylamine reduction of silver colloid (Leopold and Lendl 2003). 52 mg $AgNO_3$ was dissolved in 270 mL of dH_2O, 63 mg of hydroxylamine hydrochloride was dissolved in 15 mL dH_2O and 13.5 mL of 0.10 M of NaOH was added dropwise. The two solutions were then combined and a grey-brown solution resulted. The solution was stored in a dark cupboard until use.

3. Borohydride reduction of silver colloid (Creighton *et al.* 1976). 14 mg $NaBH_4$ was dissolved in 300 mL dH_2O to which 38 mL of a 2 x 10^{-3} M $AgNO_3$ was added dropwise with continuous stirring where the colour changed from clear to yellow. The resulting solution was stirred for a further 45 min and kept in an ice bath at all

times before storage at 4 °C. During analysis this colloid was kept in an ice bath to maintain its temperature.

4. Borohydride ($NaBh_{4)}$ reduction of gold colloid (Lee and Meisel 1982). 49 mg $HAuCl_4$ was dissolved in 25 mL of cooled dH_2O and 6 mg $NaBh_4$ was dissolved in 75 mL of cooled dH_2O. The two solutions were added together and a red colour was noted. The colloid was stored at 4 °C. During analysis the colloid was kept in an ice bath to maintain its temperature.

5. Citrate reduction of gold colloid (Grabar K.C. *et al*. 1995). 97 mg $HAuCl_4$ was dissolved in 200 mL dH_2O and boiled. 6 mL of a 1% sodium citrate solution was then added dropwise while boiling. The solution was left boiling for a further 30 min with continuous stirring and a dark red colour was noted. After cooling the colloid was kept in a dark cupboard until use.

Instrumentation

SERS spectra were recorded on four different instruments:

1. NIR Advantage Series Raman spectrometer (DeltaNu, Laramie, WY, USA) emitting up to ~ 60 mW of 785 nm radiation with 23 s integrations;

2. Advantage 200A Series Raman spectrometer (DeltaNu, Laramie, WY, USA) emitting up to ~ 3 mW of 633 nm radiation with 23 s integrations. Acquired data for the above two instruments was saved in the NuSpec ASCII XY format;

3. Ocean Optics spectrometer, (Dunedin, FL, USA) emitting up to ~ 77 mW of 532 nm radiation with 23 s integrations, for which the data was saved as csv format via RSI Scan software;

4. ChiralRAMAN spectrometer (BioTools Inc., Jupiter, FL, USA) emitting up to ~ 500 mW of 532 nm radiation with 32 scans (23 s integrations) measured, where the data were saved as ASCII txt format via Critical Link LLC software.

Spectra were measured in triplicate for all four instruments. The data from all four instruments were exported to a PC running Windows 7 and converted to xls format and csv format where required, for analysis and data processing.

Scaling factors for comparison of intensities.

During the study we had to consider how best to compare the results from each laser excitation wavelength impartially. In order to do this we considered different scaling factors and present the reasons for our choice:

Initially, data were analyzed without any scaling factor being applied. However, it was clear that the results observed overwhelmingly favoured the ChiralRAMAN spectrometer. Presumably this was due to the laser power at the sample compared to the Raman probes which emitted radiation at least eight times lower intensities than the ChiralRAMAN spectrometer.

Taking this important difference forward we scaled the data by multiplying the data observed for the Advantage 200A Series spectrometer and the NIR Advantage Series spectrometer (the Ocean Optics spectrometer was not used at this point) by appropriate scaling factors to normalize them to the intensities measured using the ChiralRAMAN spectrometer. The SERS intensities measured using the Advantage

200A spectrometer were scaled by a factor of 8.33, and the SERS intensities measured using the NIR Advantage instrument by 166.67. However, the results then greatly favoured the Advantage NIR spectrometer, indicating that this scaling factor was inappropriate.

Scaling factors were then determined based on effective scattering intensities measured on each spectrometer for a reference compound, pure ethanol. This yielded scaling factors of 5200 for the 785 nm excitation wavelength and 285 for the 633 nm laser excitation wavelength. However, these scaling factors again led to favouring of the Advantage NIR spectrometer.

In order to remove this uncertainty over appropriate scaling factors we therefore scaled the data to between 0 and 1 for all instruments, 0 being the lowest point observed for the spectrum. This negates the effect of variable instrument parameters (such as excitation wavelength, quantum efficiency of the detector and the efficiency of the optical alignment) which can dramatically differ from spectrometer-to-spectrometer and so makes our approach applicable to any SERS experiment.

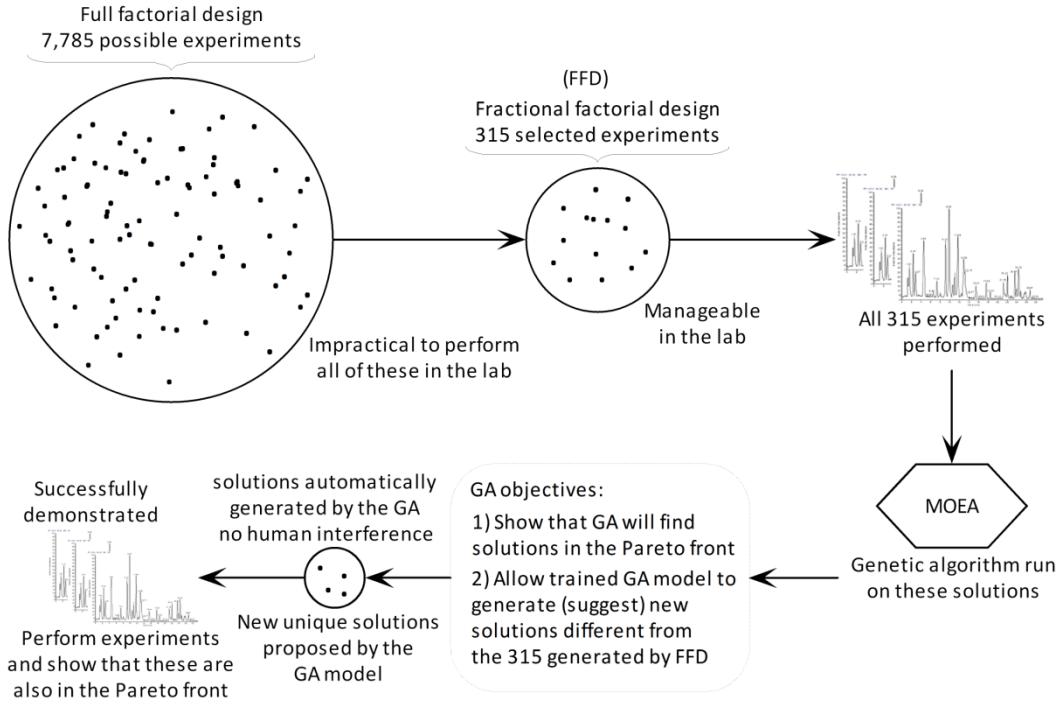


Fig. S2.1 A summary of the MOEA workflow

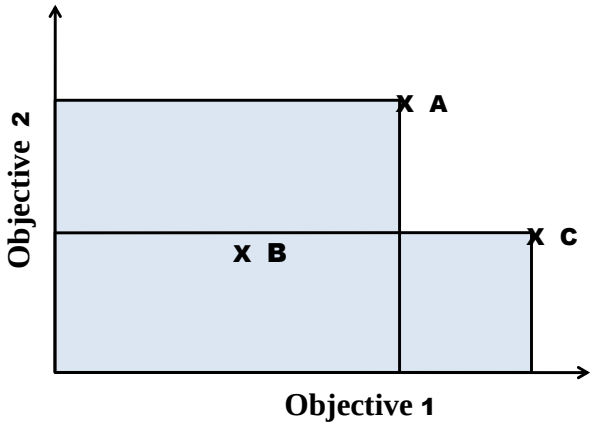
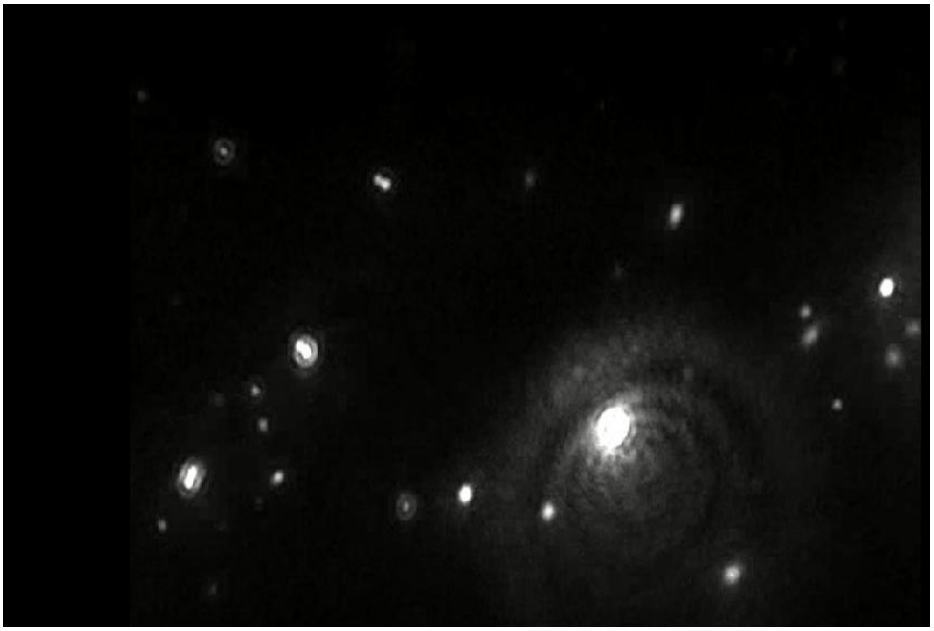
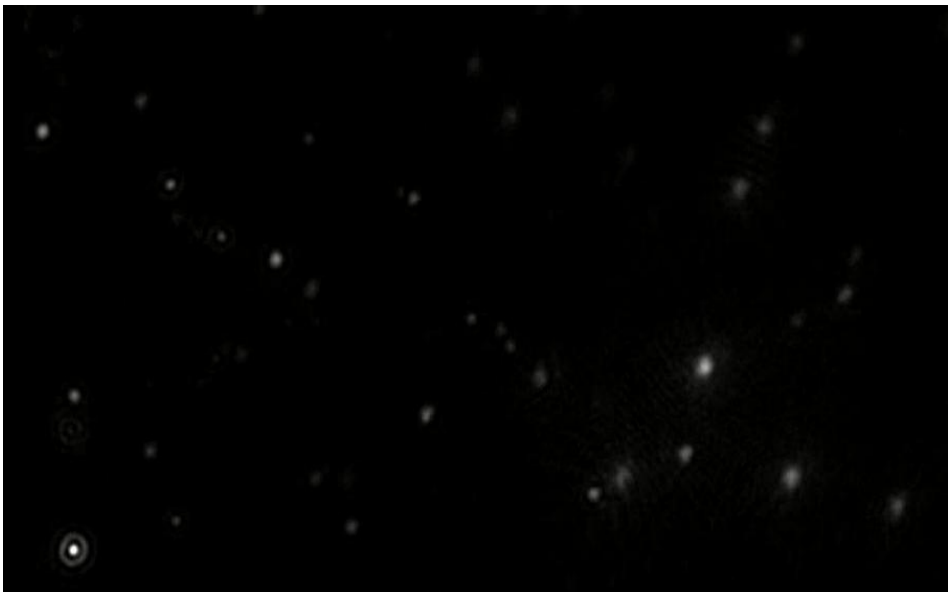


Fig. S2.2: Pareto optimality for two objective functions - Point A dominates point B in both objectives but only dominates point C in objective 2. Therefore as no solution dominates all the others in both objectives, solutions A and C would both be acceptable solutions that sit on the Pareto Front

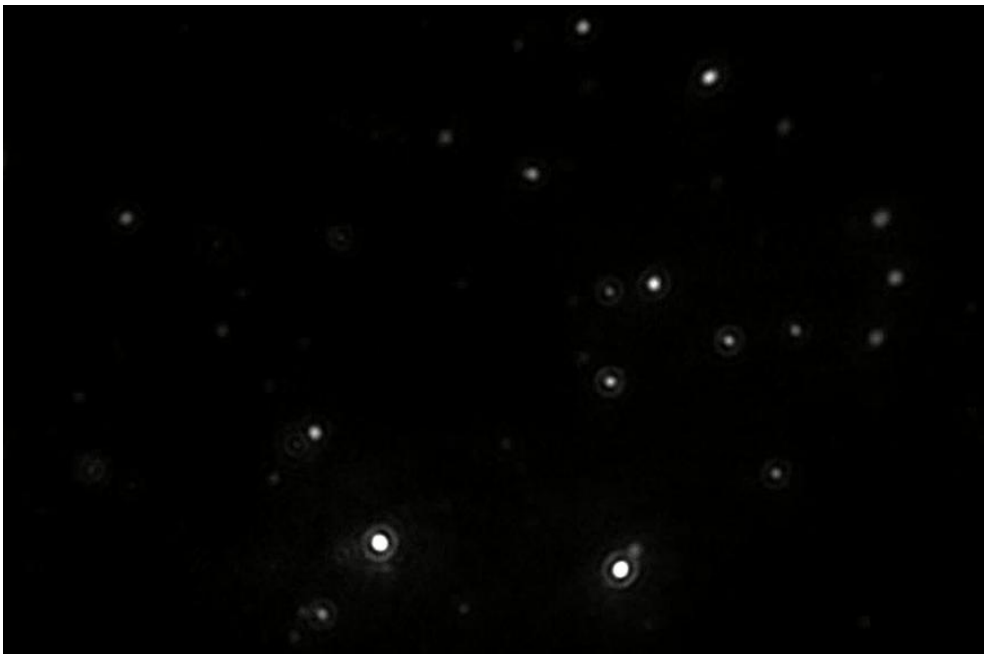


(I)

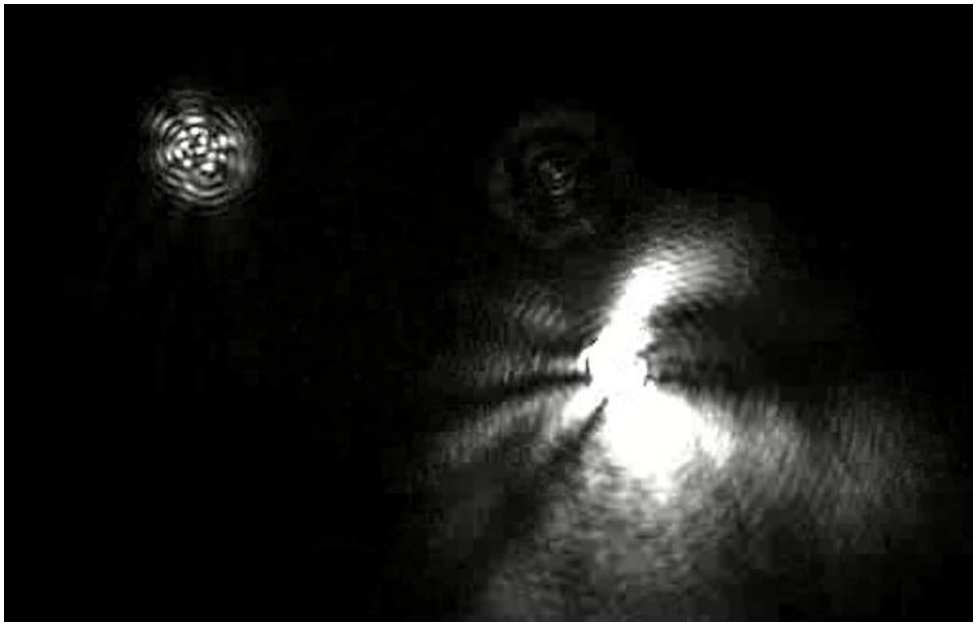
(II)



(III)



(IV)



(V)

Fig. S2.3: Representative images of colloids measured using a Nanoparticle Tracking Analysis (NTA) Version 2.0 Test Version Build 0252 (NanoSight Ltd. Amesbury, UK), in order to characterize colloid size and density in the sol:

(I) Borohydride reduction of $AgNO_3$, ~90 % of the metal particles were at least 41 nm in diameter and there were 3.70 x 10^8 particles/ml.

(II) Citrate reduction of $AgNO_3$, ~90% of the metal particles were at least 76 nm in diameter and there were 3.87 x 10^8 particles/ml.

(III) Citrate reduction of $HAuCl_4$, ~90% of the metal particles were at least 81 nm in diameter and there were 4.94 x 10^8 particles/ml.

(IV) Borohydride reduction of $HAuCl_4$, ~90% of the metal particles were at least 57 nm in diameter and there were 5.56 x 10^8 particles/ml.

(V) Colloids produced by hydroxylamine reduction of $AgNO_3$ were too large to be measured using NTA.

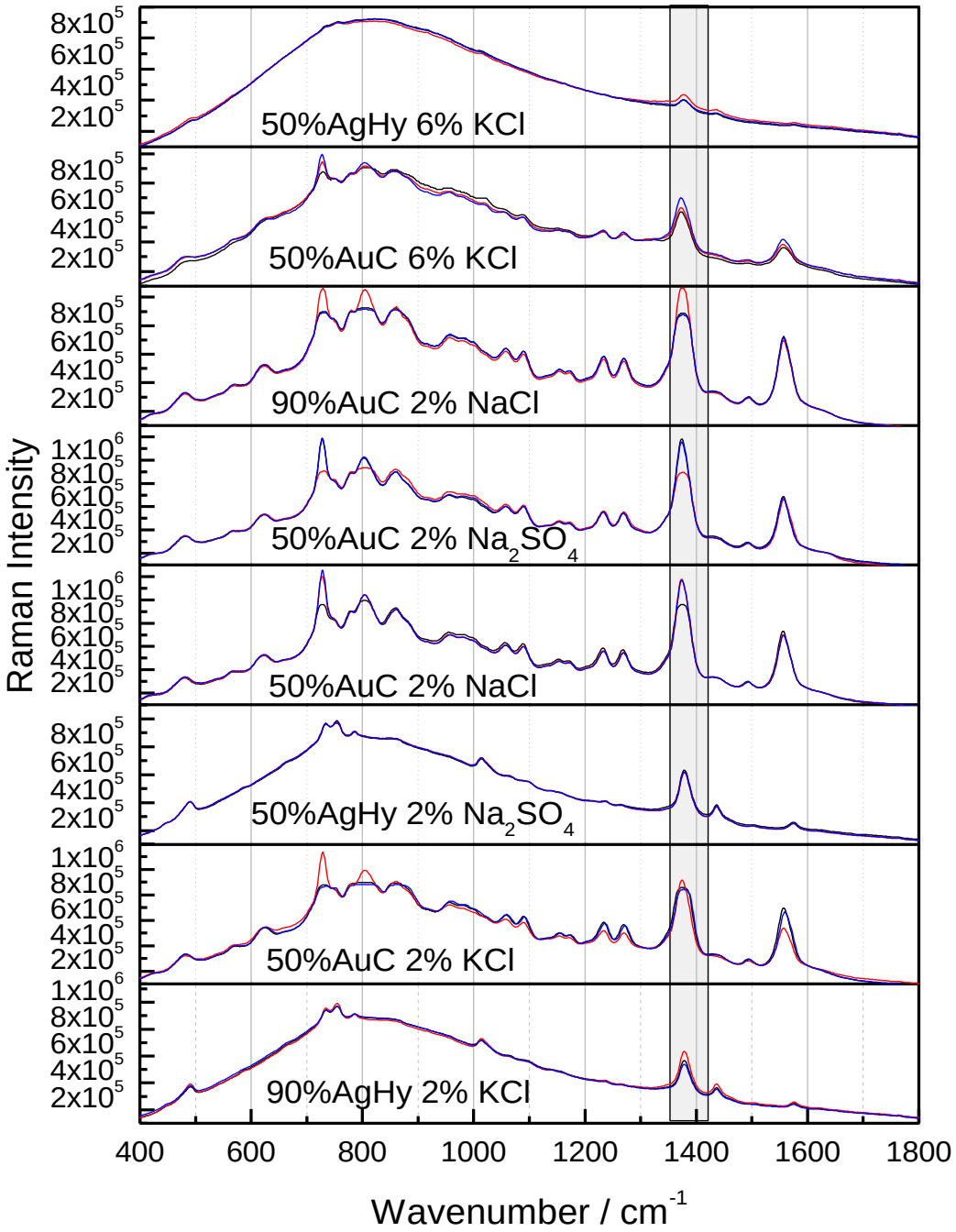
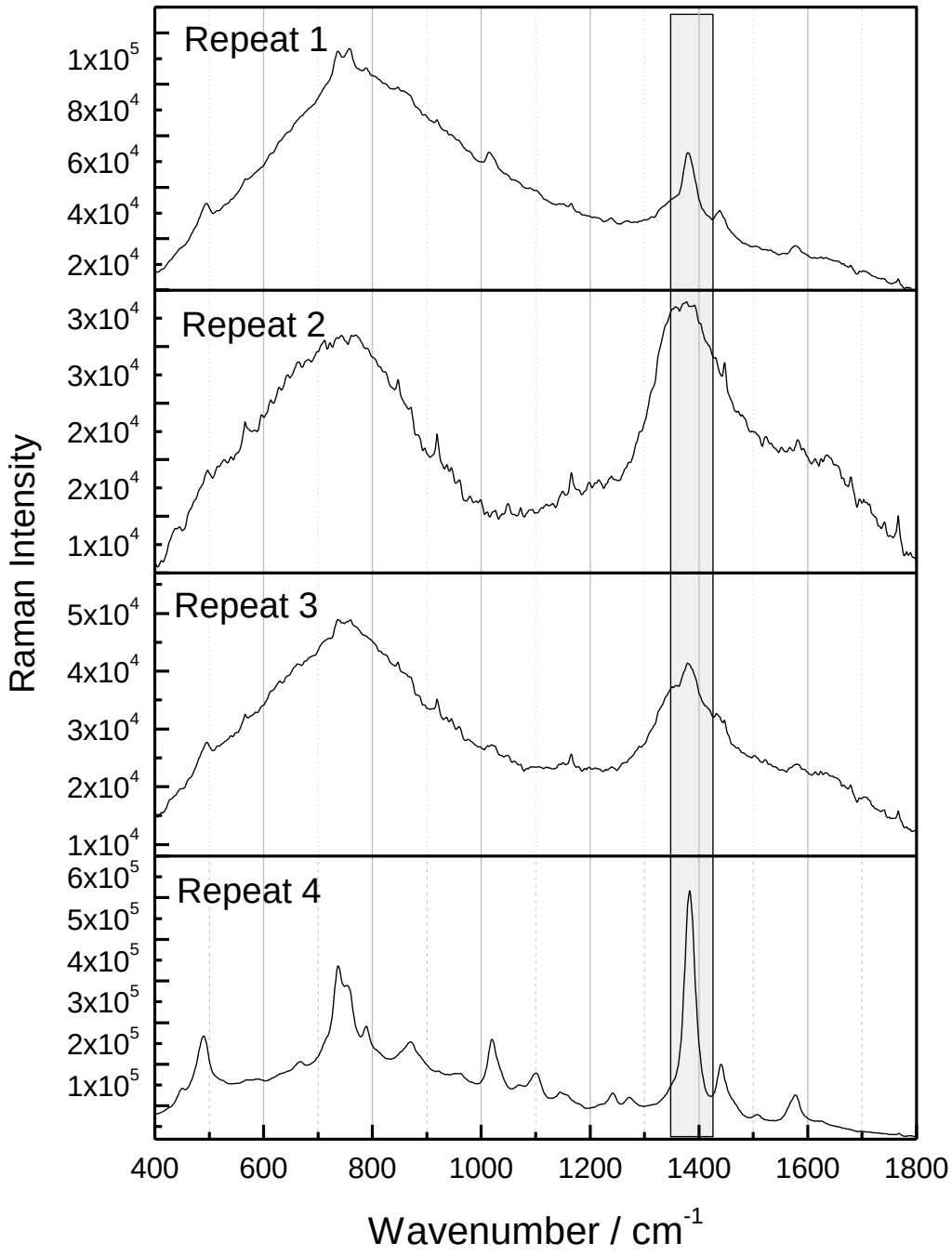
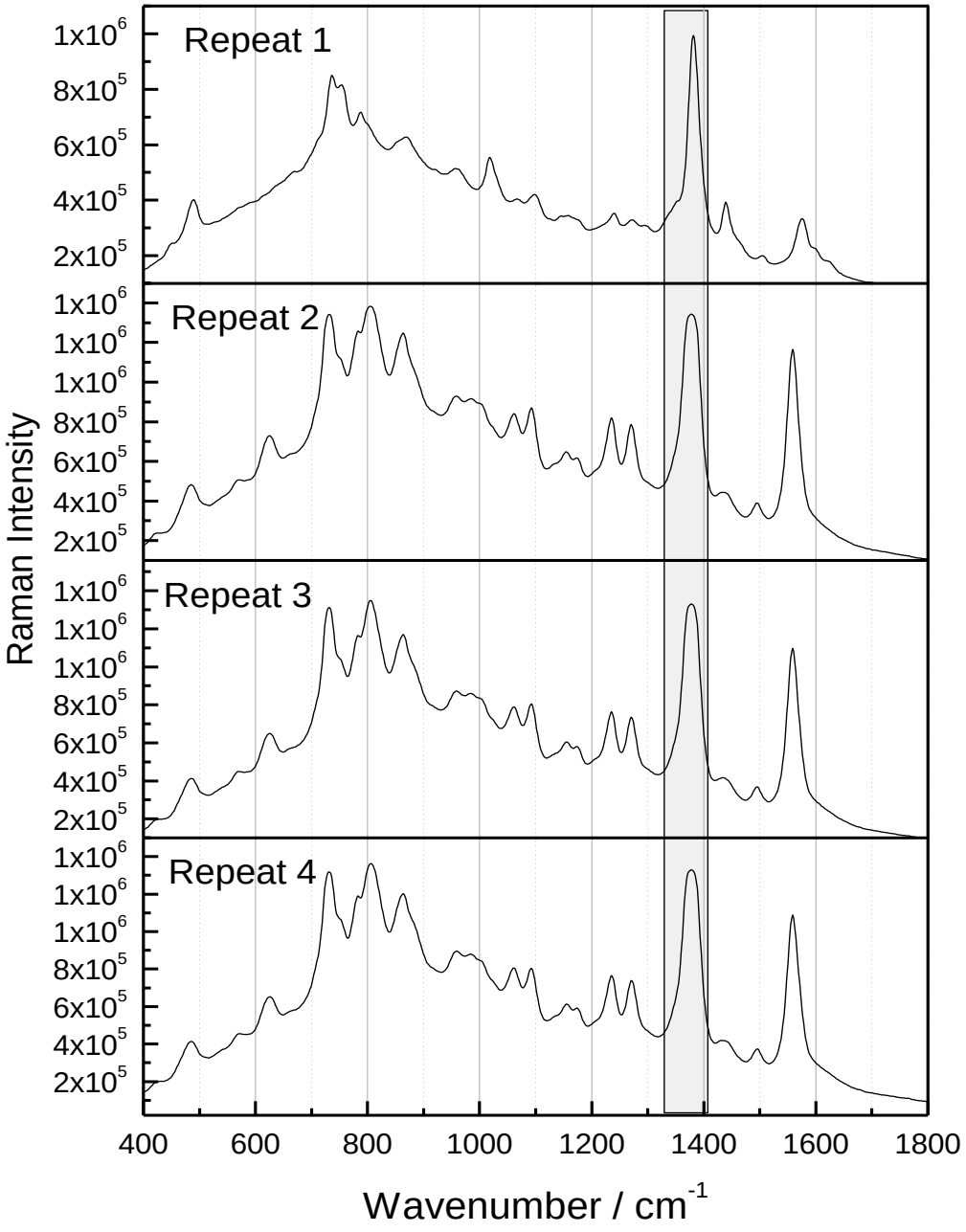


Fig. S2.4: Observed SERS spectra for the second generation of experiments (experimental conditions are listed in Table S2.3) which were performed in triplicate. An EMSC was applied to all spectra prior to analysis.


Repeat 1
Repeat 2
Repeat 3
Repeat 4
Raman Intensity
Wavenumber / cm-1
400
600
800
1000
1200
1400
1600
1800

(I)



(II)

Fig. S2.5: Representative SERS spectra demonstrating the relative differences in reproducibility in: (I) in the absence of an aggregating agent for hydroxylamine reduction of $AgNO_3$, and (II) in the absence of an aggregating agent for citrate reduction of $HAuCl_4$. These spectra illustrate that more consistent results were obtained for citrate reduction of $HAuCl_4$.

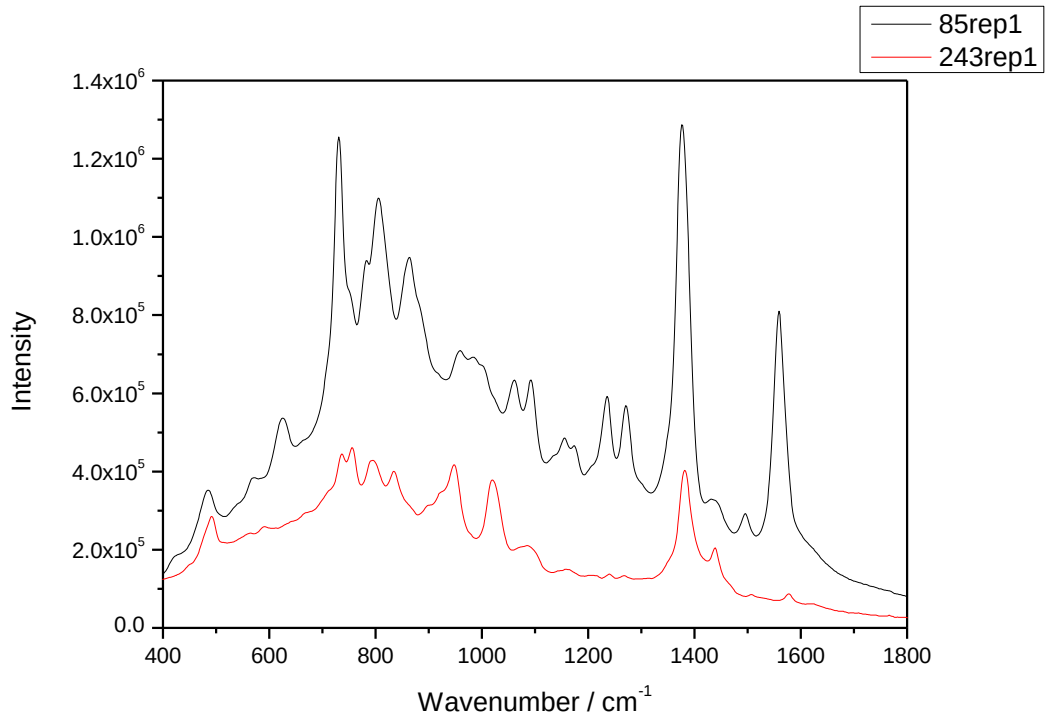


Fig. S2.6 Observed spectra of a solution on the Pareto front and a solution close to the Pareto front. Experiment 85 which is a Pareto front solution demonstrates a 26% improvement of intensity over experiment 243 which is a dominated solution. This confirms that the MOEA is working correctly and finding optimal solutions to our problem which can be important when working towards a LOD.

Table S2.1: Experimental conditions for the fractional factorial study

Exp.	% Metal	Metal	% Salt	Salt	% Analyte	H_2O	Laser (nm)
1	20	AgBh	1	NaCl	5	74	532
2	20	AgBh	7	NaCl	5	68	532
3	70	AgBh	1	KNO_3	5	24	532
4	70	AgBh	7	KNO_3	5	18	532
5	20	AgBh	1	NaCl	5	74	785
6	20	AgBh	7	NaCl	5	68	785
7	70	AgBh	1	KCl	5	24	785
8	70	AgBh	7	KCl	5	18	785
9	20	AgBh	1	Na_2SO_4	5	74	532
10	20	AgBh	7	Na_2SO_4	5	68	532
11	70	AgC	1	K_2SO_4	5	24	532
12	70	AgC	7	K_2SO_4	5	18	532
13	20	AgC	1	KNO_3	5	74	785
14	20	AgC	7	KNO_3	5	68	785
15	70	AgC	1	$MgSO_4$	5	24	785
16	70	AgC	7	$MgSO_4$	5	18	785
17	20	AgC	1	KCl	5	74	532
18	20	AgC	7	KCl	5	68	532
19	70	AgC	1	Na_2SO_4	5	24	532
20	70	AgC	7	Na_2SO_4	5	18	532

Table S2.1: Cont.

Exp.	% Metal	Metal	% Salt	Salt	% Analyte	H_2O	Laser (nm)
21	20	AgHy	1	K_2SO_4	5	74	785
22	20	AgHy	7	K_2SO_4	5	68	785
23	70	AgHy	1	NaCl	5	24	785
24	70	AgHy	7	NaCl	5	18	785
25	20	AgHy	1	$MgSO_4$	5	74	532
26	20	AgHy	7	$MgSO_4$	5	68	532
27	70	AgHy	1	KCl	5	24	532
28	70	AgHy	7	KCl	5	18	532
29	20	AgHy	1	Na_2SO_4	5	74	785
30	20	AgHy	7	Na_2SO_4	5	68	785
31	70	AuBh	1	K_2SO_4	5	24	785
32	70	AuBh	7	K_2SO_4	5	18	785
33	20	AuBh	1	NaCl	5	74	532
34	20	AuBh	7	NaCl	5	68	532
35	70	AuBh	1	KNO_3	5	24	532
36	70	AuBh	7	KNO_3	5	18	532
37	20	AuBh	1	KCl	5	74	785
38	20	AuBh	7	KCl	5	68	785
39	70	AuBh	1	Na_2SO_4	5	24	785
40	70	AuBh	7	Na_2SO_4	5	18	785
41	20	AuC	1	K_2SO_4	5	74	532
42	20	AuC	7	K_2SO_4	5	68	532
43	70	AuC	1	NaCl	5	24	532
44	70	AuC	7	NaCl	5	18	532
45	20	AuC	1	KNO_3	5	74	785
46	20	AuC	7	KNO_3	5	68	785
47	70	AuC	1	$MgSO_4$	5	24	785
48	70	AuC	7	$MgSO_4$	5	18	785
49	20	AuC	1	Na_2SO_4	5	74	532
50	20	AuC	7	Na_2SO_4	5	68	532
51	10	AuBh	0	0	5	85	532
52	50	AuBh	0	0	5	45	532
53	90	AuBh	0	0	5	5	532
54	10	AuC	0	0	5	85	532
55	50	AuC	0	0	5	45	532
56	90	AuC	0	0	5	5	532
57	10	AgBh	0	0	5	85	532
58	50	AgBh	0	0	5	45	532
59	90	AgBh	0	0	5	5	532
60	10	AgC	0	0	5	85	532
61	50	AgC	0	0	5	45	532
62	90	AgC	0	0	5	5	532
63	10	AgHy	0	0	5	85	532
64	50	AgHy	0	0	5	45	532
65	90	AgHy	0	0	5	5	532
66	10	AuBh	0	0	5	85	633
67	50	AuBh	0	0	5	45	633
68	90	AuBh	0	0	5	5	633
69	10	AuC	0	0	5	85	633
70	50	AuC	0	0	5	45	633
71	90	AuC	0	0	5	5	633
72	10	AgBh	0	0	5	85	633
73	50	AgBh	0	0	5	45	633
74	90	AgBh	0	0	5	5	633
75	10	AgC	0	0	5	85	633
76	50	AgC	0	0	5	45	633

Table S2.1: Cont.

Exp.	% Metal	Metal	% Salt	Salt	% Analyte	H_2O	Laser (nm)
77	90	AgC	0	0	5	5	633
78	10	AgHy	0	0	5	85	633
79	50	AgHy	0	0	5	45	633
80	90	AgHy	0	0	5	5	633
81	10	AuBh	0	0	5	85	785
82	50	AuBh	0	0	5	45	785
83	90	AuBh	0	0	5	5	785
84	10	AuC	0	0	5	85	785
85	50	AuC	0	0	5	45	785
86	90	AuC	0	0	5	5	785
87	10	AgBh	0	0	5	85	785
88	50	AgBh	0	0	5	45	785
89	90	AgBh	0	0	5	5	785
90	10	AgC	0	0	5	85	785
91	50	AgC	0	0	5	45	785
92	90	AgC	0	0	5	5	785
93	10	AgHy	0	0	5	85	785
94	50	AgHy	0	0	5	45	785
95	90	AgHy	0	0	5	5	785
96	10	AuBh	10	K_2SO_4	5	75	532
97	10	AuBh	10	KNO_3	5	75	532
98	10	AuBh	10	$MgSO_4$	5	75	532
99	10	AuBh	10	Na_2SO_4	5	75	532
100	10	AuC	10	NaCl	5	75	532
101	10	AuC	10	$MgSO_4$	5	75	532
102	10	AuC	10	KCl	5	75	532
103	10	AgBh	10	K_2SO_4	5	75	532
104	10	AgBh	10	KNO_3	5	75	532
105	10	AgBh	10	KCl	5	75	532
106	10	AgC	10	K_2SO_4	5	75	532
107	10	AgC	10	NaCl	5	75	532
108	10	AgC	10	$MgSO_4$	5	75	532
109	10	AgC	10	Na_2SO_4	5	75	532
110	10	AgHy	10	NaCl	5	75	532
111	10	AgHy	10	KNO_3	5	75	532
112	10	AgHy	10	KCl	5	75	532
113	10	AuBh	10	K_2SO_4	5	75	633
114	10	AuBh	10	NaCl	5	75	633
115	10	AuBh	10	KNO_3	5	75	633
116	10	AuBh	10	$MgSO_4$	5	75	633
117	10	AuBh	10	KCl	5	75	633
118	10	AuBh	10	Na_2SO_4	5	75	633
119	10	AuC	10	K_2SO_4	5	75	633
120	10	AuC	10	NaCl	5	75	633
121	10	AuC	10	KNO_3	5	75	633
122	10	AuC	10	$MgSO_4$	5	75	633
123	10	AuC	10	KCl	5	75	633
124	10	AuC	10	Na_2SO_4	5	75	633
125	10	AgBh	10	K_2SO_4	5	75	633
126	10	AgBh	10	NaCl	5	75	633
127	10	AgBh	10	KNO_3	5	75	633
128	10	AgBh	10	$MgSO_4$	5	75	633
129	10	AgBh	10	KCl	5	75	633
130	10	AgBh	10	Na_2SO_4	5	75	633
131	10	AgC	10	K_2SO_4	5	75	633
132	10	AgC	10	NaCl	5	75	633

Table S2.1: Cont.

Exp.	% Metal	Metal	% Salt	Salt	% Analyte	H_2O	Laser (nm)
133	10	AgC	10	KNO_3	5	75	633
134	10	AgC	10	$MgSO_4$	5	75	633
135	10	AgC	10	KCl	5	75	633
136	10	AgC	10	Na_2SO_4	5	75	633
137	10	AgHy	10	K_2SO_4	5	75	633
138	10	AgHy	10	NaCl	5	75	633
139	10	AgHy	10	KNO_3	5	75	633
140	10	AgHy	10	$MgSO_4$	5	75	633
141	10	AgHy	10	KCl	5	75	633
142	10	AgHy	10	Na_2SO_4	5	75	633
143	10	AuBh	10	K_2SO_4	5	75	785
144	10	AuBh	10	KNO_3	5	75	785
145	10	AuBh	10	$MgSO_4$	5	75	785
146	10	AuBh	10	Na_2SO_4	5	75	785
147	10	AuC	10	NaCl	5	75	785
148	10	AuC	10	$MgSO_4$	5	75	785
149	10	AuC	10	KCl	5	75	785
150	10	AgBh	10	K_2SO_4	5	75	785
151	10	AgBh	10	KNO_3	5	75	785
152	10	AgBh	10	KCl	5	75	785
153	10	AgC	10	K_2SO_4	5	75	785
154	10	AgC	10	NaCl	5	75	785
155	10	AgC	10	$MgSO_4$	5	75	785
156	10	AgC	10	Na_2SO_4	5	75	785
157	10	AgHy	10	NaCl	5	75	785
158	10	AgHy	10	KNO_3	5	75	785
159	10	AgHy	10	KCl	5	75	785
160	50	AuBh	6	K_2SO_4	5	39	532
161	50	AuBh	6	NaCl	5	39	532
162	50	AuBh	6	KNO_3	5	39	532
163	50	AuBh	6	$MgSO_4$	5	39	532
164	50	AuBh	6	KCl	5	39	532
165	50	AuBh	6	Na_2SO_4	5	39	532
166	50	AuC	6	K_2SO_4	5	39	532
167	50	AuC	6	NaCl	5	39	532
168	50	AuC	6	KNO_3	5	39	532
169	50	AuC	6	$MgSO_4$	5	39	532
170	50	AuC	6	KCl	5	39	532
171	50	AuC	6	Na_2SO_4	5	39	532
172	50	AgBh	6	K_2SO_4	5	39	532
173	50	AgBh	6	NaCl	5	39	532
174	50	AgBh	6	KNO_3	5	39	532
175	50	AgBh	6	$MgSO_4$	5	39	532
176	50	AgBh	6	KCl	5	39	532
177	50	AgBh	6	Na_2SO_4	5	39	532
178	50	AgC	6	K_2SO_4	5	39	532
179	50	AgC	6	NaCl	5	39	532
180	50	AgC	6	KNO_3	5	39	532
181	50	AgC	6	$MgSO_4$	5	39	532
182	50	AgC	6	KCl	5	39	532
183	50	AgC	6	Na_2SO_4	5	39	532
184	50	AgHy	6	K_2SO_4	5	39	532
185	50	AgHy	6	NaCl	5	39	532
186	50	AgHy	6	KNO_3	5	39	532
187	50	AgHy	6	$MgSO_4$	5	39	532
188	50	AgHy	6	KCl	5	39	532

Table S2.1: Cont.

Exp.	% Metal	Metal	% Salt	Salt	% Analyte	H_2O	Laser (nm)
189	50	AgHy	6	Na_2SO_4	5	39	532
190	50	AuBh	6	K_2SO_4	5	39	633
191	50	AuBh	6	NaCl	5	39	633
192	50	AuBh	6	KNO_3	5	39	633
193	50	AuBh	6	$MgSO_4$	5	39	633
194	50	AuBh	6	KCl	5	39	633
195	50	AuBh	6	Na_2SO_4	5	39	633
196	50	AuC	6	K_2SO_4	5	39	633
197	50	AuC	6	NaCl	5	39	633
198	50	AuC	6	KNO_3	5	39	633
199	50	AuC	6	$MgSO_4$	5	39	633
200	50	AuC	6	KCl	5	39	633
201	50	AuC	6	Na_2SO_4	5	39	633
202	50	AgBh	6	K_2SO_4	5	39	633
203	50	AgBh	6	NaCl	5	39	633
204	50	AgBh	6	KNO_3	5	39	633
205	50	AgBh	6	$MgSO_4$	5	39	633
206	50	AgBh	6	KCl	5	39	633
207	50	AgBh	6	Na_2SO_4	5	39	633
208	50	AgC	6	K_2SO_4	5	39	633
209	50	AgC	6	NaCl	5	39	633
210	50	AgC	6	KNO_3	5	39	633
211	50	AgC	6	$MgSO_4$	5	39	633
212	50	AgC	6	KCl	5	39	633
213	50	AgC	6	Na_2SO_4	5	39	633
214	50	AgHy	6	K_2SO_4	5	39	633
215	50	AgHy	6	NaCl	5	39	633
216	50	AgHy	6	KNO_3	5	39	633
217	50	AgHy	6	$MgSO_4$	5	39	633
218	50	AgHy	6	KCl	5	39	633
219	50	AgHy	6	Na_2SO_4	5	39	633
220	50	AuBh	6	K_2SO_4	5	39	785
221	50	AuBh	6	NaCl	5	39	785
222	50	AuBh	6	KNO_3	5	39	785
223	50	AuBh	6	$MgSO_4$	5	39	785
224	50	AuBh	6	KCl	5	39	785
225	50	AuBh	6	Na_2SO_4	5	39	785
226	50	AuC	6	K_2SO_4	5	39	785
227	50	AuC	6	NaCl	5	39	785
228	50	AuC	6	KNO_3	5	39	785
229	50	AuC	6	$MgSO_4$	5	39	785
230	50	AuC	6	KCl	5	39	785
231	50	AuC	6	Na_2SO_4	5	39	785
232	50	AgBh	6	K_2SO_4	5	39	785
233	50	AgBh	6	NaCl	5	39	785
234	50	AgBh	6	KNO_3	5	39	785
235	50	AgBh	6	$MgSO_4$	5	39	785
236	50	AgBh	6	KCl	5	39	785
237	50	AgBh	6	Na_2SO_4	5	39	785
238	50	AgC	6	K_2SO_4	5	39	785
239	50	AgC	6	NaCl	5	39	785
240	50	AgC	6	KNO_3	5	39	785
241	50	AgC	6	$MgSO_4$	5	39	785
242	50	AgC	6	KCl	5	39	785
243	50	AgC	6	Na_2SO_4	5	39	785
244	50	AgHy	6	K_2SO_4	5	39	785

Table S2.1: Cont.

Exp.	% Metal	Metal	% Salt	Salt	% Analyte	H_2O	Laser (nm)
245	50	AgHy	6	NaCl	5	39	785
246	50	AgHy	6	KNO_3	5	39	785
247	50	AgHy	6	$MgSO_4$	5	39	785
248	50	AgHy	6	KCl	5	39	785
249	50	AgHy	6	Na_2SO_4	5	39	785
250	90	AuBh	2	NaCl	5	3	532
251	90	AuBh	2	$MgSO_4$	5	3	532
252	90	AuBh	2	KCl	5	3	532
253	90	AuC	2	K_2SO_4	5	3	532
254	90	AuC	2	KNO_3	5	3	532
255	90	AuC	2	KCl	5	3	532
256	90	AuC	2	Na_2SO_4	5	3	532
257	90	AgBh	2	K_2SO_4	5	3	532
258	90	AgBh	2	NaCl	5	3	532
259	90	AgBh	2	$MgSO_4$	5	3	532
260	90	AgBh	2	Na_2SO_4	5	3	532
261	90	AgC	2	NaCl	5	3	532
262	90	AgC	2	KNO_3	5	3	532
263	90	AgC	2	KCl	5	3	532
264	90	AgHy	2	K_2SO_4	5	3	532
265	90	AgHy	2	KNO_3	5	3	532
266	90	AgHy	2	$MgSO_4$	5	3	532
267	90	AgHy	2	Na_2SO_4	5	3	532
268	90	AuBh	2	K_2SO_4	5	3	633
269	90	AuBh	2	NaCl	5	3	633
270	90	AuBh	2	KNO_3	5	3	633
271	90	AuBh	2	$MgSO_4$	5	3	633
272	90	AuBh	2	KCl	5	3	633
273	90	AuBh	2	Na_2SO_4	5	3	633
274	90	AuC	2	K_2SO_4	5	3	633
275	90	AuC	2	NaCl	5	3	633
276	90	AuC	2	KNO_3	5	3	633
277	90	AuC	2	$MgSO_4$	5	3	633
278	90	AuC	2	KCl	5	3	633
279	90	AuC	2	Na_2SO_4	5	3	633
280	90	AgBh	2	K_2SO_4	5	3	633
281	90	AgBh	2	NaCl	5	3	633
282	90	AgBh	2	KNO_3	5	3	633
283	90	AgBh	2	$MgSO_4$	5	3	633
284	90	AgBh	2	KCl	5	3	633
285	90	AgBh	2	Na_2SO_4	5	3	633
286	90	AgC	2	K_2SO_4	5	3	633
287	90	AgC	2	NaCl	5	3	633
288	90	AgC	2	KNO_3	5	3	633
289	90	AgC	2	$MgSO_4$	5	3	633
290	90	AgC	2	KCl	5	3	633
291	90	AgC	2	Na_2SO_4	5	3	633
292	90	AgHy	2	K_2SO_4	5	3	633
293	90	AgHy	2	NaCl	5	3	633
294	90	AgHy	2	KNO_3	5	3	633
295	90	AgHy	2	$MgSO_4$	5	3	633
296	90	AgHy	2	KCl	5	3	633
297	90	AgHy	2	Na_2SO_4	5	3	633
298	90	AuBh	2	NaCl	5	3	785
299	90	AuBh	2	$MgSO_4$	5	3	785
300	90	AuBh	2	KCl	5	3	785

Table S2.1: Cont.

Exp.	% Metal	Metal	% Salt	Salt	% Analyte	H_2O	Laser (nm)
301	90	AuC	2	K_2SO_4	5	3	785
302	90	AuC	2	KNO_3	5	3	785
303	90	AuC	2	KCl	5	3	785
304	90	AuC	2	Na_2SO_4	5	3	785
305	90	AgBh	2	K_2SO_4	5	3	785
306	90	AgBh	2	NaCl	5	3	785
307	90	AgBh	2	$MgSO_4$	5	3	785
308	90	AgBh	2	Na_2SO_4	5	3	785
309	90	AgC	2	NaCl	5	3	785
310	90	AgC	2	KNO_3	5	3	785
311	90	AgC	2	KCl	5	3	785
312	90	AgHy	2	K_2SO_4	5	3	785
313	90	AgHy	2	KNO_3	5	3	785
314	90	AgHy	2	$MgSO_4$	5	3	785
315	90	AgHy	2	Na_2SO_4	5	3	785

AgHy, Hydroxylamine reduction of $AgNO_3$; AgC, Citrate reduction of $AgNO_3$; AgBh,Borohydride reduction of $AgNO_3$; AuC, Citrate reduction of $HAuCl_4$; AuBh, Borohydride reduction of $HAuCl_4$

Table S2.2: Experimental conditions, RHMFW and Reproducibility (see page 55 for calculations) values for the top 35 experiments for each laser excitation wavelength.

Exp	%Metal	Metal	% Salt	Salt	Laser(nm)	% H2O	RHMFW	Rep.
1	20	AgBh	1	NaCl	532	74	0.0007818	0.9860563
2	20	AgBh	7	NaCl	532	68	0.0003823	0.9694903
3	70	AgBh	1	KNO3	532	24	0.0008868	0.9935829
4	70	AgBh	7	KNO3	532	18	0.0011649	0.9110733
9	20	AgBh	1	Na2SO4	532	74	0.0013439	0.9578850
10	20	AgBh	7	Na2SO4	532	68	0.000248	0.9011291
11	70	AgC	1	K2SO4	532	24	0.0008765	0.8205746
12	70	AgC	7	K2SO4	532	18	0.0019134	0.9885939
17	20	AgC	1	KCl	532	74	0.0012186	0.4584051
18	20	AgC	7	KCl	532	68	0.0016668	0.9491515
19	70	AgC	1	Na2SO4	532	24	0.0002739	0.9731962
25	20	AgHy	1	$MgSO_4$	532	74	0.0072671	0.9806879
26	20	AgHy	7	$MgSO_4$	532	68	0.0033792	0.9919365
27	70	AgHy	1	KCl	532	24	0.0046555	0.9738544
28	70	AgHy	7	KCl	532	18	0.0029686	0.9582929
58	50	AgBh	0	None	532	45	0.0006888	0.9737556
59	90	AgBh	0	None	532	5	0.0007729	0.7314980
60	10	AgC	0	None	532	85	0.0003099	0.7663300
61	50	AgC	0	None	532	45	0.0006293	0.6599318
62	90	AgC	0	None	532	5	0.0010444	0.7256348
63	10	AgHy	0	None	532	85	0.0005966	0.9529210
64	50	AgHy	0	None	532	45	0.0004972	0.9042508
65	90	AgHy	0	None	532	5	0.0079598	0.9880928
20	70	AgC	7	Na2SO4	532	18	0.0010242	0.9927537
103	10	AgBh	10	K2SO4	532	75	0.0002664	0.9362621
104	10	AgBh	10	KNO3	532	75	0.0005739	0.9977370
105	10	AgBh	10	KCl	532	75	9.04E-17	0.7136317
106	10	AgC	10	K2SO4	532	75	0.0026595	0.9853837
107	10	AgC	10	NaCl	532	75	0.0029707	0.9981194
108	10	AgC	10	$MgSO_4$	532	75	0.0021527	0.9955688
109	10	AgC	10	Na2SO4	532	75	0.0063402	0.8938766
111	10	AgHy	10	KNO3	532	75	0.0017837	0.8014731
172	50	AgBh	6	K2SO4	532	39	2.13E-17	0.9902685
174	50	AgBh	6	KNO3	532	39	0.0014734	0.9078552
179	50	AgC	6	NaCl	532	39	0.0081306	0.9915709
67	50	AuBh	0	None	633	45	0.00124	0.9973015
68	90	AuBh	0	None	633	5	0.001716	0.9979812
70	50	AuC	0	None	633	45	0.0017893	0.9985373
71	90	AuC	0	None	633	5	0.0034376	0.9987074
73	50	AgBh	0	None	633	45	0	0.9997801
74	90	AgBh	0	None	633	5	5.65E-05	0.9984665
75	10	AgC	0	None	633	85	0.0003076	0.9963967
76	50	AgC	0	None	633	45	0.0004458	0.9840583
77	90	AgC	0	None	633	5	0.0003379	0.9907843
78	10	AgHy	0	none	633	85	0.0002767	0.9793328
79	50	AgHy	0	none	633	45	0.0009037	0.9950422
80	90	AgHy	0	none	633	5	0.0024011	0.9989047
131	10	AgC	10	K2SO4	633	75	0.0013233	0.9959887
132	10	AgC	10	NaCl	633	75	0.0006301	0.9947151
133	10	AgC	10	KNO3	633	75	0.0019738	0.9982885
134	10	AgC	10	$MgSO_4$	633	75	0.0017023	0.9973875
136	10	AgC	10	Na2SO4	633	75	0.0025131	0.9954410
191	50	AuBh	6	NaCl	633	39	0.0006537	0.9992826

Table S2.2 Cont.

Exp	%Metal	Metal	% Salt	Salt	Laser(nm)	% H2O	FWHM	Rep
209	50	AgC	6	NaCl	633	39	0.0016528	0.9661757
210	50	AgC	6	KNO3	633	39	0.002025	0.9985192
211	50	AgC	6	$MgSO_4$	633	39	0.0027389	0.9994082
212	50	AgC	6	KCl	633	39	0.0013133	0.9821819
213	50	AgC	6	Na2SO4	633	39	0.0027377	0.9967578
268	90	AuBh	2	K2SO4	633	3	0.0015259	0.9990636
269	90	AuBh	2	NaCl	633	3	0.001998	0.9992729
270	90	AuBh	2	KNO3	633	3	0.0017399	0.9989720
272	90	AuBh	2	KCl	633	3	0.0028638	0.9982551
273	90	AuBh	2	Na2SO4	633	3	0.0018061	0.9972161
275	90	AuC	2	NaCl	633	3	0.0015351	0.9988953
286	90	AgC	2	K2SO4	633	3	0.0020746	0.9989955
287	90	AgC	2	NaCl	633	3	0.000854	0.9620250
289	90	AgC	2	$MgSO_4$	633	3	0.0027859	0.9976735
290	90	AgC	2	KCl	633	3	0.0021716	0.9283899
291	90	AgC	2	Na2SO4	633	3	0.0039751	0.9973721
15	70	AgC	1	$MgSO_4$	785	24	0.0040283	0.9987636
23	70	AgHy	1	NaCl	785	24	0.0027722	0.9981028
47	70	AuC	1	$MgSO_4$	785	24	0.0054693	0.9903152
82	50	AuBh	0	None	785	45	0.0030403	0.9978472
83	90	AuBh	0	None	785	5	0.0039033	0.9989428
84	10	AuC	0	None	785	85	0.011085	0.9975688
85	50	AuC	0	None	785	45	0.0124174	0.9938986
86	90	AuC	0	None	785	5	0.005997	0.9991184
90	10	AgC	0	None	785	85	0.0021603	0.9988131
91	50	AgC	0	None	785	45	0.0006998	0.9985796
92	90	AgC	0	None	785	5	0.0002102	0.9997584
94	50	AgHy	0	None	785	45	0.0014439	0.7818621
95	90	AgHy	0	None	785	5	0.004342	0.9976340
155	10	AgC	10	$MgSO_4$	785	75	0.0022852	0.9966401
156	10	AgC	10	Na2SO4	785	75	0.0029586	0.9975039
227	50	AuC	6	NaCl	785	39	0.0100046	0.9996431
229	50	AuC	6	$MgSO_4$	785	39	0.006205	0.9987377
231	50	AuC	6	Na2SO4	785	39	0.0082573	0.9958329
241	50	AgC	6	$MgSO_4$	785	39	0.0023579	0.9963398
243	50	AgC	6	Na2SO4	785	39	0.0092114	0.9968489
245	50	AgHy	6	NaCl	785	39	0.0056831	0.9996215
249	50	AgHy	6	Na2SO4	785	39	0.0019541	0.9998693
298	90	AuBh	2	NaCl	785	3	0.0046969	0.9988741
299	90	AuBh	2	$MgSO_4$	785	3	0.0034221	0.9977913
300	90	AuBh	2	KCl	785	3	0.0043057	0.9965541
301	90	AuC	2	K2SO4	785	3	0.0038606	0.8921341
302	90	AuC	2	KNO3	785	3	0.0051249	0.9948551
303	90	AuC	2	KCl	785	3	0.0104188	0.9976928
304	90	AuC	2	Na2SO4	785	3	0.0068583	0.9896388
309	90	AgC	2	NaCl	785	3	0.0060909	0.9995025
311	90	AgC	2	KCl	785	3	0.0053014	0.9985426
312	90	AgHy	2	K2SO4	785	3	0.0058521	0.9988292
313	90	AgHy	2	KNO3	785	3	0.0045641	0.9942991
314	90	AgHy	2	$MgSO_4$	785	3	0.0005421	0.9998828
315	90	AgHy	2	Na2SO4	785	3	0.009227	0.9998469

AgHy, Hydroxylamine reduction of $AgNO_3$; AgC, Citrate reduction of $AgNO_3$; AgBh, Borohydride reduction of $AgNO_3$; AuC, Citrate reduction of $HAuCl_4$; AuBh, Borohydride reduction of $HAuCl_4$

Table S2.3: Second generation experimental conditions. The analyte concentration was fixed at 50µL 8.79 x 10^{-4}.

% Metal	Metal	% Salt	Salt	Laser (nm)	% H2O
90	AgHy	2	KCl	785	3
50	AuC	2	KCl	785	43
50	AgHy	2	Na2SO4	785	43
50	AuC	2	NaCl	785	43
50	AuC	2	Na2SO4	785	43
90	AuC	2	NaCl	785	3
50	AuC	6	KCl	785	39
50	AgHy	6	KCl	785	39

AgHy, Hydroxylamine reduction of $AgNO_3$; AuC, Citrate reduction of $HAuCl_4$

Table S2.4: Experimental conditions for the solutions observed on the Pareto front.

Exp. No.	% Metal	Metal type	% Salt	Salt	Laser (nm)	% H_2O	FWHM	Rep
249	50	Aghy	6	Na_2SO_4	785	39	0.001514	0.999869
303	90	AuC	2	KCl	785	3	0.011416	0.997693
304	90	AuC	2	Na_2SO_4	785	3	0.012608	0.989639
227	50	AuC	6	NaCl	785	39	0.010005	0.999643
303	90	AuC	2	KCl	785	3	0.010419	0.997693
314	90	AgHy	2	$MgSO_4$	785	3	0.000627	0.999882

AgHy, Hydroxylamine reduction of $AgNO_3$; AuC, Citrate reduction of $HAuCl_4$

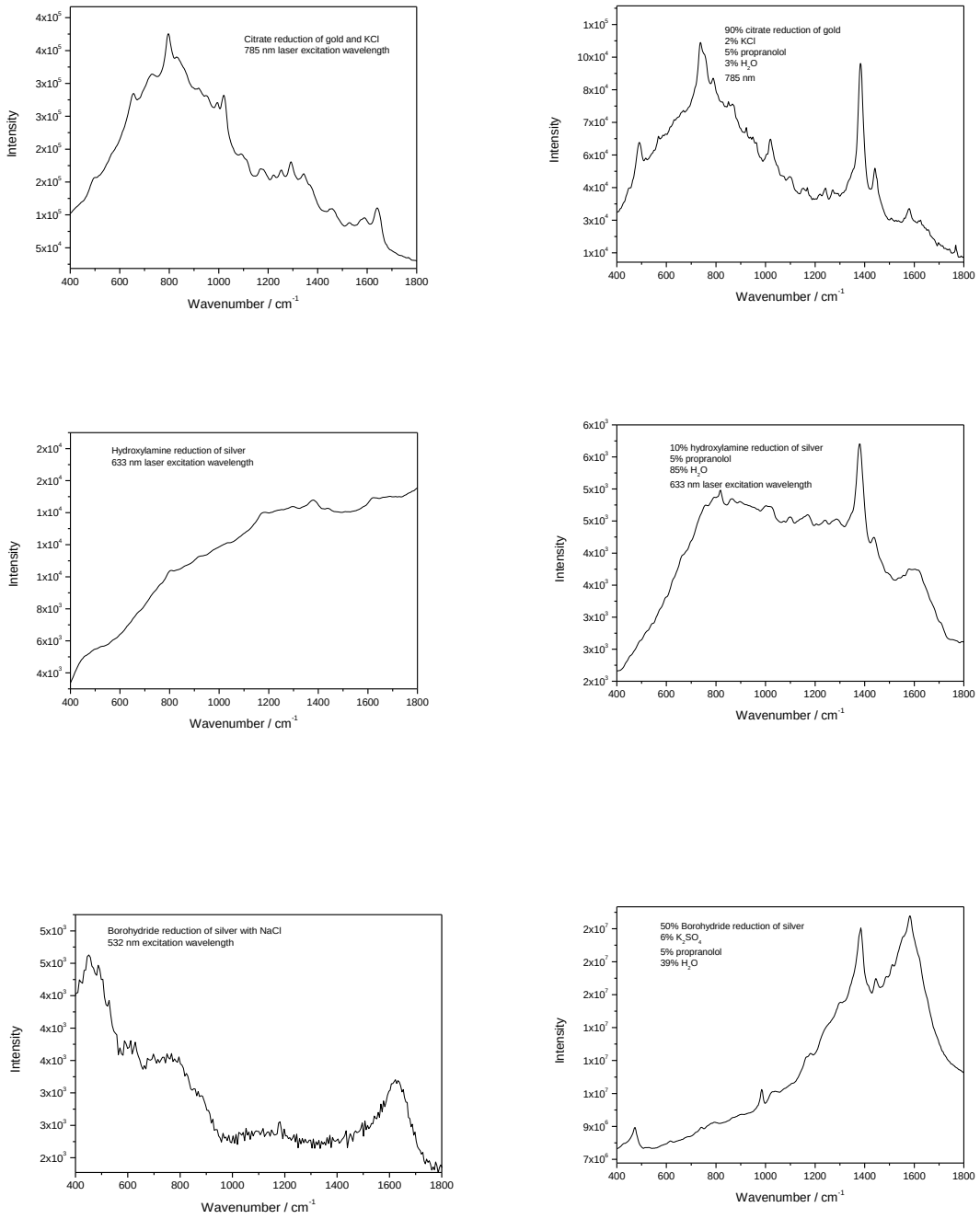


Fig S2.7 An example blank spectrum for each laser excitation wavelength (left column) and an example propranolol and colloid spectrum for each laser excitation wavelength (right column). The samples observed for 785 nm display a large background between 600 – 1000 cm^{-1} which is due to the glass band, the background observed or the 633 nm and 532 nm are due to instrument response.

Chapter 3. Surface-enhanced Raman scattering (SERS) sensing of propranolol in human body fluids.

This work will not be submitted as a paper as it was an extension that extends from the previous chapter into biofluids.

Personal contribution

I performed all sample preparation for, and measurement of SERS. I conducted analysis of the results in OriginPro 8. The discussion and conclusions are drawn from the results, a review of the literature and discussions with Dr Ewan Blanch, Dr Elon Correa and Professor Roy Goodacre.

Chapter 3.

Surface-enhanced Raman scattering (SERS) sensing of propranolol in human body fluids – a proof-of-principle study

Clare Levene [1], Elon Correa [2], Ewan Blanch [1] and Royston Goodacre[2]

[1]*Faculty of Life Sciences, University of Manchester, MIB, 131 Princes Street, Manchester, M1 7DN*

[2]*School of Chemistry, University of Manchester, MIB, 131 Princess Street, Manchester, M1 7DN*

3.1 Abstract

The rapid identification of pharmaceuticals in biofluids comes from the demand of patient need and commercial pressure. In this study we demonstrate the viability of surface-enhanced Raman scattering (SERS) for the detection of trace molecule analysis. Using experimental conditions that were previously determined with the use of a multiobjective evolutionary algorithm (MOEA) SERS measurements were taken for plasma spiked with propranolol which was detected to levels lower than 130 ng/mL. Further optimization of the experimental conditions for SERS of propranolol are required, however, this proof-of-principle study is a further development towards SERS detection of biologically relevant molecules in biofluids.

Keywords: surface-enhanced Raman scattering; biofluids; plasma

3.2 Introduction

The implementation of an analytical technique that can provide reliable and fast measurements of analytes in biofluids comes from the demands of patient need and commercial pressure. Vital medical information is often secreted through biofluids and is frequently used to help patient diagnosis and determine pharmaceutical dosage. The traditional techniques such as high-performance liquid chromatography (HPLC) or high-performance capillary electrophoresis (HPCE) are time consuming separation techniques that can potentially take days for results to be produced and as a result biomedical and pharmaceutical analysis can no longer afford to rely solely on these technologies (Ambrose *et al.* 1983; Johansson 2007). It is therefore imperative that an efficient, fast and reliable technique can be developed in order to meet the demands of biomedical analysis.

Raman spectroscopy is a non-invasive technique that measures molecular vibrational frequencies giving holistic fingerprints of the analyte under investigation (Goodacre, 2003; Smith & Dent 2005). With this distinct fingerprint, identification of an analyte molecule is possible even from complex samples such as biofluids which include other molecular species such as urea, CO_2, salts and proteins, amongst others. Previous literature has identified the potential of Raman spectroscopy in analyzing biofluids such as urine, serum and saliva (Premasiri *et al*. 2001; Berger *et al*. 1999; Farquharson *et al*. 2005) and cancer lesion tissue which is a relatively more complex sample than human biofluids (Bakker *et al*. 1997). However, Raman spectroscopy does have some limitations in that the Raman effect is relatively weak. Surface-enhanced Raman scattering (SERS) can overcome the inherent weakness of the

Raman effect as an increase in the Raman cross-section is observed in the presence of an electromagnetic field where molecules are in close proximity to surface plasmons generated by metal nanostructures, typically gold or silver (Otto 1984; Campion 1998; Moskovits 1985; Persson 1981). SERS was first observed in 1977 (Jeanmarie & VanDuyne 1977; Albrecht & Creighton 1977) and has been gaining popularity in physical and analytical chemistry and more recently in the biomedical field (Vo-Dinh *et al*. 2002; Yonzon *et al*. 2004; Qian & Nie 2008; Xie 2011). An increase in the Raman cross-section occurs if, when incident light is focused on the metal substrate the resulting surface plasmons (SP) oscillate at a frequency that matches the natural oscillating frequency of the free electrons on the metal surface. As a result, the SP resonance effect creates a large field-induced polarization at the metal particle surface which increases the intensity of the Raman signal (Rupérez & Laserna, 1996; Moskovits, 1985; Garcia-Vidal & Pendry 1996; Kniepp *et al*. 2002). The greatest enhancements occur between particles and are often referred to as 'hot spots', which are likely due to coupling between electrons of adjacent particles which results in new plasmon resonance generating extensive electric fields at the point of contact between the two particles (Xu & Käll 2000; Xu *et al.* 2003). SERS, however, does suffer from some limitations with the main one being reproducibility. The optimal experimental conditions of SERS can differ significantly for each analyte investigated and the SERS enhancement is dependent upon the degree of aggregation which can vary between aggregating agents and metal type used, this in turn can adversely affect the signal reproducibility between samples (Yea *et al*. 2005; Sanchez-Cortes *et al*. 1995). It is, therefore, important to ensure that the experimental conditions for the analyte under investigation are optimal or near optimal.

In our previous work we developed a multiobjective evolutionary algorithm (MOEA) based on a Pareto front towards the optimization of experimental conditions for SERS of propranolol. The purpose of the study was to find a fast and effective improvement in spectral reproducibility and enhancement, without the need for a full and exhaustive search of experimental conditions. MOEAs can speed up the process of problem solving where empirical means are impractical (Handl *et al*. 2007). The theory of MOEAs is that the best available values achieved for experimental conditions will maximize the objective functions. The objective functions to be satisfied during the study were to maximize reproducibility and enhancement of the SERS signal. It was important to weight the objective functions equally during the study to avoid bias of optimal conditions for one objective that may prejudice the experimental conditions for the other objective. During the study we worked towards optimizing the experimental conditions of SERS for propranolol in distilled H_2O (dH_2O). Without the development of the MOEA we would have had to perform an astronomical number of experiments to search for an optimal or near optimal solution. We demonstrated that you can achieve an acceptable solution for an optimization problem with a reduction of > 96% of required experiments. The results accomplished did demonstrate that there is a trade-off when observing superior spectral enhancement at the detriment of reduced spectral reproducibility and vice-versa. However, when analyzing the results presented on the Pareto front, solutions that fall in the middle of the Pareto front offer a compromise of reproducibility and enhancement and the experimental conditions chosen towards the detection of propranolol in biofluids were selected from the solutions that presented in the middle of the Pareto front. During our previous study we performed a practical limit of detection (LOD) for propranolol in dH_2O and achieved concentrations of 2.4

ng/mL which are well within typical physiological concentrations that have been previously determined using gas-liquid chromatographic assays and high performance liquid chromatography. (Pine et al. 1975; Majahan et al. 1984; Murillo Pulgarín *et al*. 1998; Mullet *et al*. 2001). This study was trained on H_2O solutions rather than biofluids and we now present results obtained using this approach on plasma using a solution selected from the middle of the Pareto front.

Propranolol is an amphipathic *β*-blocker used to treat patients with mild-to-moderate chronic heart failure, by preventing the binding of adrenaline to cell surface receptors which can increase the heart rate and constrict blood vessels (Bristow 2000) and is the only drug proven effective for the prophylaxis of migraines in children (Ludvigsson 1974). Propranolol has also been shown to be a potential treatment for post traumatic stress disorder if administered shortly after a traumatic event by blocking β-adrenergic receptors (Davidson *et al*. 2003). Propranolol is a complex pharmaceutical which has the capability to bind to membranes by three-point binding due to it hydrophobic aromatic group, the OH group and the positively charged NH group (see Fig. 3.1) (Phadke *et al*. 1981).

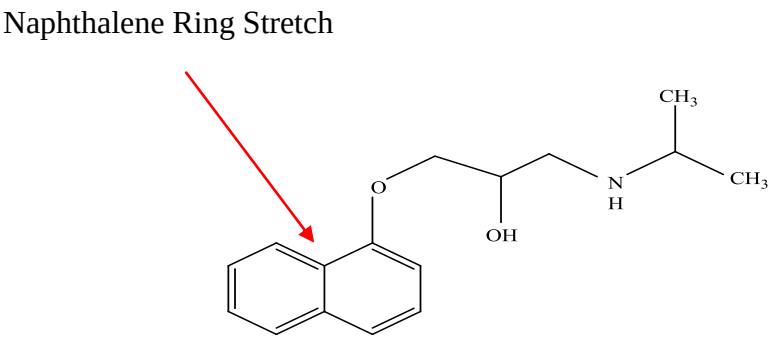


Fig. 3.1 Molecular structure of propranolol

In this study we present a detection of propranolol in plasma using the SERS technique. In plasma, propranolol binds to glycoprotein and albumin (Alvan *et al*. 1983) which makes propranolol difficult to detect at small concentrations in plasma without using a separation technique. SERS, however, requires no sample preparation and can detect small concentrations within complicated samples despite the introduction to the SERS phenomenon of other interactions from biofluids.

3.3 Materials and Methods

$HAuCl_4$ was purchased from Sigma Aldrich UK, KCl was purchased from Avocado Research Chemicals Ltd, propranolol hydrochloride was purchased from Fisher Scientific UK, human plasma and UltraPure water were purchased from Cayman Europe. All reagents were used without further purification.

The gold sol was prepared according to the procedure proposed by Grabar *et al*. (Grabar *et al*. 1995). 97 mg of $HAuCl_4$ was dissolved in 200 mL deionised H_2O (dH_2O) and then boiled. 6 mL of a 1% sodium citrate solution was then added dropwise while boiling. The solution was left boiling for a further 30 minutes with continuous stirring and a dark red colour was noted. After cooling the colloid was kept in a dark cupboard until use. A 0.500 M stock solution of KCl was prepared using dH_2O. Stock solutions of human plasma spiked with propranolol were prepared using UltraPure water to concentration levels between 1300 ng/ml and 6.5 ng/ml.

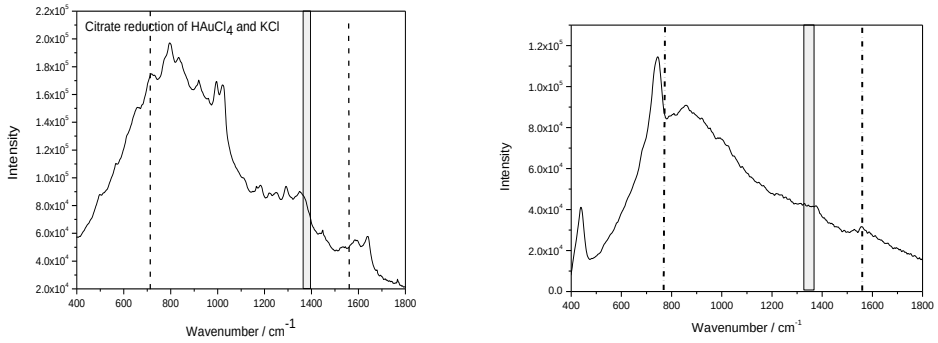
3.3.1 Instrumentation

SERS spectra were recorded on a NIR Advantage Series Raman spectrometer (DeltaNu, Laramie, WY, USA) emitting up to ~60 mW of 785 nm radiation with 23 s integrations. Acquired data was saved in the NuSpec ASCII XY format and was exported to a PC running Windows 7 and converted to xls format and csv format where required, for analysis and data processing. In order to take into account scattering and chemical variability within the samples and compare the results in an unbiased manner, an extended multiplicative scatter correction (EMSC) (Martens *et al*. 2003) process was applied as a filter to remove the variance within the spectra. All calculations were performed in R version 2.13.1 (2011): A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria). The results were plotted, analysed and interpreted using OriginPro 8 software (OriginLab Corp., Northampton, MA, USA).

3.4 Results and Discussion

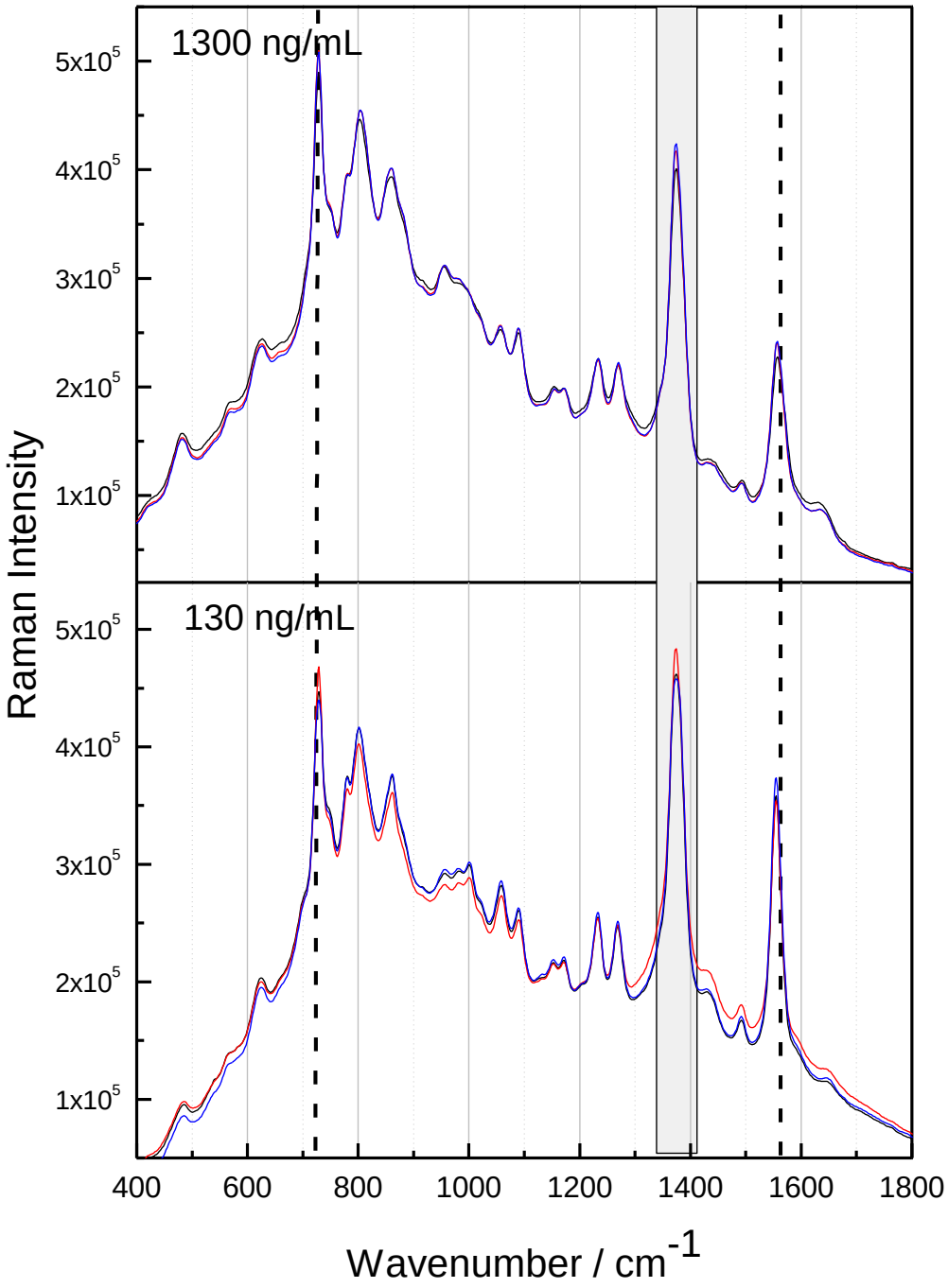
In order to in investigate the potential of SERS for spectral fingerprinting of the *β*-blocker, propranolol, in complex samples, plasma was spiked with propranolol at concentration ranges between 1300 - 6.5 ng/mL. The specific concentrations ranges were chosen based on our previous LOD work. The spiked plasma was then added to 900 μL of citrate reduction of gold colloid and 20 μL of a 0.500 M solution of KCl, the solution was then made up to a final volume of 1mL with dH_20. The spectra of the colloid with the aggregating agent and the colloid and plasma are shown in Fig

3.2, the spectra for propranolol in H_2O and the corresponding spectra for the plasma spiked with propranolol are shown in Fig. 3.3.

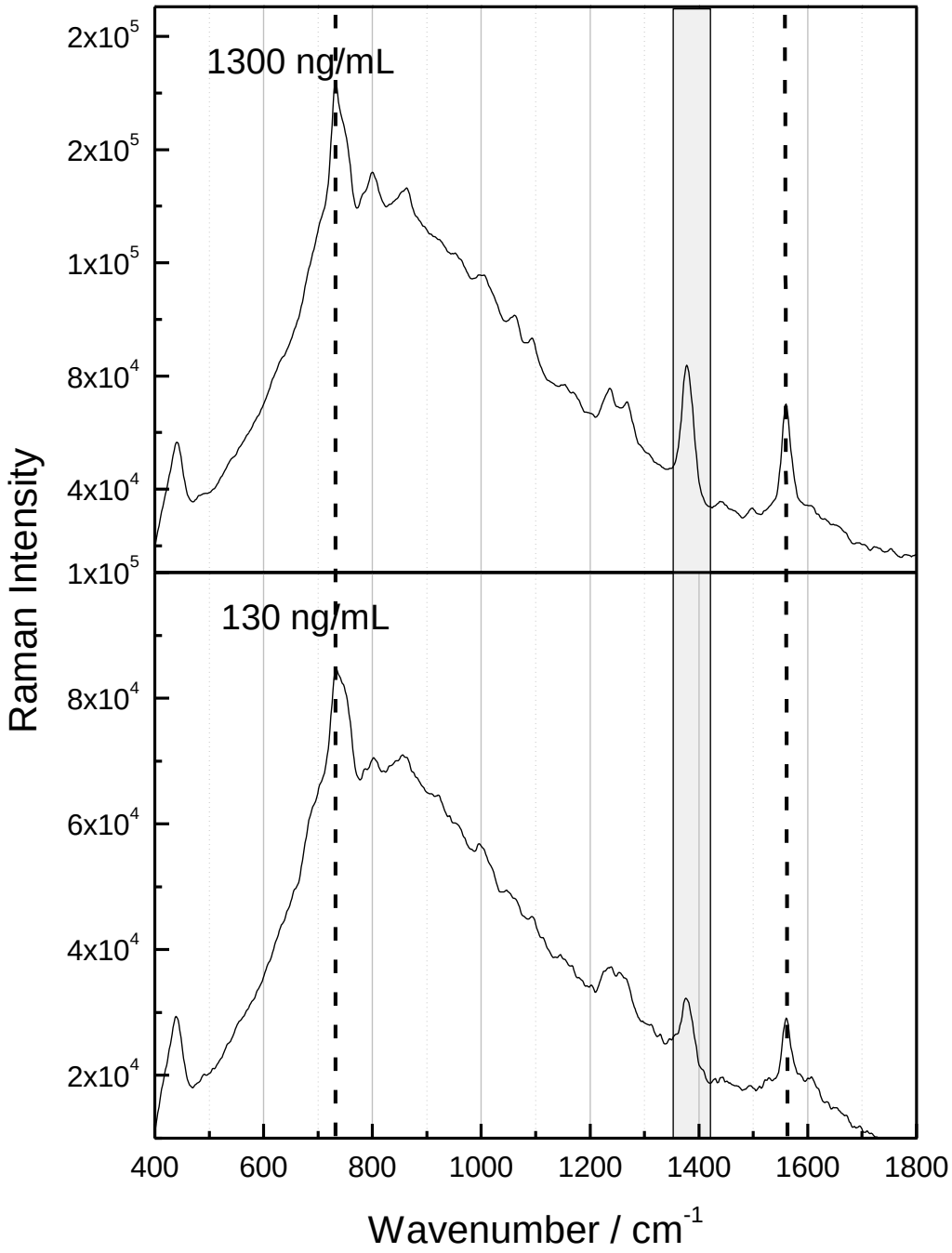


(A) (B)

Fig.3.2 (A); SERS spectrum of the colloidal system only. (B); SERS spectrum of plasma only. The dashed lines represent two other areas of the spectrum in which characteristic bands of propranolol are observed. The shaded area showing no peak in the range of interest for the propranolol naphthalene ring stretch which is observed at ~1375 cm^{-1}.


1300 ng/mL
130 ng/mL
Raman Intensity
5x10^5
4x10^5
3x10^5
2x10^5
1x10^5
400
600
800
1000
1200
1400
1600
1800
Wavenumber / cm^-1

(I)



(II)

Fig. 3.3 Spectra of propranolol in: (I) dH_2O (II) plasma. The shaded area displays the naphthalene ring stretch at ~1377 cm^{-1}. Background subtraction has been performed on the spectrum for propranolol in plasma.

In order to detect propranolol in plasma we looked for the characteristic naphthalene ring stretch that is observed in the propranolol spectrum at ~1375 cm^{-1} (Rupérez & Laserna 1996). During this investigation we prepared several concentrations of plasma spiked with propranolol and present two spectra, for the lowest concentration detection we achieved and for the highest concentration we prepared (see Fig. 3.3), the spectra that showed no SERS signal are not presented. It can be clearly seen that the characteristic peak at ~1377 cm^{-1} for propranolol detection in plasma is observed for the concentration of 130 ng/mL. Two other characteristic Raman bands for propranolol are also observed at ~736 cm^{-1} and ~1558 cm^{-1}. The reproducibility of these experiments were improved through the process of optimization of the experimental conditions for SERS of propranolol, however, the LOD we observed for propranolol in dH_2O, 2.4 ng/mL, was not observed for plasma spiked with propranolol. This is because the large local field enhancements may lead to variability in the Raman scattering signal due to at least some of the diverse molecular species present in plasma interacting with the metal surface and changing the orientation of propranolol, or restricting the propranolol molecule from absorption on the metal surface.

Plasma is the most complex human-derived proteome and is very difficult to characterize on account of the heterogeneity of the predominant glycoprotein, the high proportion of albumin (35 – 50 mg/ml) and the abundance of other proteins (Anderson & Anderson 2002). As well as dissolved proteins other molecular species found in plasma include; CO_2, glucose, clotting factors, sodium, potassium, calcium and chloride ions. The molecular species found in plasma can affect the SERS signal for a number of reasons. The phosphate content of plasma can affect the aggregation

process and as we previously modelled the optimization of experimental conditions using propranolol dissolved in dH_2O, the consideration that plasma is an ionic species which we have not previously modelled has to be taken forward in the quest for optimization of experimental conditions for SERS of propranolol in biofluids. Propranolol also binds to albumin and glycoprotein which reduces the effective free propranolol in plasma. A study carried out by Ridente *et al*. found that in a concentration of 10^{-5} M monomeric amphotericin B suspension the free drug concentration reduced to below 10^{-8} M in the presence of plasma (Ridente *et al*. 1999). In order to take the optimization process forward a separation technique is required to determine what the concentration is of the free drug in the presence of plasma. This is a consideration we expected as we are using a different solvent for propranolol, we are aware that further optimization of the experimental conditions for plasma spiked with propranolol is required as often molecules that are active in one substrate can be found to be inactive in another substrate or even totally inactive. Thus, as we are using a different solvent there are many parameters to consider when optimizing propranolol for biofluids.

3.5 Conclusion

Propranolol will compete with the molecular species found in plasma for adsorption close to the metal surface which can affect the SERS response for propranolol. As well competing for the metal surface propranolol will form a protein-ligand complex with glycoprotein and albumin which is found in plasma and reduce the effective free drug. Previous studies with a different analyte showed a reduction of three orders of magnitude of the free drug (Ridente *et al*. 1999). With these limitations in

mind the concentration that we detected, 130 ng/ml, could potentially be reduced to 0.13 ng/ml, which would be well within physiological concentrations. Taking this process forward by calculating the dissociation constant (K_d), the effective free drug can be calculated and the concentration reduction achieved. Another consideration would be to increase the concentration of the free drug in propranolol by using a technique to remove the molecular species found in plasma or increase the concentration of propranolol to account for protein-ligand binding.

From our previous studies where we demonstrated that a MOEA can be used to improve the experimental protocol needed for robust SERS spectra collection of propranolol the approach will be developed for the continuing optimization of SERS for propranolol in other biofluids. However, the approach developed will be generic and applicable to many drug compounds. By employing the MOEA to evolve new experimental protocols for analyte detection and quantification the range of samples that can be studied will increase following this proof-of-concept data, where we have detected propranolol at a physiological concentration even with the limitations we experienced in our experimental design.

3.6 References

Albrecht M.G. & Creighton J.A. (1977) *J. Am Chem. Soc.* **99**: 5215-5217

Alvan G., Bergstrom K., Iselius L. & Perdersen N. (1983) *J. Clin. Pharmacol.* **25**: 437-441

Ambrose R.T., Ketchum D.F. & Smith J.W. (1983) *Clin. Chem.* **29**: 256-259

Anderson N.L. & Anderson N.G. (2002) *Mol. Cell Proteomics.***1**: 845-867

Bakker T.C., Schut Puppels G.J., Kraan Y.M., Greve J., van der Maas L.L. & Figdor C.G. (1997) *Int. J. Cancer* **74:** 20-25

Berger A.J., Koo I., Itzkan I., Horowitz G. & Field M.S. (1999) *Appl. Opt.* **38**: 2916-2929

Bristow M.R. (2000) *Circulation.* **101**: 558-569

Campion A. & Kambhampati P. (1998) *Chem. Soc. Rev* **27**: 241-250

Coello Coello C.A., Van Veldhuizen D.A. & Lamont G.B. (2002) *Evolutionary Algorithms for Solving Multi-Objective Problems*, Kluwer Academic Publishers: Boston

Davidson J.R.T., Weisler R.H., Butterfield M.I., Casat C.D., Connor K.M., Stewart B. & van Meter S. (2003) *Biol. Psychiatry.***53**(2): 188-191

Dieringer J.A., McFarland A.D., Shah N.C., Stuart D.A., Whitney A.V., Yonzon C.R., Young, M.A., Zhang X. & Van Duyne R.P., (2005) *Faraday Discussion*. **132:** 9–26.

Farquharson S., Shende C., Inscore F.E., Maksymiuk P & Gift A. (2005) *J. Raman Spectrosc.* **26**: 208-212

Garcia-Vidal F. J. & Pendry J. B., (1996) *Phys. Rev. Lett.* **77**(6): 1163-1166

Goodacre R. (2003) *Vib. Spectrosc.* **32:** 33-45

Grabar K.C., Freeman R.G., Hommer M.B., Natan N.J., (1995) *Anal. Chem.* **67**: 735-743

Handl J. *et al*. *IEEE Trans. Comp. Biol. Bioinformatics.*(2007) **4**(2): 279-292

Jeanmarie D.L. & Van Duyne J. (1977) *Electroanal. Chem.* **84**: 1-20

Johansson J., Sparén A., Svensson O., Folestad S. & Claybourn M. (2007) *Appied. Spectroscopy.* **61**(11): 1211-1218

Keir R., Igata E., Arundell M., Smith W.E., Graham D., McHugh C., & Cooper J.M. (2002) *Anal. Chem.* **74**: 1503-1508

Kho K.W., Mei Qing K.Z., Shen Z.X., Ahmad I.B., Chin Lim S.S., Mhaisalkar S., White T.J., Watt F., Soo K.S. & Olivo M. (2008) *J. Biomedical. Optics* **13**(5): 054026-1-054026-9

Kneipp K., Kneipp H., Itzkan I., Dasari R.R., & Feld M.S., (2002) *J. Phys. Condens. Matter* **14**: R597-624

Ludvigsson J. (1974) *Acta Neurol.Scandinav.***50**: 109-115

Mahajan P., Grech E. D., Pearson R. M., Ridgway E. J. & Turner P. (1984), *Br. J. Clin. Pharmacol.* **18**: 849-852

Martens H., Nielsen P. J. & Engelsen S. B. (2003) *Anal. Chem.* **75**: 394–404.

Moskovits M., (1985) *Reviews of Modern Physics.* **57**(3): 783-826

Mullett W.M., Martin P. & Pawliszyn J. (2001) *Anal. Chem*. **73**: 2383-2389

Murillo Pulgarín J.A., Molina A.A. & López P.F. (1998) *Anal. Chimica. Acta* **370**(1): 9-18

Olson L.G., La Y.S., Beebe T.P. & Harris J.M (2001) *Anal. Chem.***73**: 4268-4276

Otto A. *Light Scattering Solids IV Electronic Scattering, Spin Effects, SERS and Morphic Effects* ed. by Cardona M. & Gunthrodt G. (1984) Springer Verlag: Berlin

Persson B.N.J. (1981) *Chem. Phys. Lett.* **82**: 561-565

Phadke R.S., Vasanth Kumar N., Hosur R.V., Saran A. & Govil G. (1981) *Int. J. Quantum Chem.* **20**(1): 85-92

Pine M., Favrot L., Smith S., McDonald K. & Chidsey C.A. (1975) *Circulation.* **52**:886-893

Premasiri W.R., Clarke R.H. & Womble M.E. (2001) *Lasers Surg. Med.* **28**: 330-334

Qian X.-M. & Nie S.M. (2008) *Chem. Soc. Rev.* **37**: 912-920

Ridente Y., Aubard J. and Bolard J. (1999) *FEBS Letters.* **446**(2): 283-286

Rupérez, A., Laserna, J.J., (1996) *Anal. Chim. Acta.* **335**: 87–94

Sanchez-cortes S., Garciaramos J.V., Morcillo G. & Tintl A. (1995) *J. Colloid Interface Sci.* **175**: 358-368

Smith E. & Dent G., (2005) *Modern Raman Spectroscopy: A practical Approach.* Jon Wiley & Sons: Chichester UK.

Strehie K.R., Cialla D., Rosch P., Henkel T., Kohler M. & Popp J. (2007) *Anal. Chem.* **79**: 1542-1547

Vo-Dinh L.R., Allain D.L. & Stokes J. (2002) *J. Raman Spectrosc.* **33**: 511-516

Xie W., Qiu P. & Mao C. (2011) *J. Mater. Chem.* **21**: 5190-5202

Xu H.X. & Käll M. (2003) *J. Phys. Chem.B* **4**: 1001-1005

Xu H.X., Aizpurua J., Käll M. & Appel P. (2000) *Phys. Rev. B.* **62**: 4318-4325

Yea K.H., Lee S., Kyong J.B., Choo J., Lee E.K., Joo S.W. & Lee S. (2005) *Analyst (Cambridge UK)* **130**: 1009-1011

Yonzon C.R., Jeoung E., Zou S., Schatz G.C., Mrksich M. & Van Duyne R.P. (2004) *Anal. Chem.* **76**: 78-85

Raman and ROA studies of acetylation in DAAPs: *α*- and *β*-*N*-acetyl-L-Asp-L-Glu

This work is due to be submitted (June 2012) to the Journal of Physical Chemistry B.

Personal contribution to the paper

I was responsible for the experimental work and the DFT calculations were provided by Dr. Trevor Dines from the University of Dundee. The diamino acid samples were provided by Dr. Nighat Kausar and Professor Babur Chowdhry from University of Greenwich. All sample preparation and measurement were performed by me and analysis in OriginPro was performed by me. The computational section of the paper was written by Dr. Trevor Dines and I wrote all other sections of the paper. The discussion and conclusions were drawn from the results, a review of the literature and discussions with Dr Ewan Blanch, Dr Trevor Dines and Professor Babur Chowdhry.

Chapter 4

Raman and ROA studies of acetylation in DAAPs: *α*- and *β*-*N*-acetyl-L-Asp-L-Glu

Clare Levene, [1] Trevor Dines, [2] Saeideh Ostovar pour, [1] Nighat Kausar, [3] Babur Z. Chowdhry, [3] and Ewan Blanch [1]*

[1] *Faculty of Life Sciences and Manchester Interdisciplinary Biocentre, University of Manchester, 131 Princess Street, Manchester, M1 7DN UK.*

[2] *Division of Electronic Engineering and Physics, University of Dundee, Dundee, DD1 4HN, UK.*

[3] *School of Science, University of Greenwich at Medway, Central Avenue, Chatham Maritime, Kent, ME4 4TB, UK.*

[*] *Correspondence to Ewan Blanch, University of Manchester, MIB, 131 Princess Street, Manchester, M1 7DN UK. E-mail*: E.Blanch@manchester.ac.uk

4.1 Abstract

We have carried out Raman and Raman optical activity (ROA) spectroscopic studies for the diamino acid peptide derivatives; *α*- and *β*-*N*-acetyl-L-Asp-L-Glu in solution. The results have been compared to nonacetylated L-Asp-L-Glu to identify the effects of site-specific acetylation on structure and ROA markers for these constitutional isomers. To help us assign the spectral features we have compared the experimental spectra to density functional theory (DFT) calculations, carried out using the *Gaussian 09* program with the B3-LYP method and the AUG-cc-pVDZ basis set. Although the Raman spectra for *α*- and *β*-*N*-acetyl-L-Asp-L-Glu are quite similar, significant differences have been observed in the ROA spectra establishing that the combination of experimental and computational ROA is a powerful technique for determining conformational preferences of acetylated peptides.

Keywords: Raman spectroscopy; ROA; DFT calculations; acetylation and *α*- and *β*-*N*-acetyl-L-Asp-L-Glu

4.2 Introduction

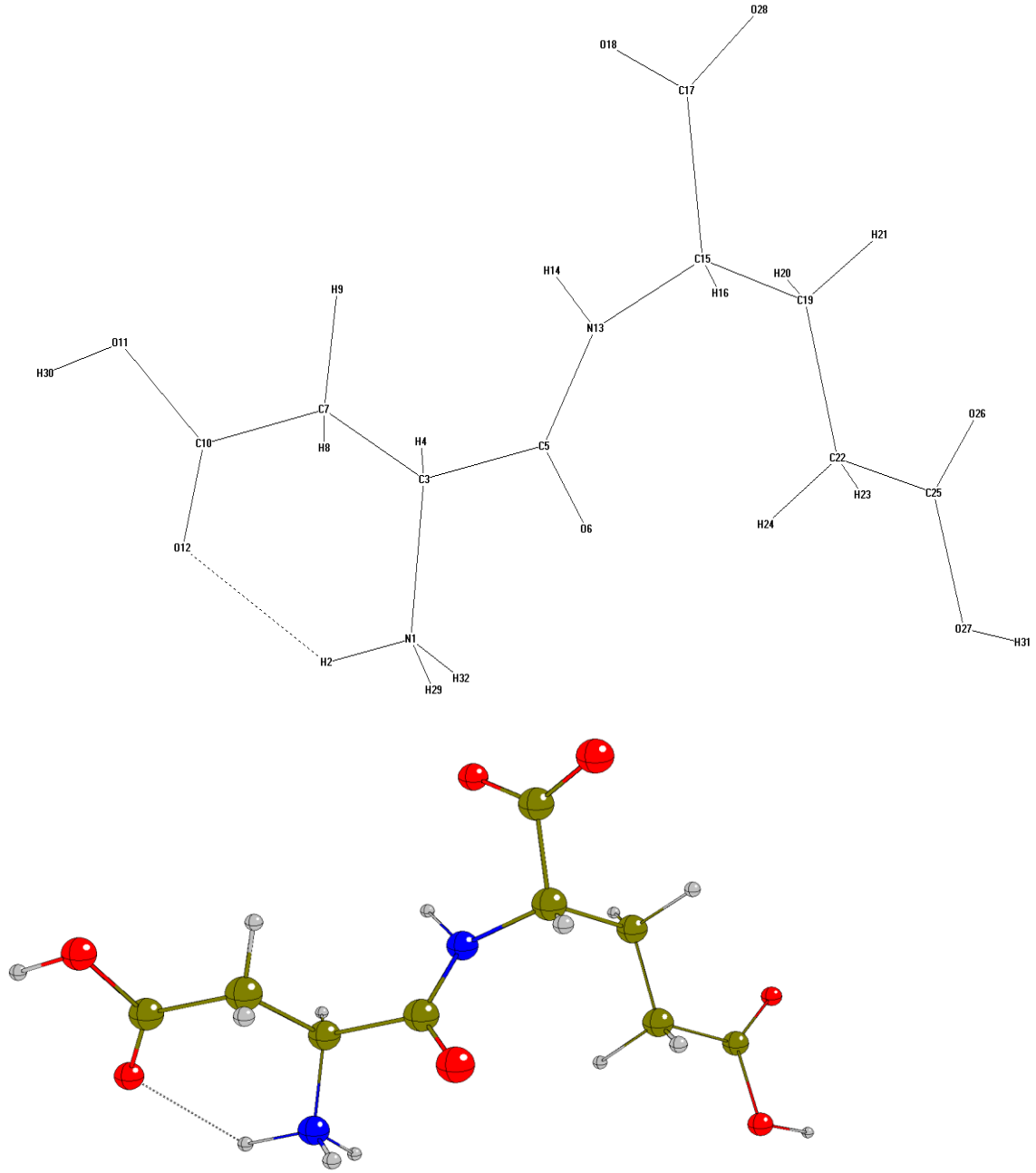
The most ubiquitous naturally occurring diamino acid peptide (DAAP) in the vertebrate central nervous system (CNS) is *α-N*-acetyl-L-Asp-L-Glu (*α*-NAAG). High micromolar and low millimolar concentrations of *α*-NAAG were first identified in horse brain and spinal cord tissue (Curatolo, Arcangelo and Lino 1965). In 1966 its presence was confirmed in bovine brain (Miyamoto, Kakimoto and Sano 1966), and then later in the brain tissue of other mammalian species (Moffett 1994, Kowalski 1987, Tieman 1987). Studies have been carried out into *α*-NAAG's biological activity as an excitory neurotransmitter, anti-histamine and mast cell de-granulator (Goldschmidt *et al* 2007, Kandel 2000). In its magnesium salt form *α*-NAAG has been approved as an over-the-counter drug in certain European countries and as an active agent against conjunctivitis and allergic rhinitis. *α*-NAAG is also used as an anti-allergenic medication in eye drops and nasal preparations (Althaus 1994). *α*-NAAG is catabolised to *N*-acetyl-L-aspartic acid and the potentially excitotoxic neurotransmitter L-Glu by the enzyme glutamate carboxypeptidase II, also referred to as *N*-acetylated-*α*-linked acidic dipeptidase (NAALADase), in both normal and pathophysiological conditions of neurological systems (Stauch 1989). Excitotoxicity is a pathological process whereby nerve cells are damaged and killed by excessive stimulation by neurotransmitters, and has been implicated in neurodegenerative diseases and neuropsychiatric disorders such as multiple sclerosis, Alzheimer's disease, amyotrophic lateral sclerosis, schizophrenia, seizure disorders and Huntingdon's disease (Tsai 1996, Neale 2000). *β-N*-acetyl-L-Asp-L-Glu (*β*-NAAG) has been less thoroughly studied than *α*-NAAG, however, *β*-NAAG may well have interesting physiological properties (Anuradah 1997). It is also known to

be a competitive inhibitor of NAALADase activity and may also act as a mimetic at some of the locations of α-NAAG pharmacological activity (Yourich 2003).

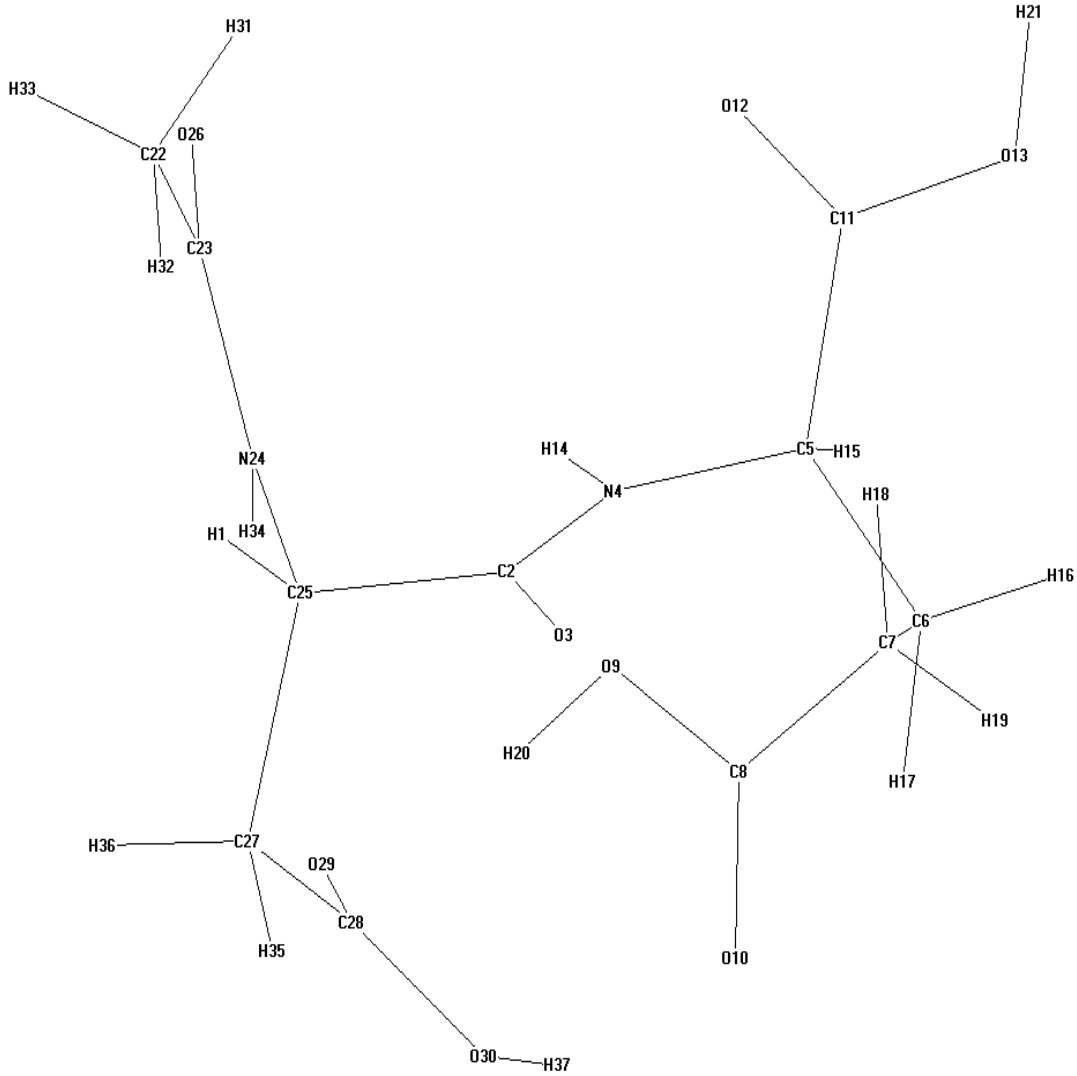
Acetylation is one of the most common post-transcriptional modifications and the identification of acetylation sites is the essential first step towards studies relating to transcriptional regulation (Mizzen & Allis 2000), gene control (Briggs *et al*. 2002) and cancer therapies (Brown & Strathdee 2002; Johnstone 2002). Acetylation is an endogenous bioprocess that substitutes an active hydrogen atom for an acetyl group (CH_3CO) at the N-teminal end with the modification being performed by N-alpha-acetyltransferases (NATs). By determining the native structures of dipeptides and their acetylation sites, for which a great deal of experimental and theoretical studies have been performed, their unique functional properties can be understood in terms of their relationship to their 3D structure (Chasse *et al* 2001). Molecular structure and dynamics of diamino acid peptides (DAAPs) can be determined using vibrational spectroscopy, which is a powerful group of techniques as is evident from numerous IR and Raman studies (Tiffany & Krimm 1969; Mirkin & Krimm 2004; Dixon *et al*. 1994; Kausar *et al* 2009a/b/c). The most frequently used bands for structural determination are the amide I and the amide III vibrations (Kauser *et al* 2009a). The amide I and III regions of vibrational spectra of peptides have been shown to be useful for determining backbone conformations (Grdadolnik 2008). To simulate the early stages of the protein folding process, a crucial area of research in molecular biophysics, the ability to calculate the relative energies of the various backbone conformations is essential (Kinalwa *et al.* 2010; Grdadolnik 2011).The significance of local geometry on the conformation of *N*-substituted amino acids and the importance of internal hydrogen bonding can be investigated using density

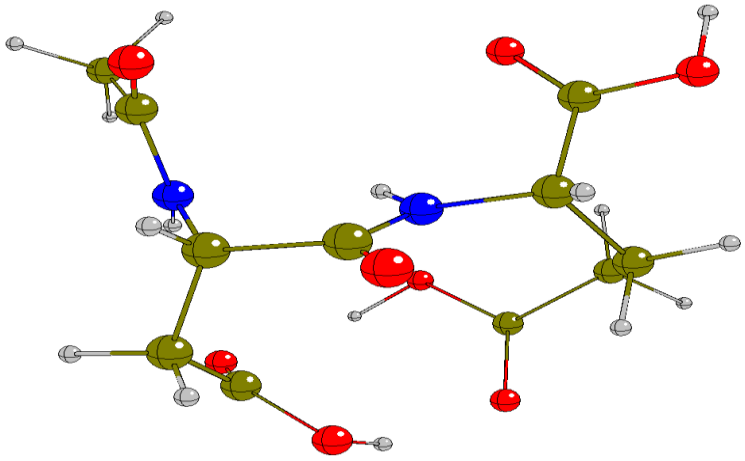
functional theory (DFT) calculations (Kauser 2007a/b/c). To determine the accuracy and reliability of results from DFT calculations high quality experimental results sensitive to the major backbone conformations are required (Elstner and Hermans 2003).

In this study we have used Raman optical activity (ROA), which measures a small difference in Raman scattering by chiral molecules using left- and right-circularly polarized scattered light (Atkins & Barron 1969; Barron, Bogaard & Buckingham 1972; Hug & Hangartner 1999), and is highly sensitive to molecular structure, as well as Raman spectroscopy to investigate the solution structures of α- and β-*N*-acetyl-L-Asp-L-Glu. With the combination of experimental and computational techniques our aim was to provide insight into the molecular conformation in solution of these two di-amino acid peptides. Isomers are difficult to distinguish using normal vibrational techniques, whereas ROA is sensitive to absolute stereochemistry (Barron *et al* 2000). Determining the conformational preferences of DAAPs in aqueous solution is essential for understanding their structure and folding (Grdadlonik 2008) and we therefore present a combined Raman and ROA study characterizing the structural differences between these two isomers. The combination of experimental Raman and ROA spectra with DFT calculations has allowed us to make detailed assignments of spectral features to the vibrational modes responsible. We have also compared the spectra of the acetylated DAAPs α- and β-*N*-acetyl-L-Asp-L-Glu to those of L-Asp-L-Glu so as to identify the effects of site specific acetylation.

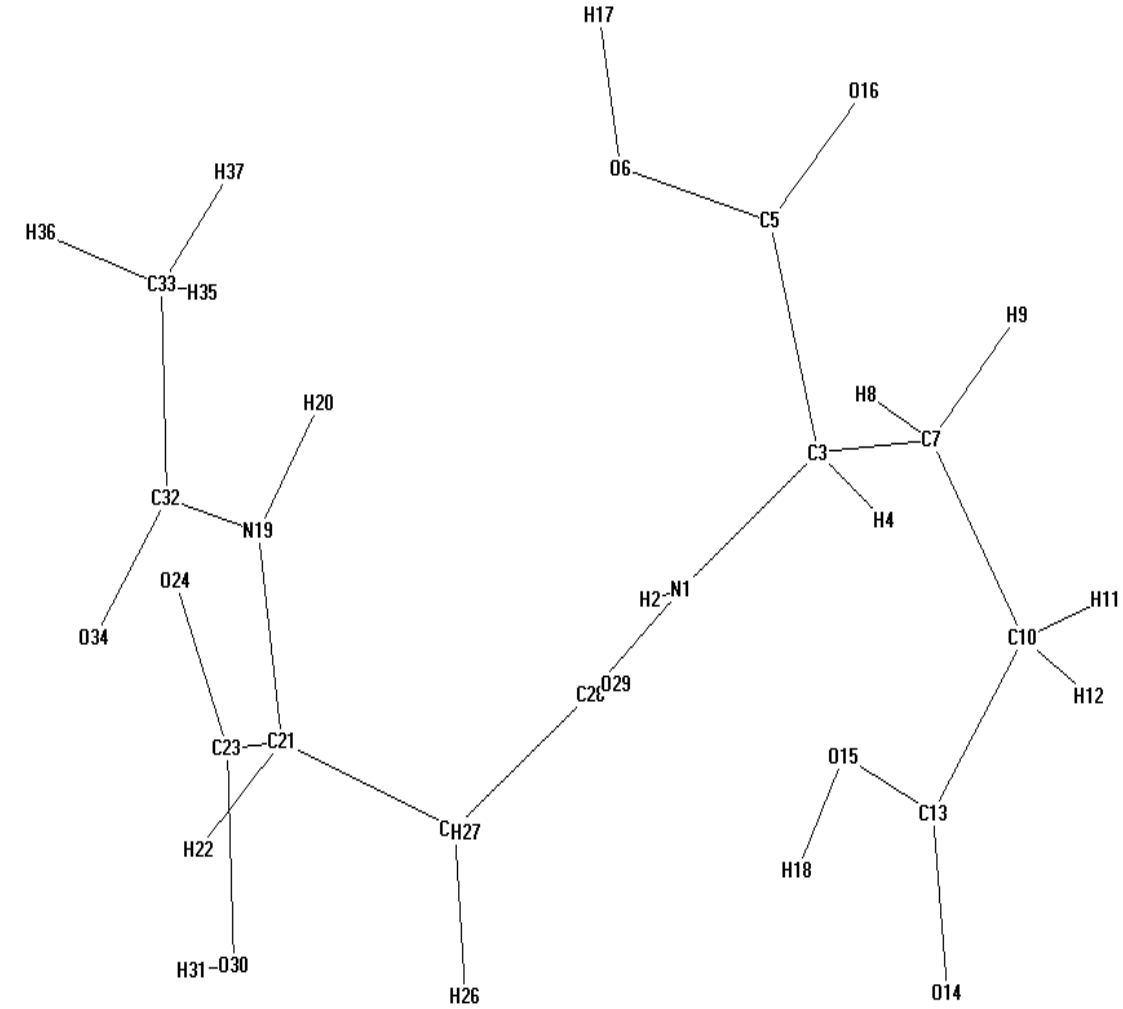

O28
O18
C17
H21
C15
H14
H20
H16
C19
H9
N13
O11
H30
C7
C10
H8
H4
C5
C3
O26
C22
H23
C25
H24
O6
O12
O27
H31
N1
H2
H32
H29

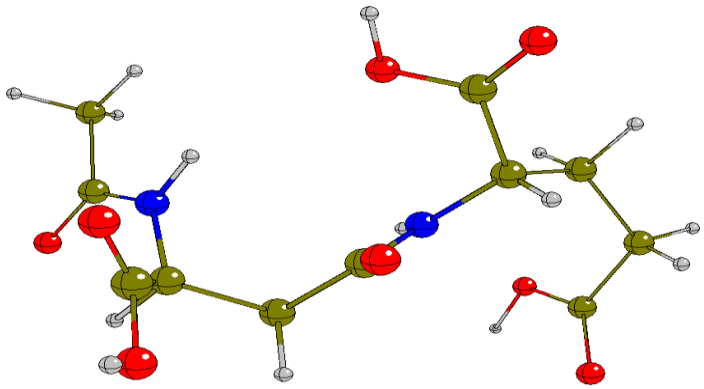
(I)


H31
H33
C22
O26
C23
H32
H21
O12
O13
C11
N24
H14
C5
H15
N4
H18
H1
H34
C2
C25
H16
O3
C6
C7
O9
H19
H20
C8
H17
H36
C27
O29
C28
H35
O10
O30
H37



(II)





(III)

Fig. 4.1 Chemical structures, atom numbering and computed molecular geometries for: (I) L-Asp-L-Glu, (II) *α*-*N*-acetyl-L-Asp-L-Glu and (III) *β*-*N*-acetyl-L-Asp-L-Glu.

4.3 Experimental

The dipeptides α- and β-*N*-acetyl-L-Asp-L-Glu and L-Asp-L-Glu (ca. 99.0% purity) were purchased from Bachem (St Helens, UK) and used without further purification. All samples were prepared in 100µl of a phosphate buffer solution (0.20M, pH 7.01) to a final concentration of 50 mg/ml and centrifuged for 5 mins at 3000 rpm to remove insoluble impurities such as dust particles prior to data collection. Deuterated samples were also prepared using D_2O under the same conditions as for samples in H_2O. Samples were then pipetted into quartz microfluorescence cells (Optiglass, Ilford, Essex, UK) for spectroscopic measurement.

Raman and ROA spectra were collected using a BioTools ChiralRAMAN spectrometer (BioTools Inc., Jupiter, FL, USA) operated via Critical Link software. The instrument has a design based on the backscattering geometry using a Nd/VO_4 laser with an excitation wavelength of 532 nm and spectral resolution of ~7 cm^{-1}. Laser power at the samples was 600 mW with an illumination time of 1.9845 seconds per scan and overall data accumulation times of 6-24hrs. The results were plotted, analysed and interpreted using OriginPro 8 software (OriginLab Corp., Northampton, MA, USA). The baseline tool in OriginPro 8 was used to perform baseline corrections, with regions of the spectra where no bands were observed being adjusted to zero intensity. All spectra were normalised with respect to experimental parameters such as accumulation time to allow for quantitative comparison between samples.

4.4 Computational methods

DFT calculations were carried out using the *Gaussian 09* program (Frisch *et al*. 2009) with the B3-LYP method (Becke 1993; Lee *et al*. 1988) and the AUG-cc-pVDZ basis set (Dunning 1989). All calculations were performed with the IEF-PCM solvation method (Cances *et al*. 1997) using the Karplus and York continuous surface charge formalism (York & Karplus 1999), with water as solvent and default polarizable continuum model (PCM) parameters were used except that Pauling atomic radii were substituted for the default UFF radii. The starting geometries for geometry optimisation were based upon those obtained from x-ray crystal structures reported in our previous publications (Kausar *et al*. 2009A; Kausar *et al*. 2009B; Kausar *et al*. 2009C).

Vibrational spectra were calculated for the optimised geometries and Raman and ROA activities were computed dynamically for an excitation wavelength of 532 nm. Relative Raman and ROA intensities were calculated from the computed Raman and ROA activities using the equations:

$$I_{fi}^{(Raman)} = \frac{\left(\tilde{\nu}_0 - \tilde{\nu}_{fi}\right)^4}{\tilde{\nu}_{fi}\left[1-\exp\left(-\frac{hc\tilde{\nu}_{fi}}{kT}\right)\right]}\left[45a^2 + 7\gamma^2\right], \quad (1)$$

$$I_{fi}^{(\mathrm{ROA})} = \frac{\left(\tilde{\nu}_0 - \tilde{\nu}_{fi}\right)^4}{\tilde{\nu}_{fi}\left[1-\exp\left(-\frac{hc\tilde{\nu}_{fi}}{kT}\right)\right]}\left[48\beta_{G'}^2 + 16\beta_{A'}^2\right], \quad (2)$$

where $\tilde{\nu}_0$ = 18,797 cm^{-1} and T = 298.15 K.

In these equations the ytensor invariants are defined as follows:

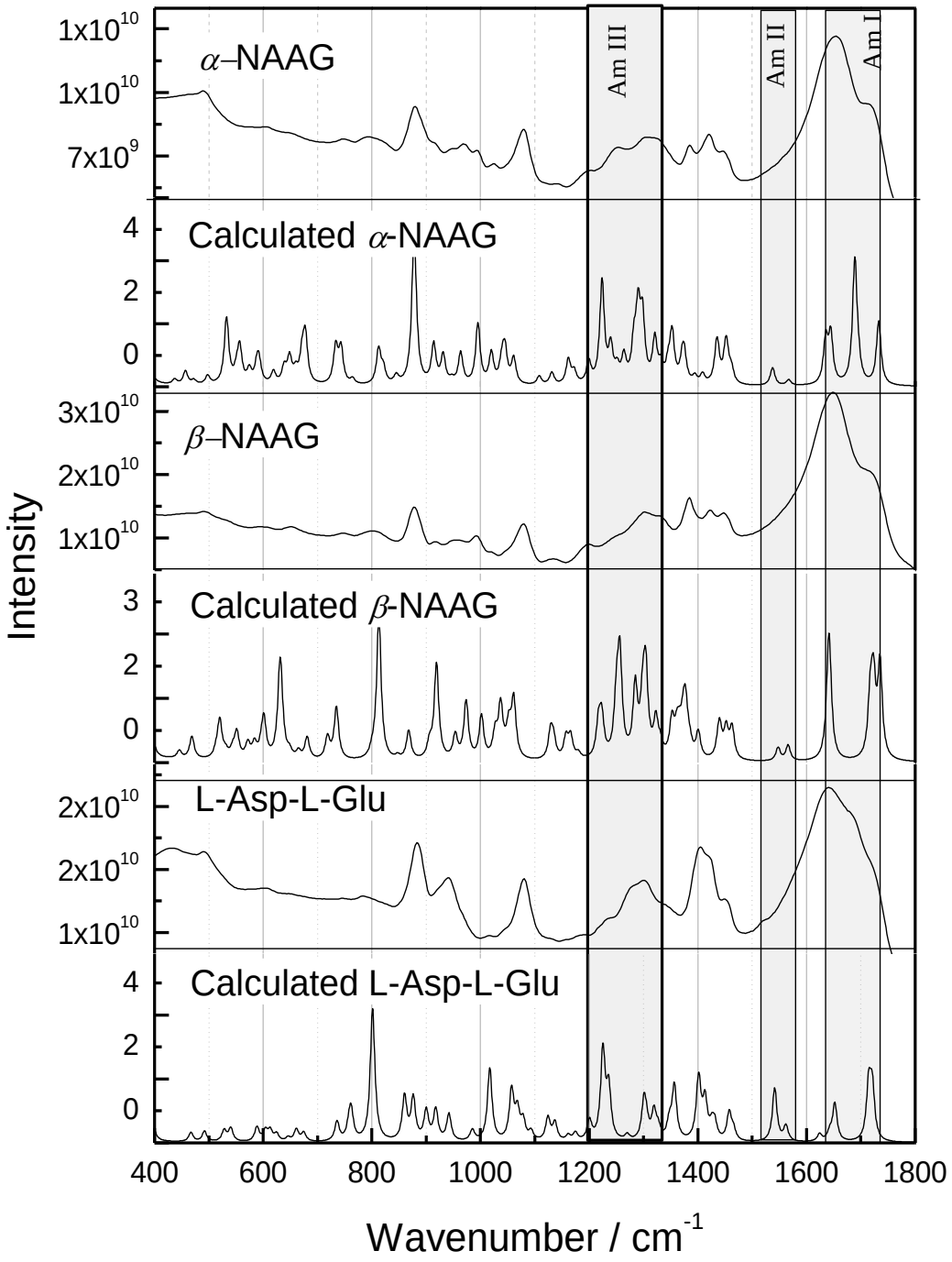
a^2 is the isotropic invariant of the electric-dipole/electricdipole polarizability tensor

γ is the symmetric anisotropic invariant of the electric-dipole/electric-dipole polarizability tensor

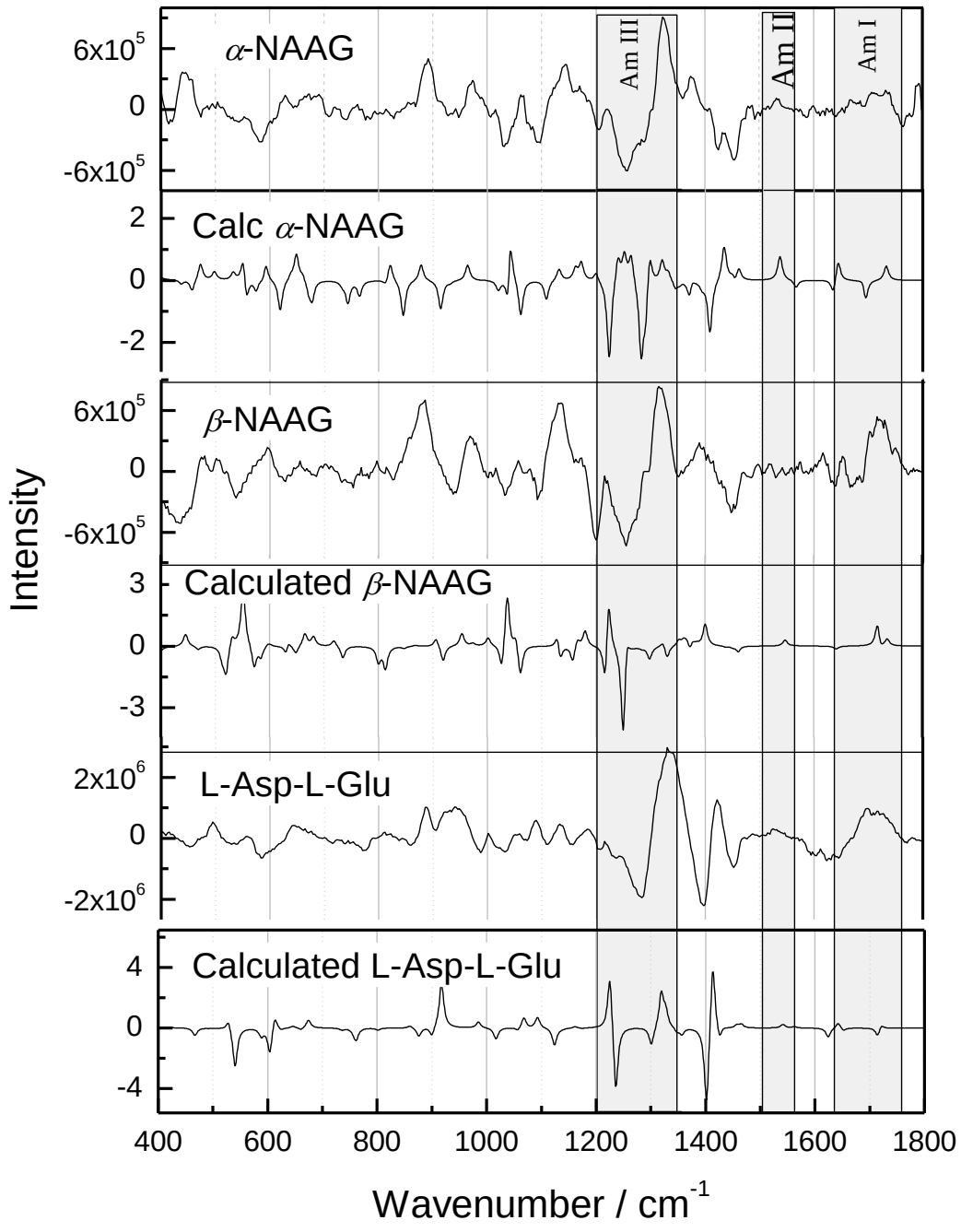
$\beta_{G'}^2$ is the anisotropic invariant of the cross-product of the electric-dipole/electric-dipole polarizability tensor with the electric-dipole/magnetic-dipole polarizability tensor

$\beta_{A'}^2$ is the anisotropic invariant of the cross-product of the electric-dipole/electric-dipole polarizability tensor with the tensor $\mathscr{A}_{\mu\nu}$ obtained by contracting the electric-dipole/electric-quadrupole polarizability tensor with the antisymmetric-unit tensor of Levi-Civita

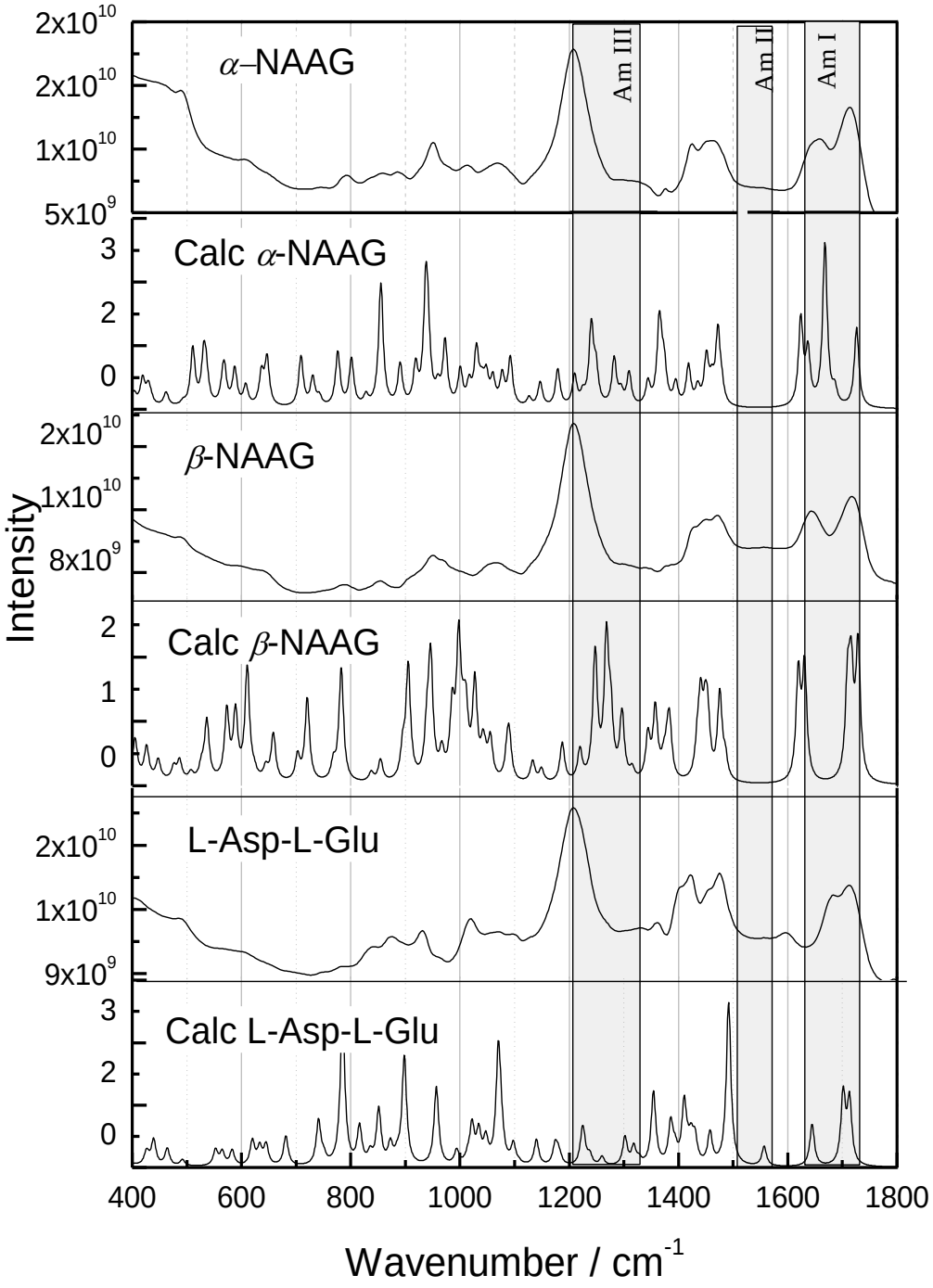
The Cartesian force constants obtained from the *Gaussian 09* output were converted to force constants expressed in terms of internal coordinates using a normal coordinate analysis program (GFXP) derived from those of Schachtsneider [i]. A full set of internal coordinates, including all bond angles and torsional angles, was reduced to a set of 3*N*-6 symmetry-adapted internal coordinates. Normal coordinate analysis was done without scaling of force constants, producing potential energy distributions for harmonic wavenumbers, according to the methods of Wilson, Decius and Cross (Wilson *et al*. 1955) Simulated Raman and ROA spectra were constructed by convolution with a Lorentzian lineshape function of 10 cm^{-1} f.w.h.m.


α–NAAG
Calculated α-NAAG
β–NAAG
Calculated β-NAAG
L-Asp-L-Glu
Calculated L-Asp-L-Glu
Am III
Am II
Am I
Intensity
Wavenumber / cm-1

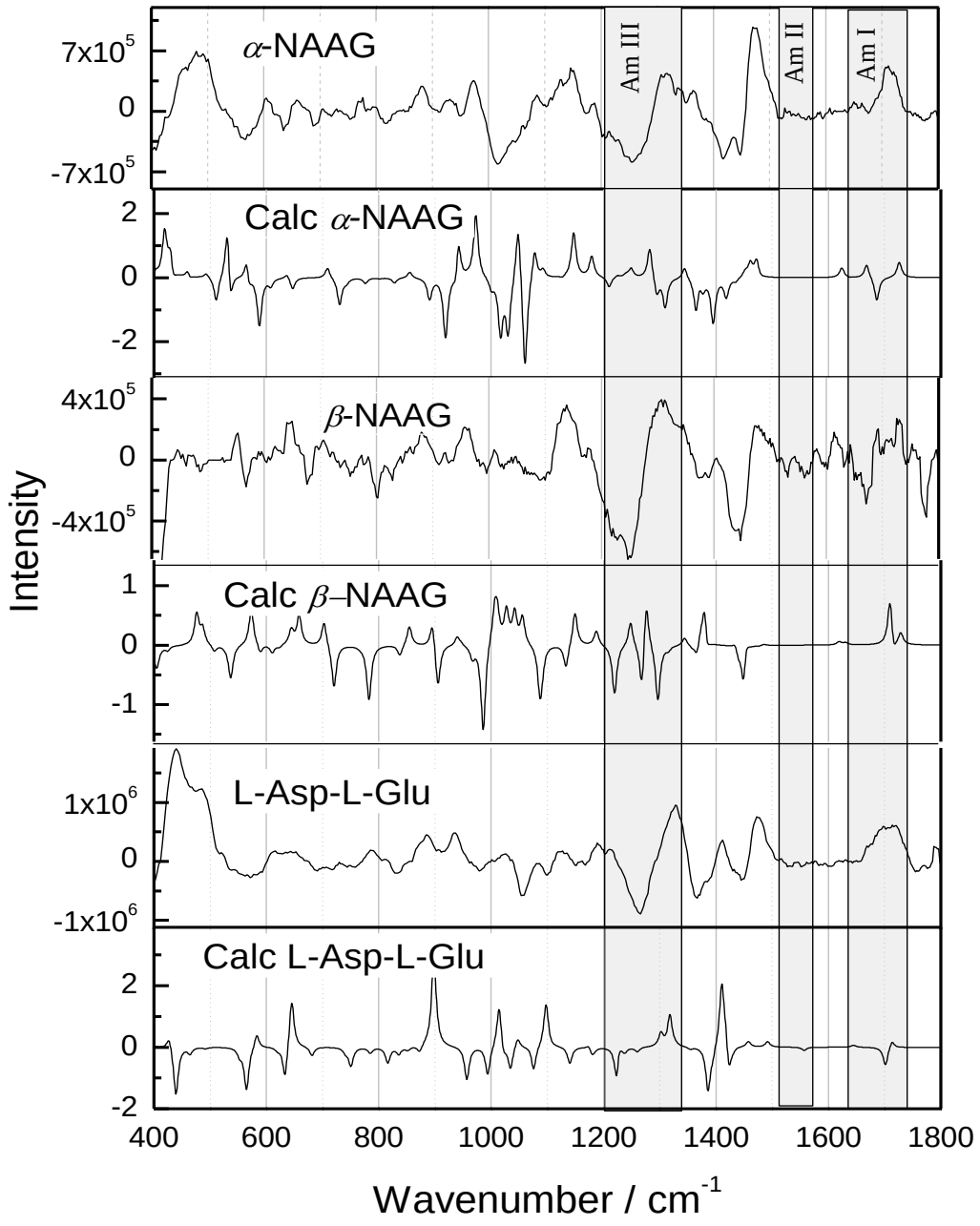
(I)


α-NAAG
Am III
Am II
Am I
6x10^5
0
-6x10^5
Calc α-NAAG
2
0
-2
β-NAAG
6x10^5
0
-6x10^5
Calculated β-NAAG
3
0
-3
L-Asp-L-Glu
2x10^6
0
-2x10^6
Calculated L-Asp-L-Glu
4
0
-4
Intensity
400
600
800
1000
1200
1400
1600
1800
Wavenumber / cm^-1

(II)


α–NAAG
Am III
Am II
Am I
Calc α-NAAG
β-NAAG
Calc β-NAAG
L-Asp-L-Glu
Calc L-Asp-L-Glu
Intensity
Wavenumber / cm-1

(III)



(IV)

Fig. 4.2 Experimental and simulated Raman and ROA, I and II for H_2O, III and IV for D_2O

4.5 Results and discussion

The atom numbering schemes for *α*- and *β*-*N*-acetyl-L-Asp-L-Glu and L-Asp-L-Glu and the computed molecular geometries are shown in Fig. 4.1. Calculated bond lengths, bond angles and selected torsion angles are shown in Tables S4.1, S4.2 and S4.3 (see supporting information). The calculated spectral features presented in wavenumbers together with their PEDs (see page 118) are compared with the experimental data for L-Asp-L-Glu and *α*- and *β*-*N*-acetyl-L-Asp-L-Glu in Tables 4.1, 4.2 and 4.3. Deuterated samples were also investigated and their calculated vibrational wavenumbers together with their PEDs are compared with the experimental data for *α*- and *β*-*N*-acetyl-L-Asp-L-Glu and L-Asp-L-Glu in Tables S4.4, S4.5 and S4.6, (see supporting information). The DFT calculations have been performed using starting geometries from previous publications (Kausar *et. al.*, a/b 2009) using B3LYP/AUG-cc-pVDZ, regarded as being capable of reproducing ROA spectra which is qualitative in respect of predicting the correct sign and approximate relative intensities (Ruud *et al*. 2002; Reiher *et al*. 2005). The DFT calculations agree reasonably although further exploration of the energy conformations is required for a better agreement. DFT calculations perform well in thermochemical predictions however, the predictions for more complicated properties such as optical rotation are complex and difficult to interpret (Crawford 2006). However, DFT, calculations are still very useful for assigning spectral features and the results have been compared to calculations that were previously performed for *α*- and *β*-*N*-acetyl-L-Asp-L-Glu in the solid form for Raman (Kausar *et al*. 2009).

4.5.1 Vibrational assignments

Both *α*- and *β*-*N*-acetyl-L-Asp-L-Glu have 105 normal vibrational modes and L-Asp-L-Glu has 90 normal vibrational modes which in principle are all Raman active, thus, given this considerable number of vibrational modes the Raman spectrum is complicated where many bands overlap resulting in strong broad bands which are difficult to differentiate. However, due to its sensitivity to chirality, ROA allows us to extract additional information from the spectrum which is useful for probing biomolecular structure and behaviour in aqueous solution. When comparing the spectra of the three DAAPs we first seek to identify bands that are specific to *α*- or *β*-*N*-acetyl-L-Asp-L-Glu and whether there are any changes that occur to the spectrum that can be assigned to acetylation when compared to the spectrum of L-Asp-L-Glu. The observed and simulated Raman and ROA spectra of *α*- and *β*-*N*-acetyl-L-Asp-L-Glu and L-Asp-L-Glu are shown in Fig. 4.2. The vibrational modes have been assigned on the basis of normal coordinate analysis derived from the computed spectra and the Raman results compare favourably with previous published literature (Kausar *et al*. 2009b).

Table 4.1. Calculated and observed wavenumber assignments (cm^{-1}) for vibrational bands of L-Asp-L-Glu.

Calculated (cm^{-1})	Experimental Raman	Experimental ROA	Assignments (%PEDs)
1721		1700	ν(C25O26) (82)
1715			ν(C10O12) (77)
1652			ν(C5O6) (58), ν(C5N13) (15), $\delta_{as}(NH_3)$ (16)
1642	1640	1610	$\delta_{as}(NH_3)$ (80)
1624			$\delta_{s}(NH_3)$ (10), $\delta_{as}(NH_3)$ (76)
1562			$\nu_{as}(CO_2)$ (68), δ_{ip}(NH) (11)
1541			$\nu_{as}(CO_2)$ (22), ν(C5N13) (26), δ_{ip}(NH) (27)
1466			$\delta_{s}(NH_3)$ (83)
1458			$\delta(CH_2')$ (93)
1431			$\delta(CH_2'')$ (88)
1426		1450	$\delta(CH_2)$ (63), δ(C3H4) (16)
1425			$\omega(CH_2)$ (39)
1413		1418	$\nu_{s}(CO_2)$ (24), $\delta(CH_2)$ (20), $\omega(CH_2')$ (10)
1402			$\nu_{s}(CO_2)$ (14), $\delta(CH_2)$ (12), δ(C3H4) (37)
1400	1401	1395	ν(C22C25) (10), $\omega(CH_2'')$ (32)
1357			$\nu_{s}(CO_2)$ (17), $\omega(CH_2')$ (26), δ(C15H16) (32)
1347			δ(C10O11H30) (19), δ(C3H4) (20)
1327			$\omega(CH_2')$ (26), δ(C15H16) (43)
1319		1338	$\nu_{s}(CO_2)$ (12), δ(C1O11H30) (11), $\tau(CH_2')$ (18), δ(C15H16) (14)
1306			δ(C25O27H31) (14), $t(CH_2)$ (36), δ(C3H4) (12)
1301			δ(C25O27H31) (29), $\tau(CH_2)$ (12), $\omega(CH_2'')$ (21)
1270			$p_{op}(NH_3)$ (10), $\omega(CH_2)$ (15), $\tau(CH_2)$ (14), δ(C3H4) (12)
1236		1281	δ_{ip}(NH) (12), $\tau(CH_2'')$ (38)
1225			δ_{ip}(NH) (26), $\tau(CH_2'')$ (15), δ(C15H16) (25)
1201		1215	$\tau(CH_2')$ (34), $\tau(CH_2'')$ (10), δ(C15H16) (18)
1174		1200	ν(C10O11) (34), δ(C10O11H30) (14)
1161		1185	ν(C3C7) (14), $p_{ip}(NH_3)$ (10), $p_{op}(NH_3)$ (14), δ(C3H4) (13)
1137		1130	ν(C25O27) (37), δ(C25O27H31) (16), $\omega(CH_2'')$ (10)
1123		1111	ν(C15N13) (25), $p_{op}(NH_3)$ (10)
1093		1086	ν(C15N13) (20), $p_{ip}(NH_3)$ (18), $p_{op}(NH_3)$ (13)
1079	1081	1070	ν(C15C19) (27), ν(C19C22) (11)
1068			ν(C3C7) (13), ν(C19C22) (10), $p(CH_2'')$ (10)
1057			ν(C19C22) (35)

Table 4.1 Continued

1017			*v*(C15C17) (10), p(CH_2') (21)
984		1002	*v*(C3C7) (24), $p_{ip}(NH_3)$ (10), $p_{op}(NH_3)$ (14)
941	938	937	*v*(C3C5) (12), $p_{op}(NH_3)$ (11), p(CH_2) (18)
918			*v*(C7C10) (33)
900			*v*(C15C17) (12), p(CH_2'') (30), δ(O26C25O27) (14)
876			*v*(C3N1) (16), *v*(C22C25) (13), p(CH_2) (15)
860	883	885	*v*(C3N1) (19), *v*(C22C25) (17)
801			*v*(C22C25) (17), p(CH_2') (12), δ(O18C17O28) (22)
762			δ_{op}(C5O6) (35), ω(O18C17O28) (17)
758			p(CH_2') (11), δ_{op}(C5O6) (13), ω(O18C17O28) (37)
736			δ(C3C7C10) (10), δ_{ip}(C5O6) (16), δ_{op}(C5O6) (10)
675			ω(O11C10O12) (23), τ(C10O11) (22)
661			p(CH_2'') (11), ω(O26C25O27) (21), τ(C25O27) (24)
46			δ(O18C17O28) (22), τ(C10O11) (11), τ(C25O27) (13)
624			δ(O11C10O12) (37), τ(C10O11) (20)
613			δ(O18C17O28) (13), τ(C10O11) (14), τ(C25O27) (15)
604			δ(O26C25O27) (27), τ(C25O27) (17)
588		586	δ_{op}(NH) (35), τ(C5N13) (32)
539			δ(O11C10O12) (12), ω(O11C10O12) (20), τ(C10O11) (16)
528		495	δ (O26C25O27) (16), ω(O26C25O27) (20), τ(C25O27) (16)
492			p(O26C25O27) (24), ω(O26C25O27) (11)
467			p,(CH_2) (18), δ(N1C3C7) (16), p,(O11C10O12) (18)

v, stretching; as, anti-symmetric; s, symmetric, δ, deformations; ω, wagging; τ, twisting; p, rocking; ip,in-plane; op, out-of-plane;

Table 4.2 Observed and calculated wavenumber assignments (cm^{-1}) for α-NAAG

Calculated (cm^{-1})	Experimental Raman	Experimental ROA	Assignments (%PEDs)
1733			ν(C11O12) (84)
1695			ν(C8O10) (47), ν(C28O29) (32)
1689			ν(C8O10) (28), ν(C28O29) (33)
1645			ν(C2O3) (52), ν(C23O25) (21)
1635			ν(C2O3) (23), ν(C23O25) (47)
1567			ν(C2N4) (26), δ_{ip}(N4H14) (47)
1538		1535	ν(C23N24) (27), δ_{ip}(N24H34) (43)
1462			$\delta(CH_2')$ (81)
1456			$\delta(CH_2'')$ (60), $\delta_{as}(CH_3)$ (11)
1453		1455	$\delta(CH_2'')$ (15), $\delta_s(CH_3)$ (11), $\delta_{as}(CH_3)$ (50)
1451			$\delta(CH_2)$ (87)
1436			δ(C8O9H20) (26), δ(C28O30H37) (21), $\delta(CH_2'')$ (11)
1434			$\delta_{as}(CH_3)$ (85)
1409		1421	δ (C28O30H37) (18)
1394			ν(C11O13) (11), δ(C5H15) (10), $\omega(CH_2)$ (13)
1375	1380	1380	$\delta_s(CH_3)$ (68), $\delta_{as}(CH_3)$ (21)
1371			δ(C5H15) (22), $\omega(CH_2)$ (56)
1352		1356	δ(C8O9H20) (13), $\tau(CH_2)$ (11), $\omega(CH_2')$ (31), $\omega(CH_2'')$ (11)
1345			δ(C25H1) (71)
1332			δ(C28O30H37) (16), δ(C25H1) (13), $\omega(CH_2')$ (35)
1321			δ(C11O13H21) (18), δ(C5H15) (26)
1298		1320	ν(C23N24) (20), δ_{ip}(N24H34) (16)
1290		1257	ν(C2N4) (12), δ_{ip}(N4H14) (14), δ(C5H15) (12), $\tau(CH_2')$ (11)
1282			δ(C11O13H21) (15), δ(C5H15) (17)
1264			ν(C28O30) (13), $\tau(CH_2'')$ (16)
1251			δ(C5H15) (11) $\tau(CH_2)$ (11), $\omega(CH_2'')$ (17)
1240			ν(C8O9) (13), δ(C5H11) (20), $\tau(CH_2)$ (27)
1224		1205	δ(C25H1) (34), $\omega(CH_2'')$ (22)
1199			ν(C28O30) (15), ν(C25N24) (18), $\tau(CH_2'')$ (22)
1172			ν(C11O13) (20), ν(C5N4) (17), δ(C11O13H21) (14)
1162			ν(C8O9) (22), $\tau(CH_2')$ (38)
1131		1130	ν(C11O13) (27), ν(C5N4) (24), δ(C11O13H21) (14)
1108			ν(C25N24) (32), $\tau(CH_2'')$ (23)
1061	1078	1096	ν(C5C6) (20), ν(C6C7) (32)
1045			ν(C25C27) (15), $p_{op}(CH_3)$ (17)
1040			$p_{op}(CH_3)$ (46), δ_{op}(C23O26) (14)

Table 4.2 Continued

1038		1031	$p(CH_2)$ (22)
1020			$p_{ip}(CH_3)$ (58)
963		970	*ν*(C25C27) (14), *ν*(C22C23) (22), $p(CH_2'')$ (15)
947			$p(CH_2')$ (32), ω(O9C8O10) (11)
931			δ(C2N4C5) (10), $r(CH_2'')$ (14)
914			*ν*(C8O9) (10) *ν*(C5C6) (13), *ν*(C6C7) (12), *ν*(C7C8) (28)
878	879	890	*ν*(C27C28) (21)
845			ω(O12C11O13) (11)
821			ω(O29C28O30) (21)
813			*ν*(C5C11) (12), *ν*(C6C7) (11), n(C7C8) (14), $p(CH_2)$ (16)
765			ω(O9C8O10) (29), τ(C8O9) (25)
743			ω(O12C11O13) (26)
734			*ν*(C5C6) (10), δ_{op}(C2O3)
678			τ(C8O9) (21), τ(C28O30) (32)
672			δ(C2C25N4) (11), δ_{op}(C23O26) (19), τ(C23N24) (11)
660			t(C8O9) (40), τ(C28O30) (25)
649			*ν*(C22C23) (18), δ_{op}(C23O26) (18)
639			δ(N4C2C25) (10), δ_{op}(C23O26) (12)
619			δ(O12C11O13) (18), τ(C11O13) (22)
592			*ν*(C27C28) (16), δ(O29C28O30) (57)
588		584	*ν*(C7C8) (16), δ(O9C8O10) (59)
575			ω(O12C11O13) (10), δ_{op}(N4H14) (13), τ(C11O13) (29)
557			δ_{op}(N4H14) (22), τ(C11O13) (16), τ(C2N4) (28)
551			δ_{op}(N24H34) (25), δ_{op}(C23O26) (27), τ(C25N24) (10), τ(C23N24) (15)
532			*ν*(C22C25) (13), δ_{ip}(C2O3) (16), δ(C22C23N24) (12), δ_{ip}(C23O26) (21)
498			δ(N4C5C6) (20), δ(C5C6C7) (14), ω(O12C11O13) (11), τ(C11O13) (15)
472		443	δ(N4C5C11) (11), δ(O12C11O13) (15), p(O12C11O13) (18)
457			$p(CH_2'')$ (11), p(O29C28O30) (47)
437			$p(CH_2')$ (15), p(O9C8O10) (79)

ν, stretching; as, anti-symmetric; s, symmetric, δ, deformations; ω, wagging; τ, twisting; p, rocking; ip,in-plane; op, out-of-plane;

Table 4.3 Observed and calculated wavenumber assignments (cm^{-1}) for β-NAAG

Calculated (cm^{-1})	Experimental Raman	ROA	Assignments (%PEDs)
		1735	ν(C23O24) (84)
1723			ν(C5O16) (72), ν(C13O14) (12)
1717		1650	ν(C5O16) (12), ν(C13O14) (70)
1643		1636	ν(C28O29) (63), δ_{ip}(N1H2) (10)
1641		1616	ν(C32O34) (53), δ_{ip}(N19H29) (21)
1566			ν(C32O34) (17), ν(C32N19) (25), δ_{ip}(N19H29) (41)
1548		1540	ν(C28N1) (31), δ_{ip}(N1H2) (48)
1463			$\delta(CH_2'')$ (45), $\delta_{as}(CH_3)$ (26)
1462			$\delta(CH_2'')$ (41), $\delta_{as}(CH_3)$ (29)
1452		1450	$\delta(CH_2)$ (95)
1440			$\delta(CH_2')$ (95)
1438			$\delta_{as}(CH_3)$ (88)
1401			ν(C23O20) (10), ν(C21C23) (10), δ(C21H22) (17)
1385	1383	1388	δ(C3H4) (10), $\omega(CH_2)$ (27), $t(CH_2)$ (13), $\omega(CH_2')$ (10)
1378			$\delta_s(CH_3)$ (56), $\delta_{as}(CH_3)$ (25)
1373			δ(C3H4) (14), δ(C21H22) (15), $\omega(CH_2'')$ (10), $\delta_s(CH_3)$ (10)
1367			ν(C5O6) (15), δ(C5O6H17) (20), $\omega(CH_2)$ (16)
1362			δ(C5O6H17) (11), $\omega(CH_2)$ (14), $\omega(CH_2')$ (21)
1352			δ(C3H4) (42), $\omega(CH_2'')$ (12)
1331		1350	δ(C13O15H18) (20), $\omega(CH_2')$ (23)
1323		1315	δ(C23O30H31) (25), δ(C21H22) (20)
1304			ν(C32N19) (25), δ_{ip}(N19H29) (22)
1298			δ(C3H4) (47), $\tau(CH_2)$ (14)
1285		1255	ν(C28N1) (14), δ_{ip}(N1H2) (14), δ(C21H22) (25)
1256			δ(C21H22) (30), $\omega(CH_2'')$ (29)
1250			$\tau(CH_2)$ (16), $w(CH_2')$ (15), $\tau(CH_2')$ (21)
1223		1223	δ(C13O15H18) (10), $\tau(CH_2')$ (13), $\tau(CH_2'')$ (21)
1217		1198	δ(C13O15H18) (14), $\tau(CH_2'')$ (10)
1180			ν(C5O6) (33), δ(C5O6H17) (41)
1166			ν(C13O15) (10), $\tau(CH_2')$ (10)
1157			ν(C23O30) (20), δ(C23O30H31) (16), $\tau(CH_2'')$ (16)
1134			ν(C3N1) (38), ν(C3C7) (13)
1129		1130	ν(C23O30) (22), ν(C21N19) (35), δ(C23O30H31) (11)
1061	1078	1091	ν(C13O15) (11), ν(C7C10) (10), ν(C3C7) (10), $p(CH_2)$ (10), $p(CH_2')$ (10)
1052		1055	ν(C7C10) (39), ν(C3C7) (15)
1045			$p_{op}(CH_3)$ (65), δ_{op}(C32O34) (19)
1037			ν(C21C25) (17), $p_{ip}(CH_3)$ (37)
1027			$p(CH_2'')$ (11), $p_{ip}(CH_3)$ (24)
1002		966	ν(C32N19) (11), δ(C21N19C32) (11)
974	990	966	ν(C25C28) (13), ν(C32C33) (19), $p(CH_2'')$ (24)

Table 4. 3 Continued

954		966	*v*(C21C25) (10), p(CH_2') (16), ω(O14C13O15) (12)
919		935	*v*(C5O6) (15), *v*(C3C7) (12), *v*(C3C5) (15)
906	880	880	*v*(C13O15) (10), p(CH_2) (15)
868			p(CH_2') (14)
813			*v*(C13O15) (11), *v*(C10C13) (20), p(CH_2) (11), δ_{op}(C28O29) (11)
800			*v*(C10C13) (11), *v*(C21C23) (13), p(CH_2) (19), δ_{op}(C28O29) (12)
735			*v*(C3C7) (11), ω(O6C15O16) (41)
718			δ_{op}(C28O29) (13), ω(O24C23O30) (39)
681			p(CH_2') (14), ωO14C13O15) (28), τ(C13O15) (12)
665			δ(N1C3C5) (14)
648			p_{op}(CH_3) (11), δ_{op}(C32O34) (45), τ(N19C32) (20)
634			δ(O6C5O16) (29)
631			*v*(C32C33) (14), δ(O24C23O30) (26), δ_{op}(C32O34) (13)
601			*v*(C5O6) (11)
597		597	δ(O14C13O15) (13), τ(C13O15) (66)
584			*v*(C5O6) (54)
572			τ(C23O30) (41), τ(N1C28) (11)
551			*v*(C10C13) (12), δ(O14C13O15) (41)
543			δ_{op}(N19H20) (14), δ_{op}(C32O34) (15), τ(N19C32) (17)
531			τ(C23O30) (19)
520		537	δ_{ip}(C28O29) (11), δ_{op}(N1H2) (22), τ(N1C28) (15)
513			p(O24C23O30) (13), δ(N19C32C33) (16), δ_{ip}(C32O34) (14)
469			δ(C11C01C12) (11), p(O14C13O15) (43), ω(O14C13O15) (12)
446			δ(N19C21C25) (29), ω(O24C23O30) (11)

v, stretching; as, anti-symmetric; s, symmetric, δ, deformations; ω, wagging; τ, twisting; p, rocking; ip,in-plane; op, out-of-plane;

4.5.2 Spectral Region 1800 – 1200 cm^{-1}

In this region of the spectra the bands observed are primarily due to the amide I, II and III vibrations which are sensitive to secondary structure. *α*- and *β*-*N*-acetyl-L-Asp-L-Glu have two amide moieties and L-Asp-L-Glu has one amide moieties. However, this does not infer that the amide vibrational mode will be observed as a result of the extra amide moiety as the observance of vibrational moieties is dependent upon the stability of the vibrational mode and is determined by hydrogen bonding and interaction with side chains. It is generally the amide I and III bands that are used in Raman spectroscopy for peptide secondary-structure analysis and the determination of conformational preferences. The amide I vibrations, which occur between 1730 and 1650 cm^{-1}, arise mainly due to the C=O stretch (Bandekar 1992; Krimm & Bandekar 1986; Watson & Hirst 2002), the amide II vibrations, occur between 1570 and 1515 cm^{-1}, and arise mainly due to the C-N stretch and the N-H deformation (Chen *et al*. 1995; Mirkin & Krimm 1991; Miyazawa *et al*. 1956) and the amide III vibrations, which occur between 1340 and 1200 cm^{-1}, originate mainly from the N-H in-plane bending and the C-N stretching modes. CH wagging, twisting and deformations also contribute to bands appearing in this region (Chen *et al*. 1995; Mirkin & Krimm 1991; Miyazawa *et al*. 1956).

For the Raman spectrum of L-Asp-L-Glu shown in Fig. 2 there is a large broad amide I band between 1750 and 1600 cm^{-1} and in the ROA spectrum there is a corresponding broad band between 1750 and 1650 cm^{-1}. It is a known consideration for ROA that bands between 1500 and 1800 cm^{-1} can become distorted due to the influence of the O-H bending vibrations of water molecules and caution should be

exercised in assigning this ROA feature to a definitive amide I vibration (Mukhopadhyay *et al*. 2008) from the dipeptide. The spectral features observed for the two deuterated samples are very similar. The amide I signal observed in the Raman spectra for *α*- and *β*-*N*-acetyl-L-Asp-L-Glu in H_2O are very similar to the results shown for L-Asp-L-Glu, however, for the deuterated samples of *α*- and *β*-*N*-acetyl-L-Asp-L-Glu there is a band at 1653 and 1643 cm^{-1}, respectively, which we have assigned to this spectral feature. The amide I vibration for *α*-*N*-acetyl-L-Asp-L-Glu is weak in the ROA spectrum for H_2O, however, for the deuterated sample there is a broad band observed at 1680 cm^{-1} which we have tentatively assigned to the amide I mode. The amide I ROA bands for these three dipeptides measured in H_2O are not sufficiently different to make them sensitive markers of acetylation; however the spectra measured in D_2O are more promising. For *β*-*N*-acetyl-L-Asp-L-Glu in H_2O the amide I vibration is observed at 1652 cm^{-1}, whereas in D_2O this band is observed at 1620 cm^{-1}, which indicates that there is significant interaction of the carbonyl groups with solvent water molecules. The clear observation of the amide I band for *β*-*N*-acetyl-L-Asp-L-Glu in D_2O appears to be a ROA marker band for distinction between the two DAAPs. The observation that the amide I band for *β*-*N*-acetyl-L-Asp-L-Glu in D_2O has the opposite sign to that measured for the non- and *α*-acetylated forms also when in D_2O is most likely explained by a different environment and greater conformational mobility of the C=O groups in the *β*-acetylated peptide. This also suggests that for *α*-*N*-acetyl-L-Asp-L-Glu there is more restricted conformational mobility of the carbonyl groups, possibly due to stronger intramolecular interactions with the side chain which leads to a different and sharper ROA profile.

The amide II vibration is not usually observed or is weak in ROA spectra of polypeptides and proteins (Barron *et al* 2000). In agreement with this general trend the Raman spectra for all three DAAPs do not display obvious or strong amide II vibrations. However, there is a weak ROA band observed at 1535 cm^{-1} for *α-N*-acetyl-L-Asp-L-Glu which we assign to the amide II vibration. *β-N*-acetyl-L-Asp-L-Glu displays a similar ROA band at 1540 cm^{-1}. DFT calculations predict this band at 1538 and 1566 cm^{-1}, respectively. Though these features are relatively weak, it is therefore possible to observe amide II ROA marker bands for small peptides. Since the amide II band is not observed for L-Asp-L-Glu, acetylation appears to be the reason this vibration is observed in the ROA spectra of *α*- and *β-N*-acetyl-L-Asp-L-Glu. This indicates that this ROA amide II band in each case is originating specifically from the NH group that is the site of acetylation, i.e. N24 for of *α*- and N19 for *β-N*-acetyl-L-Asp-L-Glu. Although this feature is relatively weak for these two acetylated peptides it may potentially be used to monitor the conformational dynamics of the acetyl group with respect to the rest of the peptide.

Complex bands result from coordinate displacements from the interaction of side chains and hydrogen bonding for the amide III motions. The Raman spectrum for L-Asp-L-Glu displays a broad feature between 1220-1300 cm^{-1} which arises from a number of contributing bands within this spectral feature, including NH deformations, CH twisting, wagging and deformations. The ROA spectrum of L-Asp-L-Glu shows a corresponding negative band at ~1281 cm^{-1}. For *α*- and *β-N*-acetyl-L-Asp-L-Glu there is a weak broad Raman band around the same position as for L-Asp-L-Glu, with a corresponding ROA band with negative sign at 1255 cm^{-1}, a downshift of 30 cm^{-1}. The deuterated samples for the three DAAPs display

qualitatively similar profiles in the amide III region, though the relative intensity of the amide III bands is weaker for the α-acetylated form than for either the non- or β-acetylated forms. As this relative difference in amide III ROA signal intensities was not observed for the same samples measured here in H_2O we conclude that this observation is due to a higher level of H/D exchange occurring for the α-acetylated DAAP.

Bands originating from CH_3 symmetric and anti-symmetric deformations and COH deformations are also observed in this spectral region. For the Raman spectra of α- and β-*N*-acetyl-L-Asp-L-Glu there is an observed band at ~1380 cm^{-1} which we assign to CH_3 deformations. In the ROA spectrum there are corresponding bands at 1380 and 1388 cm^{-1}, respectively. The ROA band observed for β-*N*-acetyl-L-Asp-L-Glu is broader which suggests interaction with the acetyl group and the Raman and ROA spectra of L-Asp-L-Glu do not display corresponding bands. COH deformations are observed clearly for α-*N*-acetyl-L-Asp-L-Glu at 1421 cm^{-1} with a band at 1455 cm^{-1} being assigned to CH_3, CH_2 anti-symmetric and symmetric deformations. For β-*N*-acetyl-L-Asp-L-Glu again this latter band is broad which suggests more hydrogen bonding interaction between CH_3, CH_2 and COH groups.

In the spectral region there are also bands observed for symmetric stretching, deformations, and wagging motions of CH_2 groups. For L-Asp-L-Glu a ROA band is observed at 1418 cm^{-1} which agrees closely with the DFT calculations which predict this band at 1413 cm^{-1}. The corresponding band for α-*N*-acetyl-L-Asp-L-Glu is observed at 1356 cm^{-1} and for β-*N*-acetyl-L-Asp-L-Glu is observed at 1388 cm^{-1}, and again the position of these bands agrees favourably with the DFT calculations which

predict a negative band at 1352 cm^{-1} for *α-N*-acetyl-L-Asp-L-Glu and a positive band at 1385 cm^{-1} for *β-N*-acetyl-L-Asp-L-Glu. The observation of opposite signs for these ROA bands indicates that the CH_2 motions with respect to the nearest chiral centre have opposite stereochemistries in the two acetylated DAAPS, and that this ROA marker band is distinctive for these constitutional isomers. Similar CH_3 deformation ROA bands for *α*- and *β-N*-acetyl-L-Asp-L-Glu are observed at 1434 cm^{-1} and 1450 cm^{-1}, respectively. These experimental bands agree favourably with the calculations where they are predicted at 1434 cm^{-1} and 1462 cm^{-1}, respectively.

In order to define which vibrations are influenced by deuteration of labile hydrogens, Raman and ROA spectra were measured for samples of these acetylated and non-acetylated dipeptides dissolved in D_2O. The computed bands for the amide I, II and III vibrations generally agree with the experimental results for the Raman and ROA spectra observed in H_2O although the deuterated samples do display a lower intensity than the samples analysed using H_2O and a general downward shift of ~5-10 cm^{-1}, however, this is not linear. This intensity difference can be explained by the fact that the ROA and Raman cross sections are influenced by water. However, water makes little to no contribution to the Raman and ROA spectral intensities from 800 cm^{-1} to 1600 cm^{-1}, thus, the experimental and simulated bands over this range agree more favourably (Kapitán *et al* 2006).

4.5.3 Spectral Region 1200 – 700 cm^{-1}

Band assignments in this region of the spectrum are usually tentative owing to the wide range of skeletal vibrations spanning this region. Without DFT calculations

assignments of bands in this region are usually difficult to make. The ROA spectra for all three DAAPs display a clear positive band at ~1130 cm^{-1} which we have assigned to an overlap of bands corresponding to CO and CN stretching and deformations of COH groups as indicated by the predicted bands at 1137, 1131 and 1129 cm^{-1} for L-Asp-L-Glu, *α*-*N*-acetyl-L-Asp-L-Glu and *β*-*N*-acetyl-L-Asp-L-Glu, respectively. The intensities of these ROA bands is notable considering that in the Raman spectra for all three DAAPs there are only very weak corresponding bands.

In the Raman spectrum for L-Asp-L-Glu in H_2O there is a band observed at 1081 cm^{-1} which corresponds to the band predicted at 1068 cm^{-1}, and is assigned to C-C stretching. This Raman band gives rise to a corresponding positive ROA band at 1086 cm^{-1} and 1082 cm^{-1} in H_2O and D_2O, respectively. The corresponding Raman bands for *α*- and *β*-*N*-acetyl-L-Asp-L-Glu in D_2O are observed with a slight downshift at 1078 cm^{-1} and at a lower intensity. The DFT calculations predict this band at 1061 cm^{-1}, in the ROA spectra for *α*- and *β*-*N*-acetyl-L-Asp-L-Glu the band is observed at 1096 cm^{-1} and 1091 cm^{-1}, respectively. In the Raman and ROA spectra rocking out of plane NH_3 motions, CC stretching and rocking CH_2 motions are observed at ~930 cm^{-1}, originating from the amino terminus of L-Asp-L-Glu, the observed band agrees with the calculated spectra for both Raman and ROA, in which it is predicted at 941 cm^{-1}. Bands observed in the Raman and ROA spectrum for *α*- and *β*-*N*-acetyl-L-Asp-L-Glu around 950 cm^{-1} and 880 cm^{-1}, respectively, are assigned as contributions from rocking CH_2 motions, CC stretching, CNC deformations and OCO wagging motions.

4.5.4 Bands below 700 cm^{-1}

Assignments for bands observed in this region are due to NCC deformations, CO out-of-plane deformations, CC stretching, OCO deformations, O-C-O rocking motions, CO twisting motions, NH out-of-plane deformations, NC twisting motions, CCC deformations, CO in-plane deformations, and OCO wagging motions. The majority of the ROA bands in this region are weak which is typical for peptides and proteins in aqueous solutions as lower wavenumber bands are influenced by the solvent (Kapitán *et al* 2006). The ROA bands for *α*- and *β-N*-acetyl-L-Asp-L-Glu are clearer and more defined than the Raman bands, however, the ROA bands for L-Asp-L-Glu are weaker and broader than those of the acetylated samples. The bands observed for the deuterated samples of *α*- and *β-N*-acetyl-L-Asp-L-Glu are more defined compared to samples dissolved in H_2O. This follows the general trend that vibrations below 800 cm^{-1} interact with the O-H bending of water molecules. The observed spectra for both *α*- and *β-N*-acetyl-L-Asp-L-Glu are in reasonable agreement with the calculated spectra.

4.6 Conclusion

ROA spectra are known to provide more detailed information about peptide conformations than Raman spectra do (Barron *et al* 2000), and we demonstrate here that this is also the case for acetylated peptides. The Raman spectra for *α*- and *β-N*-acetyl-L-Asp-L-Glu cannot be distinguished whereas the ROA spectra have clear marker bands of acetylation site for each dipeptide. *β-N*-acetyl-L-Asp-L-Glu has a richer band structure and provides more information at higher wavenumbers, the

spectral regions most commonly used for conformational analysis, there is a clear amide I band at 1652 cm^{-1} which isn't observed for *α-N*-acetyl-L-Asp-L-Glu. The band assigned to amide II type vibrations for the two acetylated DAAPs is quite weak in H_2O, however, for *β-N*-acetyl-L-Asp-L-Glu analysed in D_2O it is clear and strong whereas for *α-N*-acetyl-L-Asp-L-Glu the band is not observed. This spectral feature is reliant on hydrogen bonding and whilst the H_2O samples display a weak band, due to interaction with the water molecules, the difference in the conformation for *β-N*-acetyl-L-Asp-L-Glu suggests more hydrogen bonding occurs regardless of the solvent. We have also shown in this work that ROA provides information on acetylation. The CH_3 symmetric and anti-symmetric deformations bands are clearly observed in the ROA spectrum of *α*- and *β-N*-acetyl-L-Asp-L-Glu however, for *β-N*-acetyl-L-Asp-L-Glu this band is broader which suggests some overlap from the carboxyl bands, whereas the carboxyl band for *α-N*-acetyl-L-Asp-L-Glu is observed at 1421 cm^{-1}. We have ascertained that ROA provides more structural information about modifications of peptides than conventional Raman spectra do, in this case for site-specific acetylation and marker bands for constitutional isomers. In addition by comparing the experimental results to DFT calculations we can assign most of the spectral features.

Acknowledgements

The authors thank the Royal Society of Chemistry and the Engineering and Physical Sciences Research Council for funding CL's Analytical Trust Fund PhD Studentship..

4.7 References

Althaus M.A. & Pichler W.J. (1994) *Allergy* **49**: 184-188

Anuradah M.V. & Ravindranath B. (1997) *Tetrahedran* **53**: 1123-1130

Atkins P.W. & Barron L.D. (1969) *Molecular Physics* **16**: 452-456

Bandekar J. (1992) *J. Biochim. Biophys. Acta* **1120**: 123-143

Barron L.D., Bogaard M.P. & Buckingham A.D. (1972) *J. Am. Chem. Soc.* **95**: 603-605

Barron L.D., Hecht L., Blanch E.W. & Bell A.F. (2000) *Prog. Biophys. Mol. Bio.* **73**: 1-49

Becke A.D. (1993) *J. Chem. Phys.* **98**: 1372-1377

Briggs S.D., Xiao T., Sun Z.W. & Caldwell J.A. (2002) *Nature* **418**: 498

Brown R. & Strathdee G. (2002) *Trends Mol. Med.* **8**: S43–S48

Cances M.T., Mennucci V. & J. Tomasi (1997) *J. Chem. Phys.* **107**: 3032-3040

Chasse G.A. , Mak M.L., Deretey E., Farkas I., Torday L.L., & Papp J.G. (2001) *J. Mol. Struc. (Theochem)* **571**: 27–37

Chen X.-G., Schweitzer-Stenner R., Asher S.A., Mirkin N. & Krimm S. (1995) *J. Phys. Chem.* **99**: 3074-3083

Chin W., Piuzzi F., Dimicoli I. & Mons M. (2006) *Phys. Chem. Chem. Phys.* **8**: 1033-1048

Choi J.H., Hahn S. & Cho M. (2005) *Int. J. Quantum Chem.* **104**: 616-634

Curatolo A., D'Arcangelo P., Lino,A. & Brancati A. (1965) *J. Neurochemistry* **12**: 339-342

Dunning T.H. (1989) *J. Chem. Phys.* **90**: 1007-1023

Edler J. & Hamm P. (2002) *J. Chem. Phys.* **117**: 2415-2424

Elstner M., Frauenheim T. & Suhai S. (2003) *J. Mol. Struc. (Theochem)* **632:** 29–41

Fang C., Wang J., Charnley A.K., Barber-Armstrong, W., Smith A.B., Decatur S.M. & Hochstrasser R.M. (2003) *Chem. Phys. Lett.* **2003**: 586-592

Farkas A.D., Foresman O., Ortiz J.B., Cioslowski J.V. & Fox D.J., Gaussian, Inc., Wallingford CT, 2009

Gaussian 09, Revision ***A.1***, Frisch M.J., Trucks G.W., Schlegel H.B.,. Scuseria G.E, Robb M.A., Cheeseman J.R., Scalmani G., Barone V., Mennucci B., Petersson G.A., Nakatsuji H., Caricato M., Li X., Hratchian H.P.,. Izmaylov A.F, Bloino J.,. Zheng G, Sonnenberg J.L., Hada M., Ehara M., Toyota K., Fukuda R., Hasegawa J., Ishida M., Nakajima T., Honda Y., Kitao O., Nakai H., Vreven T.,. Montgomery Jr. J.A, Peralta J.E., Ogliaro F., Bearpark M., Heyd J.J., Brothers E., Kudin K.N., Staroverov V.N., Kobayashi R., Normand J., Raghavachari K., Rendell A., Burant J.C., Iyengar S.S., Tomasi J., Cossi M., Rega N., Millam J.M., Klene M.,. Knox J.E, Cross J.B., Bakken V., Adamo C., Jaramillo J., Gomperts R., Stratmann R.E., Yazyev O., Austin A.J., Cammi R., Pomelli C., Ochterski J.W., Martin R.L., Morokuma K., Zakrzewski V.G., Voth G.A., Salvador P., Dannenberg J.J., Dapprich S., Daniels
Goldschmidt P.L., Normand Vullez-Le B. Briquet I. & Dray F. (2007 *Allergy* **45**: 363

Grdadolnik J., Grdadolnik S.G. & Avbelj F. (2008) *J. Phys. Chem.* **112**: 2712-2718

Grdadolnik J., Mohacek-Grosev V., Baldwin R.L. & Avbelj F. (2011) *PNAS* **108**(5): 1794-1798

Hug W. & Hangartner G. (1999) *J. Raman Spectrosc.* **30**: 841-852

Johnstone R.W. (2002) *Nat. Rev. Drug Discov.* **1**: 287–299

Kandel E.R., Schwartz J.H. & Jessell T.M. (2000) *Principles of Neural Science (4th Ed)* McGraw-Hill Co. New York pp. 281

Kapitán J., Baumruk V. & Bouř P. (2006) *J. Am. Chem. Soc.* **128** (7): 2438–2443

Kausar K., Alexander B.D., Dines T.J., Withnall R. & Chowdhry B.Z. (2009A) *J. Raman Spectrosc.* **40**(6): 670-678

Kausar N., Alexander B D., Dines T.J., Withnall R. & Chowdhry B.Z. (2009B) *J. Raman Spectrosc.* **40**: 661-669

Kausar N., Dines T.J., Chowdhry B.Z. & Alexander B.D. (2009C) *Phys. Chem. Chem. Phys.* **11**: 6389-6400

Kinalwa M.N., Blanch E.W., & Doig A. J. (2010) *Anal. Chem.* **82***:* 6347–6349

Kowalski M.M., Cassidy M., Namboodiri M.A.A. & Neale J.H. (1987) *Brain Res.* **406**: 397-401

Lee C., Yang W. & Parr R.G. (1988) *Phys. Rev. B,* **37**: 785-789

Mirkin N.G. & Krimm S. (1991) *J. Am. Chem.* **113**: 9742-9747

Mirkin N.G. & Krimm S. (2004) *J. Phys. Chem.* **108**: 5438-5448

Miyamoto E., Kakimoto Y. & Sano I. (1966) *J. Neuro. Chem.* **13**: 999-1003

Miyazawa T., Shimanouchi T. & Mizushima S.I. (1956) *J. Chem. Phys.***24**: 408

Mizzen C.A. & Allis C.D. (2000) *Science.***289**: 2290–2291

Moffett J.R., Palkovits M., Namboodiri M.A.A. & Neale J.H. (1994) *J. Comp. Neurol.* **347**: 598-618

Mukhopadhyay P., Zuber G. & David N. Beratan (2008) *Biophysical Journal* **95**(12): 5574–5586

Neale J.H., Bzdega T. & Wroblewska B. (2000) *J. Neurochemistry.* **75**: 443-452

Pozo Ramajo A., Petty S.A. & Volk M. (2006) *Chem. Phys.* **323**: 11-20

Reihe M., Liégeois V. & Ruud K. (2005) *J. Phys. Chem. A.* **109**(33): 7567–7574

Ruud K., Helgaker T. & Bouř P. (2002) *J. Phys. Chem. A.* **106**(32): 7448–7455

Schachtschneider J. A. (1964 & 1965) *Vibrational Analysis of Polyatomic Molecules*, Parts V and VI, Technical Report Nos. 231 and 57, Shell Development Co. Houston TX

Stauch B.L., Robinson M.B., Forloni G., Tsai G. & Coyle J.T. (1989) *Neurosci. Lett.* **100**: 295-300

Tieman S.B., Cangro C.B. & Neale J.H. (1987) *Brain Res.* **420**: 188-193

Tiffany M.L. & Krimm S. (1969) *Biopolymers* **8**: 347-59

Tsai G., Coyle J.T., Kleinman J.E., Baer L., Carter R., Slusher B.S. & Passani L.A. (1996) *Arch. Gen. Psychiatry* **52**: 829-836

Watson T.M. & Hirst J.D. (2002) *J. Phys. Chem.* **106**: 7858-7867

Wilson E.B., Decius J.C. & Cross P.C., *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*, McGraw-Hill, New York, 1955.

York D.M. & Karplus M. (1999) *J. Phys. Chem. A* **103:** 11060-11079

4.8 Supplementary Information

Table S4.1 L-Asp-L-Glu calculated geometry

Calculated bond lengths (Å)

r(N1H29)	1.0253
r(N1C3)	1.4973
r(H2N1)	1.0294
r(C3C7)	1.5342
r(C3C5)	1.5365
r(H4C3)	1.0930
r(C5O6)	1.2442
r(C7H8)	1.0995
r(C7H9)	1.0943
r(C10C7)	1.5110
r(C10O11)	1.3382
r(O12C10)	1.2240
r(N13C5)	1.3342
r(N13C15)	1.4629
r(H14N13)	1.0149
r(C15C17)	1.5479
r(C15H16)	1.0960
r(C17O28)	1.2628
r(O18C17)	1.2651
r(C19C22)	1.5328
r(C19C15)	1.5476
r(H20C19)	1.0967
r(H21C19)	1.0976
r(C22H23)	1.0962
r(H24C22)	1.1012
r(C25C22)	1.5086
r(C25O27)	1.3491
r(O26C25)	1.2213
r(H30O11)	0.9740
r(H31O27)	0.9732
r(H32N1)	1.0219

Calculated bond angles (°)

θ(H29N1C3)	110.00
θ(H29N1H2)	109.88
θ(H29N1H32)	107.72
θ(C3N1H2)	108.72
θ(N1C3C7)	110.59
θ(N1C3C5)	107.62
θ(N1C3H4)	106.90
θ(C3N1H32)	111.69
θ(H2N1H32)	108.82
θ(C7C3C5)	110.16
θ(C7C3H4)	110.54
θ(C3C7H8)	110.41
θ(C3C7H9)	108.87
θ(C3C7C10)	114.23
θ(C5C3H4)	110.94
θ(C3C5O6)	119.57
θ(C3C5N13)	115.48
θ(O6C5N13)	124.91
θ(H8C7H9)	107.06
θ(H8C7C10)	107.17
θ(H9C7C10)	108.86
θ(C7C10O11)	111.97
θ(C7C10O12)	125.37
θ(O11C10O12)	122.64
θ(C10O11H30)	108.99
θ(C5N13C15)	124.53
θ(C5N13H14)	120.13
θ(C15N13H14)	115.04
θ(N13C15C17)	108.34
θ(N13C15H16)	108.04
θ(N13C15C19)	110.15
θ(C17C15H16)	108.37
θ(C15C17O28)	116.58
θ(C15C17O18)	117.47
θ(C17C15C19)	112.89
θ(H16C15C19)	108.92
θ(O28C17O18)	125.94
θ(C22C19C15)	115.07
θ(C22C19H20)	107.93
θ(C22C19H21)	110.22
θ(C19C22H23)	111.70
θ(C19C22H24)	108.95
θ(C19C22C25)	115.38
θ(C15C19H20)	106.89
θ(C15C19H21)	109.19
θ(H20C19H21)	107.19
θ(H23C22H24)	106.37
θ(H23C22C25)	109.09
θ(H24C22C25)	104.74
θ(C22C25O27)	112.25
θ(C22C25O26)	125.79
θ(O27C25O26)	121.88
θ(C25O27H31)	108.49

Calculated torsion angles (^{o})	
τ(H29N1C3C7)	-73.71
τ(H29N1C3C5)	46.66
τ(H29N1C3H4)	165.90
τ(C7C3N1H2)	46.64
τ(C5C3N1H2)	167.01
τ(H4C3N1H2)	-73.75
τ(H32N1C3C7)	166.73
τ(N1C3C7H8)	64.28
τ(N1C3C7H9)	-178.47
τ(N1C3C7C10)	-56.58
τ(H32N1C3C5)	-72.90
τ(N1C3C5O6)	-29.97
τ(N1C3C5N13)	152.13
τ(H32N1C3H4)	46.34
τ(H8C7C3C5)	-54.57
τ(H9C7C3C5)	62.69
τ(C10C7C3C5)	-175.42
τ(C7C3C5O6)	90.67
τ(C7C3C5N13)	-87.23
τ(H8C7C3H4)	-177.54
τ(H9C7C3H4)	-60.29
τ(C10C7C3H4)	61.61
τ(C3C7C10O11)	-159.89
τ(C3C7C10O12)	21.66
τ(O6C5C3H4)	-146.59
τ(N13C5C3H4)	35.51
τ(C3C5N13C15)	172.30
τ(C3C5N13H14)	-1.02
τ(O6C5N13C15)	-5.48
τ(O6C5N13H14)	-178.79
τ(H8C7C10O11)	77.48
τ(H8C7C10O12)	-100.98
τ(H9C7C10O11)	-37.99
τ(H9C7C10O12)	143.56
τ(C7C10O11H30)	-176.68
τ(H30O11C10O12)	1.83
τ(C5N13C15C17)	-150.11
τ(C5N13C15H16)	-32.91
τ(C5N13C15C19)	85.96
τ(C17C15N13H14)	23.51
τ(H16C15N13H14)	140.71
τ(C19C15N13H14)	-100.42
τ(N13C15C17O28)	153.13
τ(N13C15C17O18)	-27.66
τ(N13C15C19C22)	-172.91
τ(N13C15C19H20)	-53.09
τ(N13C15C19H21)	62.55
τ(O28C17C15H16)	36.15
τ(O18C17C15H16)	-144.65
τ(C19C15C17O28)	-84.59
τ(C19C15C17O18)	94.61
τ(C17C15C19C22)	65.84
τ(C17C15C19H20)	-174.34
τ(C17C15C19H21)	-58.70
τ(H16C15C19C22)	-54.59
τ(H16C15C19H20)	65.23
τ(H16C15C19H21)	-179.13
τ(H23C22C19C15)	40.90
τ(H24C22C19C15)	158.12
τ(C25C22C19C15)	-84.45
τ(H23C22C19H20)	-78.34
τ(H24C22C19H20)	38.88
τ(C25C22C19H20)	156.31
τ(H23C22C19H21)	164.90
τ(H24C22C19H21)	-77.88
τ(C25C22C19H21)	39.55
τ(C19C22C25O27)	153.76
τ(C19C22C25O26)	-29.46
τ(H23C22C25O27)	27.08
τ(H23C22C25O26)	-156.14
τ(H24C22C25O27)	-86.47
τ(H24C22C25O26)	90.31
τ(C22C25O27H31)	176.43
τ(H31O27C25O26)	-0.49

Table S4.2 α-N-acetyl-L-Asp-L-Glu calculated geometry

Calculated bond lengths (Å)	
r(H1C25)	1.0948
r(C2N4)	1.3443
r(C2C25)	1.5459
r(O3C2)	1.2417
r(N4H14)	1.0118
r(C5N4)	1.4522
r(C5C11)	1.5265
r(C5C6)	1.5506
r(C6C7)	1.5494
r(C7C8)	1.5062
r(C7H19)	1.0964
r(C7H18)	1.0952
r(C8O9)	1.3373
r(O9H20)	0.9818
r(O10C8)	1.2295
r(C11O12)	1.2179
r(O13H21)	0.9734
r(O13C11)	1.3419
r(H15C5)	1.0964
r(H16C6)	1.0955
r(H17C6)	1.0965
r(C22H32)	1.0950
r(C22H31)	1.0978
r(C23N24)	1.3546
r(C23C22)	1.5101
r(N24H34)	1.0129
r(C25N24)	1.4541
r(C25C27)	1.5521
r(O26C23)	1.2443
r(C27H36)	1.0949
r(C27C28)	1.5088
r(C28O29)	1.2298
r(O30H37)	0.9861
r(O30C28)	1.3322
r(H33C22)	1.0971
r(H35C27)	1.0926
Calculated bond angles (°)	
θ(H1C25C2)	105.04
θ(H1C25N24)	108.08
θ(H1C25C27)	107.14
θ(N4C2C25)	116.52
θ(N4C2O3)	124.09
θ(C2N4H14)	117.56
θ(C2N4C5)	124.36
θ(C25C2O3)	119.38
θ(C2C25N24)	114.34
θ(C2C25C27)	111.33
θ(H14N4C5)	117.91
θ(N4C5C11)	109.19
θ(N4C5C6)	114.90
θ(N4C5H15)	106.94
θ(C11C5C6)	112.17
θ(C5C11O12)	125.24
θ(C5C11O13)	111.67
θ(C11C5H15)	106.67
θ(C5C6C7)	117.57
θ(C6C5H15)	106.49
θ(C5C6H16)	107.04
θ(C5C6H17)	107.46
θ(C6C7C8)	112.07
θ(C6C7H19)	107.51
θ(C6C7H18)	111.82
θ(C7C6H16)	107.83
θ(C7C6H17)	109.55
θ(C8C7H19)	106.70
θ(C8C7H18)	110.00
θ(C7C8O9)	114.28
θ(C7C8O10)	123.53
θ(H19C7H18)	108.51
θ(C8O9H20)	108.95
θ(O9C8O10)	122.17
θ(O12C11O13)	123.09
θ(H21O13C11)	108.78
θ(H16C6H17)	106.89
θ(H32C22H31)	109.26
θ(H32C22C23)	112.98
θ(H32C22H33)	109.33
θ(H31C22C23)	108.63
θ(H31C22H33)	107.84
θ(N24C23C22)	116.46
θ(C23N24H34)	118.78
θ(C23N24C25)	123.22
θ(N24C23O26)	122.01
θ(C22C23O26)	121.53
θ(C23C22H33)	108.66
θ(H34N24C25)	117.85
θ(N24C25C27)	110.47
θ(C25C27H36)	108.05
θ(C25C27C28)	111.06
θ(C25C27H35)	109.01
θ(H36C27C28)	108.97
θ(H36C27H35)	109.44
θ(C27C28O29)	123.55
θ(C27C28O30)	113.45
θ(C28C27H35)	110.26
θ(O29C28O30)	122.91
θ(H37O30C28)	109.43

Calculated torsion angles (°)	
τ(H1C25C2N4)	-138.22
τ(H1C25C2O3)	42.69
τ(H1C25N24C23)	30.08
τ(H1C25N24H34)	-145.46
τ(H1C25C27H36)	62.00
τ(H1C25C27C28)	-178.53
τ(H1C25C27H35)	-56.86
τ(H14N4C2C25)	1.97
τ(C5N4C2C25)	176.95
τ(N4C2C25N24)	-19.91
τ(N4C2C25C27)	106.16
τ(H14N4C2O3)	-178.99
τ(C5N4C2O3)	-4.01
τ(C2N4C5C11)	-128.53
τ(C2N4C5C6)	104.45
τ(C2N4C5H15)	-13.49
τ(N24C25C2O3)	161.00
τ(C27C25C2O3)	-72.93
τ(C2C25N24C23)	-86.49
τ(C2C25N24H34)	97.98
τ(C2C25C27H36)	176.32
τ(C2C25C27C28)	-64.21
τ(C2C25C27H35)	57.47
τ(H14N4C5C11)	46.43
τ(H14N4C5C6)	-80.59
τ(H14N4C5H15)	161.48
τ(N4C5C11O12)	-9.16
τ(N4C5C11O13)	170.96
τ(N4C5C6C7)	61.38
τ(N4C5C6H16)	-177.21
τ(N4C5C6H17)	-62.68
τ(O12C11C5C6)	119.39
τ(O13C11C5C6)	-60.49
τ(C11C5C6C7)	-64.10
τ(C11C5C6H16)	57.31
τ(C11C5C6H17)	171.84
τ(H15C5C11O12)	-124.38
τ(H15C5C11O13)	55.74
τ(C5C11O13H21)	179.34
τ(H15C5C6C7)	179.56
τ(C5C6C7C8)	-80.18
τ(C5C6C7H19)	162.88
τ(C5C6C7H18)	43.87
τ(H16C6C5H15)	-59.02
τ(H17C6C5H15)	55.50
τ(H16C6C7C8)	158.81
τ(H17C6C7C8)	42.82
τ(C6C7C8O9)	105.94
τ(C6C7C8O10)	-72.40
τ(H16C6C7H19)	41.87
τ(H17C6C7H19)	-74.12
τ(H16C6C7H18)	-77.14
τ(H17C6C7H18)	166.87
τ(O9C8C7H19)	-136.63
τ(O10C8C7H19)	45.03
τ(O9C8C7H18)	-19.12
τ(O10C8C7H18)	162.55
τ(C7C8O9H20)	-167.49
τ(O10C8O9H20)	10.87
τ(O12C11O13H21)	-0.55
τ(H32C22C23N24)	3.35
τ(H32C22C23O26)	-176.90
τ(H31C22C23N24)	-118.06
τ(H31C22C23O26)	61.69
τ(H34N24C23C22)	-1.76
τ(C25N24C23C22)	-177.26
τ(N24C23C22H33)	124.86
τ(O26C23N24H34)	178.49
τ(O26C23N24C25)	2.99
τ(C23N24C25C27)	147.00
τ (H33C22C23O26)	-55.39
τ (H34N24C25C27)	-28.54
τ (N24C25C27H36)	-55.50
τ (N24C25C27C28)	63.97
τ (N24C25C27H35)	-174.35
τ (C25C27C28O29)	-73.61
τ (C25C27C28O30)	103.16
τ (H36C27C28O29)	45.31
τ (H36C27C28O30)	-137.92
τ (H35C27C28O29)	165.44
τ (H35C27C28O30)	-17.79
τ (C27C28O30H37)	-164.67
τ (O29C28O30H37)	12.12

Table S4.3 *β*-*N*-acetyl-L-Asp-L-Glu calculated geometry

Calculated bond lengths (Å)	
r(N1H2)	1.0112
r(C3N1)	1.4530
r(C3C7)	1.5436
r(H4C3)	1.0962
r(C5C3)	1.5313
r(C5O6)	1.3400
r(C7H8)	1.0975
r(C7C10)	1.5378
r(H9C7)	1.0955
r(C10C13)	1.5091
r(C10H11)	1.1006
r(H12C10)	1.0953
r(C13O15)	1.3480
r(O14C13)	1.2220
r(O15H18)	0.9732
r(O16C5)	1.2195
r(H17O6)	0.9736
r(N19C32)	1.3536
r(H20N19)	1.0113
r(C21C25)	1.5498
r(C21N19)	1.4482
r(C21H22)	1.0960
r(C23C21)	1.5271
r(O24C23)	1.2174
r(C25H27)	1.0949
r(H26C25)	1.0961
r(C28C25)	1.5202
r(C28N1)	1.3501
r(O29C28)	1.2439
r(O30C23)	1.3427
r(H31O30)	0.9732
r(C32O34)	1.2452
r(C32C33)	1.5100
r(C33H35)	1.0976
r(H36C33)	1.0972
r(H37C33)	1.0952
Calculated bond angles (°)	
θ(H2N1C3)	118.34
θ(H2N1C28)	118.50
θ(N1C3C7)	112.15
θ(N1C3H4)	108.02
θ(N1C3C5)	113.31
θ(C3N1C28)	123.08
θ(C7C3H4)	108.91
θ(C7C3C5)	109.04
θ(C3C7H8)	109.85
θ(C3C7C10)	114.38
θ(C3C7H9)	107.27
θ(H4C3C5)	105.09
θ(C3C5O6)	114.31
θ(C3C5O16)	122.68
θ(O6C5O16)	122.99
θ(C5O6H17)	108.44
θ(H8C7C10)	110.86
θ(H8C7H9)	106.96
θ(C10C7H9)	107.15
θ(C7C10C13)	117.25
θ(C7C10H11)	107.77
θ(C7C10H12)	110.79
θ(C13C10H11)	105.68
θ(C13C10H12)	107.36
θ(C10C13O15)	113.40
θ(C10C13O14)	124.72
θ(H11C10H12)	107.50
θ(O15C13O14)	121.83
θ(C13O15H18)	108.32
θ(C32N19H20)	118.10
θ(C32N19C21)	123.54
θ(N19C32O34)	122.46
θ(N19C32C33)	116.24
θ(H20N19C21)	118.35
θ(C25C21N19)	112.61
θ(C25C21H22)	106.79
θ(C25C21C23)	113.46
θ(C21C25H27)	107.54
θ(C21C25H26)	108.80
θ(C21C25C28)	113.40
θ(N19C21H22)	106.65
θ(N19C21C23)	110.66
θ(H22C21C23)	106.13
θ(C21C23O24)	125.27
θ(C21C23O30)	111.59
θ(O24C23O30)	123.12
θ(H27C25H26)	108.13
θ(H27C25C28)	111.31
θ(H26C25C28)	107.53
θ(C25C28N1)	116.26
θ(C25C28O29)	121.43
θ(N1C28O29)	122.32
θ(C23O30H31)	108.71
θ(O34C32C33)	121.30
θ(C32C33H35)	108.57
θ(C32C33H36)	108.65
θ(C32C33H37)	112.93
θ(H35C33H36)	107.85
θ(H35C33H37)	109.33
θ(H36C33H37)	109.39
Calculated torsion angles (°)	
τ(H2N1C3C7)	-27.94

τ(H2N1C3H4)	-147.97
τ(H2N1C3C5)	96.04
τ(H2N1C28C25)	3.22
τ(H2N1C28O29)	-176.85
τ(C28N1C3C7)	155.32
τ(N1C3C7H8)	54.20
τ(N1C3C7C10)	-71.20
τ(N1C3C7H9)	170.11
τ(C28N1C3H4)	35.29
τ(C28N1C3C5)	-80.70
τ(N1C3C5O6)	-20.19
τ(N1C3C5O16)	161.25
τ(C3N1C28C25)	179.96
τ(C3N1C28O29)	-0.12
τ(H8C7C3H4)	173.70
τ(C10C7C3H4)	48.30
τ(H9C7C3H4)	-70.38
τ(H8C7C3C5)	-72.13
τ(C10C7C3C5)	162.47
τ(H9C7C3C5)	43.78
τ(C7C3C5O6)	105.47
τ(C7C3C5O16)	-73.08
τ(C3C7C10C13)	71.28
τ(C3C7C10H11)	-169.75
τ(C3C7C10H12)	-52.40
τ(H4C3C5O6)	-137.90
τ(H4C3C5O16)	43.54
τ(C3C5O6H17)	-178.34
τ(H17O6C5O16)	0.21
τ(H8C7C10C13)	-53.59
τ(H8C7C10H11)	65.38
τ(H8C7C10H12)	-177.27
τ(C13C10C7H9)	-169.96
τ(H11C10C7H9)	-51.00
τ(H12C10C7H9)	66.36
τ(C7C10C13O15)	48.06
τ(C7C10C13O14)	-134.41
τ(O15C13C10H11)	-72.02
τ(O14C13C10H11)	105.51
τ(O15C13C10H12)	173.46
τ(O14C13C10H12)	-9.00
τ(C10C13O15H18)	177.87
τ(H18O15C13O14)	0.25
τ(O34C32N19H20)	179.96
τ(C33C32N19H20)	0.21
τ(O34C32N19C21)	0.99
τ(C33C32N19C21)	-178.76
τ(C32N19C21C25)	111.11
τ(C32N19C21H22)	-5.71
τ(C32N19C21C23)	-120.72
τ(N19C32C33H35)	120.19
τ(N19C32C33H36)	-122.77
τ(N19C32C33H37)	-1.23
τ(H20N19C21C25)	-67.86
τ(H20N19C21H22)	175.32
τ(H20N19C21C23)	60.31
τ(H27C25C21N19)	-60.28
τ(H26C25C21N19)	-177.17
τ(C28C25C21N19)	63.22
τ(H27C25C21H22)	56.45
τ(H26C25C21H22)	-60.44
τ(C28C25C21H22)	179.96
τ(H27C25C21C23)	173.03
τ(H26C25C21C23)	56.14
τ(C28C25C21C23)	-63.46
τ(C25C21C23O24)	124.06
τ(C25C21C23O30)	-57.82
τ(C21C25C28N1)	-112.93
τ(C21C25C28O29)	67.15
τ(N19C21C23O24)	-3.64
τ(N19C21C23O30)	174.48
τ(H22C21C23O24)	-118.98
τ(H22C21C23O30)	59.15
τ(C21C23O30H31)	-179.43
τ(O24C23O30H31)	-1.26
τ(H27C25C28N1)	8.49
τ(H27C25C28O29)	-171.44
τ(H26C25C28N1)	126.75
τ(H26C25C28O29)	-53.18
τ(O34C32C33H35)	-59.56
τ(O34C32C33H36)	57.48
τ(O34C32C33H37)	179.02

Table S4.4 Calculated and observed wavenumber assignments (cm^{-1}) for vibrational bands of L-Asp-L-Glu (D_2O).

Calculated (cm^{-1})	Experimental Raman	ROA	Assignments (%PEDs)
1713	1700	1700	ν(C25O26) (86)
1701			ν(C10O12) (85)
1645	1682		ν(C5O6) (69), ν(C5N13) (20)
1557			$\nu_{as}(CO_2)$ (90)
1492			ν(C5O6) (12), ν(C5N13) (40), ν(C3C5) (12), δ_{ip}(ND) (10)
1458	1474	1471	$\delta(CH_2')$ (94)
1430		1448	$\delta(CH_2'')$ (93)
1423	1419	1409	$\delta(CH_2)$ (88)
1417			ν(C7C10) (11), $\omega(CH_2)$ (54)
1410			$\nu_s(CO_2)$ (39), ν(C15H17) (11), $\omega(CH_2')$ (16)
1394			ν(C22C25) (11), $\tau(CH_2')$ (12), $\omega(CH_2'')$ (42)
1385			δ(C3H4) (73)
1354		1362	$\nu_s(CO_2)$ (16), $\omega(CH_2')$ (29), δ(C15H16) (31)
1328		1328	$\omega(CH_2')$ (28), δ(C15H16) (34)
1318			$\tau(CH_2)$ (18), δ(C3H4) (27), δ(C15H16) (13)
1302			$\tau(CH_2)$ (11), $\tau(CH_2')$ (14), δ(C3H4) (10), δ(C15H16) (32)
1260			ν(C10O11) (19), $\tau(CH_2)$ (42)
1237			ν(C25O27) (12), $\omega(CH_2'')$ (15), $\tau(CH_2'')$ (45)
1227			ν(C10O11) (11), ν(C25O27) (10), $\omega(CH_2)$ (16), $\tau(CH_2')$ (14), δ(C3H4) (25)
1224			$\omega(CH_2)$ (14), $\tau(CH_2')$ (11), δ(C3H4) (12), δ(C15H16) (12)
1183	1208	1266	$\delta_s(ND_3)$ (37), $\delta_{as}(ND_3)$ (52)
1179			$\delta_{as}(ND_3)$ (95)
1174		1185	ν(C25O27) (11), $\tau(CH_2')$ (20), $\omega(CH_2'')$ (13), $\tau(CH_2'')$ (16), δ(C15H16) (11)
1140			ν(C3N1) (22), ν(C3C7) (14), $\delta_s(ND_3)$ (16), $\delta_{as}(ND_3)$ (23)
1108			ν(C15N13) (19), ν(C3C7) (17),$p(CH_2)$ (10)
1097			ν(C15N13) (22), ν(C3C7) (15), $\delta_s(ND_3)$ (19)
1075			ν(C15C19) (18), ν(C19C22) (35)
1070			ν(C15C19) (22),$p(CH_2')$ (11),$p(CH_2'')$ (17)
1047			ν(C19C22) (11), δ(C10O11H30) (18)
1034		1052	ν(C19C22) (13)
1022	1018		ν(C10O11) (15), δ(C10O11H30) (20)
1014			ν(C25O27) (15, ν(C3C5) (12), δ(C25O27H31) (23)
994			δ(C25O27H31) (17)

Table S4.4 Continued

898	930	930	*ν*(C15C17) (15), δ_{ip}(ND) (10),p(CH_2'') (25), δ(O26C25O27) (13)
885		886	*ν*(C7C10) (25), δ(C10O11H30) (12)
872			$p_{ip}(ND_3)$ (24), δ_{ip}(C5O6) (12)
851			*ν*(C22C25) (10)
835		835	$p_{op}(ND_3)$ (19),p(CH_2) (14)
816			*ν*(C3N1) (18),$p_{op}(ND_3)$ (40)
785			*ν*(C22C25) (19), δ(O18C17O28) (18), ω(O18C17O28) (12)
750			*ν*(C22C25) (12),p(CH_2') (16), ω(O18C17O28) (41)
741			δ_{op}(C5O6) (55)
681			*ν*(C3C5) (13),$p_{ip}(NH_3)$ (23), δ(O18C17O28) (12)
633			δ(O18C17O28) (11), ω(O26C25O27) (22)
620			δ_{ip}(C5O6) (15), ω(O11C10O12) (23)
583			δ(O11C10O12) (20)
565			δ(CH_2) (11), δ(O11C10O12) (26), ω(O11C10O12) (12)
552			δ(O26C25O27) (37), ω(O26C25O27) (10)
492			δ(O26C25O27) (21),p(O26C25O27) (10)
464			δ(N1C3C7) (10),p(O11C10O12) (11),p(O18C17O28) (13)
439			δ_{op}(ND) (43), τ(C5N13) (21)
427			τ(C25O27) (77)
424			p(CH_2) (10), ω(O11C10O12) (13), τ(C10O11) (67)

ν, stretching; as, anti-symmetric; s, symmetric, δ, deformations; ω, wagging; τ, twisting; p, rocking; ip,in-plane; op, out-of-plane;

Table S4.5 Calculated and observed wavenumber assignments (cm^{-1}) for vibrational bands α-NAAG (D_2O).

Calculated (cm^{-1})	Experimental Raman	Experimental ROA	Assignments (%PEDs)
1726	1712	1710	ν(C11O12) (87)
1686			ν(C8O10) (43), ν(C28O29) (41)
1668			ν(C8O10) (39), ν(C28O29) (42)
1637	1653		ν(C2O3) (73)
1624			ν(C23O25) (75)
1479			ν(C2N4) (19), ν(C23N24) (17), $\delta_{as}(CH_3)$ (15)
1472	1463	1473	ν(C2N4) (24), ν(C23N24) (11), $\delta_{as}(CH_3)$ (15)
1461			$\delta(CH_2')$ (87)
1452			$\delta(CH_2'')$ (85)
1451		1450	$\delta(CH_2)$ (91)
1435			$\delta_{as}(CH_3)$ (87)
1418	1421	1415	ν(C23N24) (22), $\delta_{as}(CH_3)$ (46)
1395			$\tau(CH_2)$ (21), $\omega(CH_2')$ (15), $\tau(CH_2')$ (11)
1376			δ(C5H15) (10), $\omega(CH_2)$ (53)
1374			$\delta_s(CH_3)$ (47), $\delta_{as}(CH_3)$ (10)
1369			ν(C28O30) (12), $\omega(CH_2'')$ (15), $\delta_s(CH_3)$ (32)
1365			δ(C5H15) (32), $\omega(CH_2')$ (10)
1344		1309	δ(C25H1) (74)
1310			δ(C5H15) (28), $\omega(CH_2')$ (21), $\tau(CH_2')$ (19)
1295			ν(C28O30) (20), δ(C25H1) (18), $\omega(CH_2'')$ (30), $\tau(CH_2'')$ (13)
1282			ν(C8O9) (16), δ(C5H15) (27), $\omega(CH_2')$ (19)
1249			δ(C5H15) (24), $\tau(CH_2)$ (43)
1240			ν(C25C27) (10), δ(C25H1) (38), $\omega(CH_2'')$ (24)
1225			ν(C11O13) (35), δ(C5H15) (26), $\tau(CH_2)$ (12), δ(C12C11O13) (10)
1210	1208	1252	ν(C25N24) (14), δ(C25H1) (13), $\tau(CH_2'')$ (38)
1179			δ(C5H15) (12), $\omega(CH_2)$ (10), $\tau(CH_2')$ (50)
1147			ν(C25N24) (34), ν(C5C6) (11)
1126			ν(C25N24) (20), $\tau(CH_2'')$ (14), $p_{ip}(CH_3)$ (11)
1092			δ(C28O30H37) (25), $p(CH_2'')$ (12)
1077			δ(C8O9H20) (19)
1060		1096	ν(C5C6) (11), ν(C6C7) (25)
1048		1145	ν(C8O9) (11), ν(C6C7) (11), δ_{ip}(N4D14) (11)
1042			$p_{op}(CH_3)$ (55), δ_{op}(C23O26) (16)
1039	1072		ν(C28O30) (12), δ(C28O30D37) (24), $p_{op}(CH_3)$ (14)
1030			ν(C25C27) (12), $p_{ip}(CH_3)$ (20)

Table S4.5 Continued

1017		1015	$p_{ip}(CH_3)$ (11)
973			*v*(C2N4) (11), *v*(C25C27) (16), δ_{ip}(N4D14) (25)
959	950		*v*(C25C27) (10), *v*(C22C23) (14), p(CH_2'') (23)
942		974	p(CH_2') (27), ω(O9C8O10) (10)
938		880	*v*(C23N24) (15), δ_{ip}(N24D34) (23), p(CH_2') (10), $p_{ip}(CH_3)$ (11)
919			δ(C2N4C5) (11)
890			*v*(C5C6) (17), *v*(C6C7) (15), *v*(C7C8) (19)
855			*v*(C27C28) (12)
828			ω(O12C11O13) (10), ω(O29C28O30) (13)
801			*v*(C28O30) (12), *v*(C27C28) (27), δ_{op}(C2O3) (11), ω(O29C28O30) (10)
776			*v*(C5C11) (11), *v*(C7C8) (23), p(CH_2) (17)
742			p(CH_2) (12), ω(O9C8O10) (42)
731			δ_{op}(C2O3), ω(O12C11O13) (13), ω(O29C28O30) (23)
709		690	*v*(C5C11) (11), *v*(C5C6) (14), δ_{op}(C2O3) (12), ω(O12C11O13) (14)
647		659	δ(C2C25N24) (15), δ_{ip}(C23O26) (12), δ_{op}(C23O26) (20)
637		625	*v*(C22C23) (14), δ_{ip}(C2O3) (12), δ_{ip}(C23O26) (15)
608		602	$p_{op}(CH_3)$ (10), δ_{op}(C23O26) (39)
587			δ(O12C11O13) (27), ω(O12C11O13) (11)
569			*v*(C27C28) (10), δ(O29C28O30) (58)
565			*v*(C7C8) (11), δ(O9C8O10) (61)
536			τ(C8O9) (17), τ(C28O30) (19)
531		573	δ(N4C5C6) (14), δ(C5C6C7) (11)
511			δ_{ip}(C23O26) (13), τ(C8O9) (13), τ(C28O30) (31)
493		493	τ(C8O9) (54), τ(C28O30) (24)
463			δ(N24C25C27) (10), p(CH_2'') (10), δ(O29C28O30) (37)
461			δ(C2C25N24) (10), δ(O12C11O13) (16), p(O12C11O13) (21)
434			p(O9C8O10) (34), τ(C2N4) (14)
429		443	p(O9C8O10) (41), δ_{op}(N4D14) (14), δ_{op}(N24D34) (10)
419			p(O29C28O30) (10), δ_{op}(N4D14) (15), δ_{op}(N24D34) (20), τ(C23N24) (12)
404			τ(C11O13) (55), τ(C2N4) (11)

v, stretching; as, anti-symmetric; s, symmetric, δ, deformations; ω, wagging; τ, twisting; p, rocking; ip,in-plane; op, out-of-plane;

Table S4.6 Calculated and observed wavenumber assignments (cm^{-1}) for vibrational bands of β-NAAG (D_2O)

Calculated (cm^{-1})	Experimental Raman	ROA	Assignments (%PEDs)
1729	1716	1720	ν(C23O24) (87)
1716			ν(C5O16) (78)
1710			ν(C5O16) (10), ν(C13O14) (76)
1631	1643	1620	ν(C28O29) (78)
1619			ν(C32O34) (77)
1486		1530	ν(C32N19) (31), $\delta_{as}(CH_3)$ (20)
1476	1471	1480	ν(C28N1) (32), ν(C25C28) (10), $\delta(CH_2'')$ (23)
1454			$\delta(CH_2)$ (58), $\delta(CH_2'')$ (32)
1449		1439	ν(C28N1) (11), $\delta(CH_2)$ (35), $\delta(CH_2'')$ (38)
1441			$\delta(CH_2')$ (18), $\delta_{as}(CH_3)$ (50)
1440		1406	$\delta(CH_2')$ (75), $\delta_{as}(CH_3)$ (11)
1434			$\delta_{as}(CH_3)$ (59)
1384			δ(C3H4) (11), δ(C21H22) (12), $\delta_{as}(CH_3)$ (11)
1379		1374	δ(C3H4) (12), $\omega(CH_2)$ (27), $\tau(CH_2)$ (10), $\delta_s(CH_3)$ (10)
1373			$\omega(CH_2)$ (11), $\delta_s(CH_3)$ (52)
1366			δ(C3H4) (10), $\omega(CH_2)$ (28), $\tau(CH_2')$ (10)
1357			δ(C3H4) (18), $\omega(CH_2')$ (30)
1344			δ(C3H4) (18), δ(C21H22) (27), $\omega(CH_2'')$ (28)
1315			ν(C5O6) (16), δ(C3H4) (36), $\omega(CH_2')$ (10)
1297			ν(C5O6) (17), ν(C13O15) (12), ν(C3C5) (11), $\tau(CH_2')$ (17)
1276			ν(C13O15) (15), δ(C3H4) (22)
1268		1309	δ(C21H22) (32), $\omega(CH_2'')$ (26)
1248			δ(C21H22) (26), $\tau(CH_2)$ (12), $\tau(CH_2'')$ (10)
1247	1208	1252	δ(C21H22) (15), $\tau(CH_2)$ (17), $\omega(CH_2')$ (14)
1219			ν(C23O20) (18), $\tau(CH_2'')$ (52)
1187			$\omega(CH_2)$ (10), $\tau(CH_2)$ (10), $\tau(CH_2')$ (33)
1149			ν(C21N19) (33)
1133		1135	ν(C3N1) (32)
1090			ν(C21N19) (10), δ_{ip}(N19D29) (30), $p_{ip}(CH_3)$ (16)
1087		1091	ν(C7C10) (12), $p(CH_2)$ (15), $p(CH_2')$ (10)
1056			ν(C7C10) (15), ν(C3C7) (22)
1046			ν(C7C10) (16), $p_{op}(CH_3)$ (36), δ_{op}(C32O34) (10)
1042			$p_{op}(CH_3)$ (30)
1027			ν(C32C33) (14), $p_{ip}(CH_3)$ (25)
1011			ν(C23O30) (10), δ(C5O6D17) (20), δ(C23O30D31) (17)
1006			ν(C5O6) (12), δ(C5O6D17) (34)
998			ν(C13O15) (20), δ(C13O15D18) (39)
986		990	ν(C21C25) (11), ν(C32C33) (12), δ_{ip}(N1D2) (13)
967			δ(C23O30D31) (10), δ_{ip}(N1D2) (13), $p(CH_2'')$ (13)
946	948	956	ν(C32N19) (24), δ_{ip}(N19D29) (24), $p_{ip}(CH_3)$ (11)

Table S4.6 Continued

905			*v*(C3C7) (13)
895			p(CH_2) (20)
854	854	880	*v*(C3C5) (11), p(CH_2') (12)
838			δ_{op}(C28O29) (11)
783		801	*v*(C21C23) (19), δ_{op}(C28O29) (17)
768			*v*(C10C13) (26), δ(C13O15D18) (26), p(CH_2) (17)
720			*v*(C3C7) (11), ω(O6C15O16) (43) ω(O14C13O15) (10)
702			ω(O24C23O30) (42)
658		677	p(CH_2') (14), ω(O14C13O15) (34)
644		646	*v*(C25C28) (13), δ_{ip}(C28O29) (13)
623			p_{op}(CH_3) (17), δ_{op}(C32O34) (69)
611			*v*(C32C33) (19), δ_{ip}(C32O34) (31)
589			δ(C5O6D17) (10), δ(O6C5O16) (20), δ(O24C23O30) (13)
573		565	δ(O6C5O16) (23)
536			δ(O14C13O15) (32)
527			δ(O14C13O15) (35)
507			p(O24C23O30) (17), δ(N19C32C33) (18), δ_{ip}(C32O34) (15)
487			p(O14C13O15) (20), τ(C13O15) (13)
476			δ(N19C21C25) (15), p(O14C13O15) (10), τ(C23O30) (11)
448			τ(C5O6) (82)
427		537	p(O14C13O15) (14), τ(C13O15) (65)
426			δ_{op}(N19D20) (20), δ_{op}(C23O30) (12), τ(N19C32) (17)
405			δ_{op}(N1D2) (35), τ(N1C28) (20)

v, stretching; as, anti-symmetric; s, symmetric, δ, deformations; ω, wagging; τ, twisting; p, rocking; ip,in-plane; op, out-of-plane;

Chapter 5. Raman and ROA studies of the amino acid derivatives: *N*-acetyl-L-Asp and *N*-acetyl-L-Glu

No decision has been made about the submission of this paper.

Personal contribution to the paper

I was responsible for the experimental work and the DFT calculations were provided by Dr. Trevor Dines from the University of Dundee. The amino acid samples were provided by Dr. Nighat Kausar and Professor Babur Chowdhry from University of Greenwich. All sample preparation and measurement were performed by me and analysis in OriginPro was performed by me. I wrote the paper and the discussion and conclusions were drawn from the results, a review of the literature and discussions with Dr Ewan Blanch, Dr Trevor Dines and Professor Babur Chowdhry.

Chapter 5.

Raman and ROA studies of the amino acid derivatives: *N*-acetyl-L-Asp and *N*-acetyl-L-Glu in solution

Clare Levene, [1] Trevor Dines, [2] Saeideh Ostovar pour, [1] Nighat Kausar, [3] Babur Z. Chowdhry, [3] and Ewan Blanch [1]*

[*] *Correspondence to Ewan Blanch, E-mail*: E.Blanch@manchester.ac.uk

[1] *Faculty of Life Sciences, MIB, University of Manchester, 131 Princess Street, Manchester, M1 7DN UK.*

[2] *Division of Electronic Engineering and Physics, University of Dundee, Dundee, DD1 4HN, UK.*

[3] *School of Science, University of Greenwich at Medway, Central Avenue, Chatham Maritime, Kent, ME4 4TB, UK.*

5.1 Abstract

Experimental Raman and Raman optical activity (ROA) studies have been carried out for the amino acid derivatives; *N*-acetyl-L-Asp and *N*-acetyl-L-Glu. These results have been compared to the experimental spectra of L-Asp and L-Glu to identify site-specific markers of acetylation, vibrational modes that interact with the solvent and structural differences observed experimentally. Significant differences were measured in the Raman and ROA spectra for the acetylated and nonacetylated amino acids. Site-specific acetylation ROA marker bands were identified in the amide I region and several bands were observed of opposite sign in the 800 - 1000 cm^{-1} spectral region for the acetylated amino acids, thus providing ROA marker bands for structural differences between these amino acids. The combination of Raman and ROA spectroscopy allows for the detection of chemical modification of amino acids, however ROA provides more structural information and is a powerful technique for future studies relating to chemical modification of peptides and proteins.

Keywords: Raman spectroscopy; Raman Optical Activity; *N*-acetyl-L-Asp; *N*-acetyl-L-Glu; acetylation; deuteration.

5.2 Introduction

There have been many computational and experimental studies into the conformations of native amino acids over the years in the solid state, gas phase and in solution (Pawlukojc *et al* 2002; Chitra *et al* 2003; Shigeto & Dlott 2007). Amino acids are the building blocks of proteins and typically appear in eukaryotes in the L form. Aspartic acid (L-Asp) and glutamic acid (L-Glu) are the only two naturally occurring amino acids that have one amino group and two carboxylic acid groups and are also biologically important in their own right. L-Asp is a metabolite in the urea cycle and also participates in gluconeogenesis, one of two main mechanisms that prevent glucose levels in humans and many animals from dropping too low. L-Asp may serve as an excitatory transmitter in the brain and may lead to physical endurance by providing resistance to fatigue (Matsres *et al*. 1978; Garner & Spector 1978). Studies on L-Glu have shown that it has a crucial role in nitrogen metabolism in biological systems and it also plays an active role in parental nutrition (Maeda *et al*. 1958). L-Glu is also a major degradation product of tumour cells (Taday *et* al. 2003) and a neurotransmitter in the central nervous system (CNS) (Hudson *et al*. 1976).

N-acetyl-L-aspartic acid (*N*-acetyl-L-Asp) is a nervous system-specific metabolite present in the brain at exceptionally high concentrations (Tallan *et al.* 1956; Tallan 1957). It is one of the most concentrated molecules in the CNS and one of the most reliable markers for brain magnetic resonance spectroscopy (MRS) studies (Barany *et al*. 1987; Fan *et al*. 1986). Increased high levels of *N*-acetyl-L-Asp have been connected to the rare, but fatal, genetic disorder known as Canavan disease (Bartalini

et al. 1992; Divry & Mathieu 1989; Matalon *et al.* 1988) and when observed in decreased levels *N*-acetyl-L-Asp has been associated with brain cancer, brain injury, multiple sclerosis (MS) and Alzheimer's disease (Danielsen & Ross 1999). *N*-acetyl-L-glutamic acid (*N*-acetyl-L-Glu) was first detected in the human brain in 1966 (Audtoe *et al.* 1966). A conformational and electronic study on *N*-L-acetyl-L-glutamate-*N*-methylamide has been performed at the RB3LYP/6-31G(d) level of theory (Masman *et al*. 1988). Under biological conditions *N*-acetyl-L-Glu can deprotonate easily, which then leaves the negatively charged side-chain available for salt-bridge formation. Salt-bridge formation is typically observed to contribute stability to the folded conformations of proteins, although these noncovalent interactions are known to be relatively weak small stabilizing interactions make an important contribution to the overall stability of the conformer (Székely *et al*. 1996; Doherty 2006).

N-terminal acetylation was first reported in 1958 by Narita who found an acetyl peptide in a natural protein (Narita 1958). Since then, acetylation has been correlated with hormonal regulation (Jungmann & Schweppe 1972; Libby 1972), transcriptional regulation (Mizzen & Allis 2000), gene control (Briggs *et al.* 2002), and cancer therapies (Brown & Strathdee 2002; Johnstone 2002). Acetylation is a co-translational and post-translational modification to proteins and the identification of acetylation sites is the essential first step towards molecular biological studies into acetylation (Mizzen & Allis 2000). The acetylation bioprocess involves the substitution of an active hydrogen atom by an acetyl group (CH_3CO) at the N-terminal of L-Asp or L-Glu.

Molecular structure and dynamics of amino acids can be investigated using vibrational spectroscopy, which is a powerful group of techniques that can be applied to a wide range of biological molecules as is evident from numerous IR, Raman and ROA studies (Blanch *et* al. 2004; Choi *et al*. 2005; Edler 2002; Fang *et al*. 2003; Bowen & Brunge 2006; Pozo Ramajo 2006; Chin *et al*. 2006; Ashton *et al*. 2009). Precise structure determination using vibrational spectroscopy relies upon the selection of vibrational modes that are sensitive to conformational changes. Conformationally sensitive bands in aqueous solution include amide I vibrations (Krimm & Bandekar 1986; Bandekar 1992; Watson & Hirst 2002), amide III vibrations (Krimm & Bandekar 1986; Bandekar 1992; Watson & Hirst 2002; Watson & Hirst 2004), NH stretching vibrations (Burgess & Scheraga 1973; Maxfield *et al*. 1979) and Raman skeletal vibrations (Deng *et al*. 1996; Takekiyo *et al*. 2004). The most frequently used bands for structural determination are the amide I and the amide III bands as the amide I and III regions of vibrational spectra have been shown to be useful for determining backbone conformations (Grdadolnik 2008).

In this study we have used Raman spectroscopy and Raman optical activity (ROA), which measures a small difference in Raman scattering by chiral molecules using left- and right-circularly polarized scattered light (Atkins & Barron 1969; Barron *et al*. 1972; Hug & Hangartner 1999), and is highly sensitive to molecular structure, to investigate the solution structures of *N*-acetyl-L-Asp and *N*-acetyl-L-Glu. In order to understand the structures of amino acids it is essential to determine their conformational preferences (Grdadlonik 2008) and we therefore present a combined Raman and ROA study characterizing the structural differences between these two *N*-protected amino acids, which differ only with the addition of an extra CH_2 group

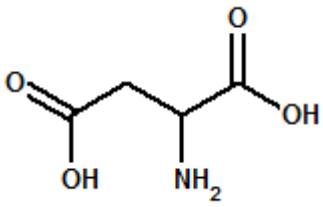
in *N*-acetyl-L-Glu. We have also compared the spectra of the acetylated amino acids to those of the native amino acids, L-Asp and L-Glu, to determine marker bands of site-specific acetylation. Deuterated samples were also investigated to determine which vibrational modes involve functional groups that interact with the solvent. ROA analysis has not previously been presented for L-Asp, L-Glu, *N*-acetyl-L-Asp or *N*-acetyl-L-Glu.

5.3 Experimental

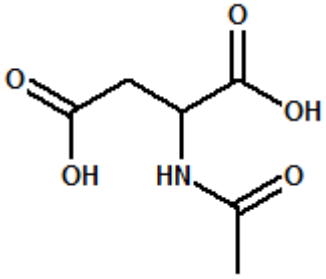
The amino acid derivatives, *N*-acetyl-L-Asp and *N*-acetyl-L-Glu, were purchased from Bachem, St Helens, UK (ca. 99.0% purity). L-Asp and L-Glu were purchased from Sigma-Aldrich Company Ltd, Dorset, UK. All reagents were used without further purification. All samples were prepared in 100µl of a phosphate buffer solution (0.20 M and pH 7.01) to a final concentration of 50 mg/ml and centrifuged for 5 mins at 3000 rpm to remove insoluble impurities such as dust particles prior to data collection. Deuterated samples were also prepared using D_2O under the same conditions as for samples in H_2O. Samples were then pipetted into quartz microfluorescence cells (Optiglass, Ilford, Essex, UK) for spectroscopic measurement.

Raman and ROA spectra were collected using a BioTools ChiralRAMAN spectrometer (BioTools Inc., Jupiter, FL, USA) operated via Critical Link software. The instrument has a design based on the backscattering geometry using a Nd/VO_4 laser with an excitation wavelength of 532.5 nm and spectral resolution of ~7 cm^{-1}. Laser power at the samples was 600 mW with an illumination time of 1.9845

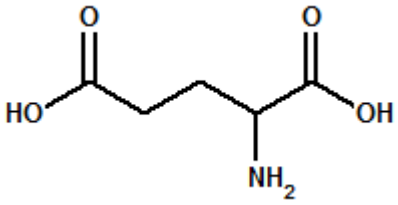
seconds per scan and overall data accumulation times of 6-24 hrs. The results were plotted, analyzed and interpreted using OriginPro 8 software (OriginLab Corp., Northampton, MA, USA). The baseline tool in OriginPro 8 was used to perform baseline corrections. Regions of the spectra where no bands were observed were adjusted to zero intensity. All spectra were normalized with respect to experimental parameters such as accumulation time to allow for quantitative comparison between samples.



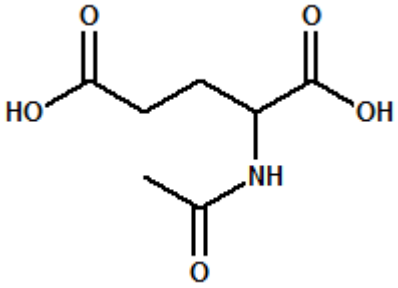
(1)



(2)



(3)



(4)

Fig. 5.1 Schematic representations of: (1) L-Asp, (2) N-acetyl-L-Asp (3) L-Glu and (4) *N*-acetyl-L-Glu.

5.4 Results and discussion

5.4.1 Vibrational Assignments

The molecular structures of L-Asp, L-Glu, *N*-acetyl-L-Asp and *N*-acetyl-L-Glu are shown in Fig. 5.1. The band assignments observed for L-Asp, L-Glu, *N*-acetyl-L-Asp and *N*-acetyl-L-Glu are shown in Tables 5.1 and 5.2, in H_2O and D_2O, respectively.

N-acetyl-L-Asp is comprised of 21 atoms which give rise to 57 normal vibrational modes while *N*-acetyl-L-Glu is comprised of 23 atoms with 63 normal vibrational modes. L-Asp and L-Glu are comprised of 16 and 19 atoms, respectively, giving rise to 42 and 51 normal vibrational modes, respectively. In principle, all of these vibrational modes are Raman active, however, due to the considerable amount of Raman active vibrations Raman spectra can be complicated and difficult to interpret, because of the overlap of bands. ROA, which is sensitive to chirality and which generates its strongest bands from the most chiral and rigid parts of a molecule (Barron *et al*. 2000), can provide additional and more incisive information which is highly useful for probing biomolecular structure and behaviour in aqueous solution.

It is first of interest to identify marker bands that are specific to L-Asp and *N*-acetyl-L-Asp and similarly bands specific to L-Glu and *N*-acetyl-L-Glu that is the identity of the amino acid. The ability of Raman and ROA spectral details to identify site-specificity of acetylation of these simple amino acids can then be assessed. The observed Raman and ROA spectra for the four samples measured in H_2O are shown in Fig. 5.2, while the corresponding Raman and ROA spectra for these samples in

D_2O are shown in Fig. S5.1 (see Supplementary Information). Deuterated samples were measured for all four amino acids in order to define which vibrations are influenced by deuteration of labile hydrogens.

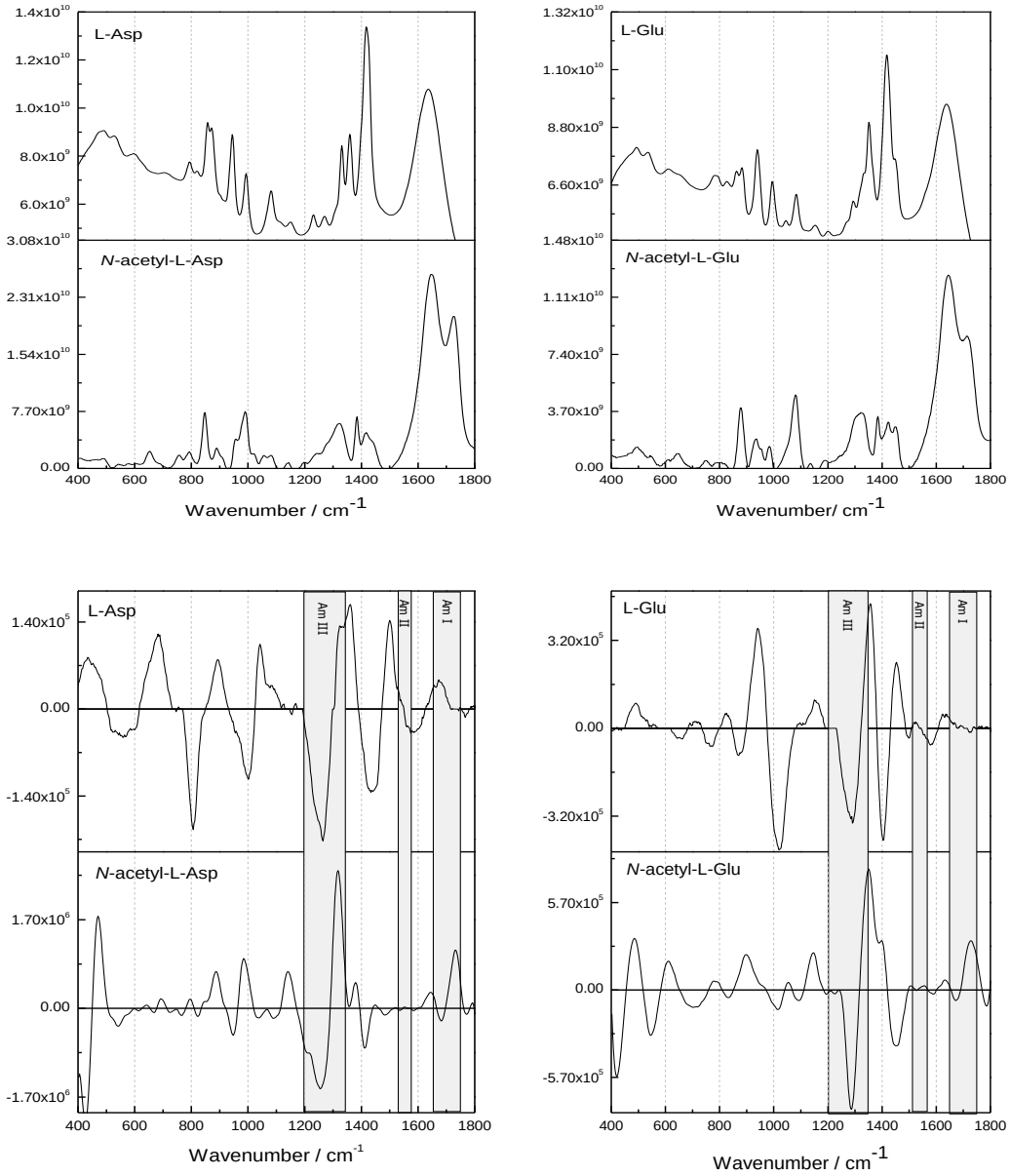


Fig. 5.2 Observed spectra in H_2O. Raman spectra of; (Top) L-Asp *N*-acetyl-L-Asp, L-Glu and *N*-acetyl-L-Glu, ROA spectra of; (Bottom) L-Asp, *N*-acetyl-L-Asp, L-Glu and *N*-acetyl-L-Glu.

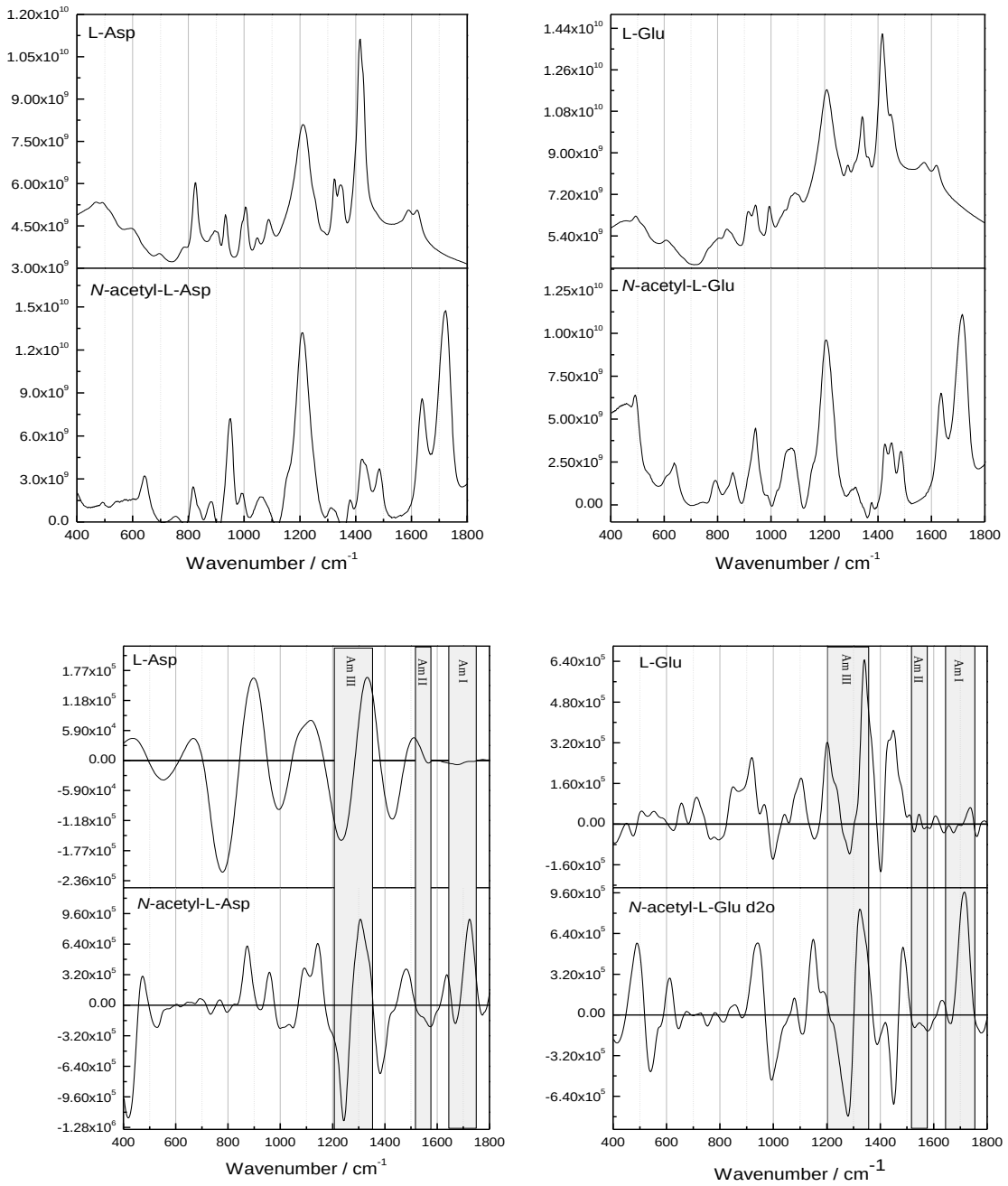


Fig. 5.3 Observed spectra in D_2O, (Top) Raman spectra; L-Asp *N*-acetyl-L-Asp, L-Glu and *N*-acetyl-L-Glu and (Bottom) ROA spectra; L-Asp, *N*-acetyl-L-Asp, L-Glu and *N*-acetyl-L-Glu.

Table 5.1. Observed Raman and ROA bands for the amino acids measured in H_2O (cm^{-1}).

L-Asp		***N*-acetyl-L-Asp**		**L-Glu**		***N*-acetyl-L-Glu**	
Raman	**ROA**	**Raman**	**ROA**	**Raman**	**ROA**	**Raman**	**ROA**
485	+483		+471	492	+489		+485
535			-540	538		504	
	-562					550	-544
596							+608
		653			-653	646	
	+681		+695		+704		-704
788	+758	754		785	-763	747	
	-804	795	+795				+781
859		845	-822	831	+822		-832
				866	-868	878	
	+891	890	+886	886			+895
940			-946	941	+941	934	
		982	+981			984	
1005	-1000			992	-1019		-1014
	+1042		-1037				+1055
1076		1080		1082	+1082	1080	-1096
1157		1142	+1142	1156	+1151		+1147
1228							
1268	-1261		-1256				
			+1315		-1292		-1283
1334		1320			+1356	1318	
1359	+1356			1351			+1347
		1379	+1380		-1397	1384	
1420	-1434	1416	-1407	1416		1424	+1402
					+1453	1450	-1452
	+1500				-1498		
	-1584				-1584		
1632		1649	+1639	1634	+1630	1647	+1630
	+1676		-1681				-1676
		1721	+1731			1712	+1726

+ = positive observed ROA bands, - = negative observed ROA bands

Table 5.2. Observed Raman and ROA bands for the amino acids in D_2O (cm^{-1}).

L-Asp		***N*-acetyl-L-Asp**		**L-Glu**		***N*-acetyl-L-Glu**	
Raman	**ROA**	**Raman**	**ROA**	**Raman**	**ROA**	**Raman**	**ROA**
	+435						
			+472	485			+489
	-544		-528				-538
				602			+614
	+633	645			+653	645	-639
			+700	698	+714		+720
		756	-731			756	
	-777		+762				+781
827		816		829		816	-810
		845	+873				+850
898	+895	887		900	+916	882	
933			-923	935			+941
		948	+954		+962	948	
1004	-996	993		1006	-1002	993	-991
1040				1052	+1043		
1085		1064	+1090	1087	+1098	1064	+1082
	+1114						-1105
			+1141		-1149		+1151
1205		1206		1210	+1199	1206	+1186
	-1233		-1238				
			+1300		-1283		-1279
1323	+1329	1317		1322		1317	+1316
1348				1347	+1342		
		1373	-1380		-1397	1378	-1388
1414	-1424	1423		1411		1423	
		1484	+1480		+1448	1484	-1452
	+1507						+1490
					+1544		
			-1571				
1591				1591			
1621		1635		1621	+1609	1635	
			-1672		-1675		-1672
		1721	+1723		+1736	1723	+1711

+ = positive observed ROA bands, -= negative observed ROA bands

5.4.2 Spectral Region 1800-1200 cm^{-1}

For the acetylated amino acids the vibrations observed in this region of the spectrum are predominantly due to amide I, II and III vibrations which are sensitive to secondary structure. Native amino acids do no have amide moieties, however, they do have vibrational transistions which are similar to amide vibrations and these vibrational transitions were compared to the amide vibrations of the acetylated amino acids to distinguish differences between these amino acids. Amide I vibrations, which usually occur between 1640 and 1730 cm^{-1}, are sensitive to changes in conformation due to the large contribution from backbone C=O stretching coordinates (Bandekar 1992; Krimm & Bandekar 1986; Watson & Hirst 2002). The amide II vibrations, which are due to contributions from C-N stretch and N-H deformations, occur between 1530 and 1580 cm^{-1} depending upon the solvent used, however, for samples analyzed in D_2O where the amide proton can be replaced by deuterium, the amide II can shift to values below 1500 cm^{-1}. This is due to interactions with CH deformation vibrational modes (Chen *et al*. 1995; Mirkin & Krimm 1991; Miyazawa *et al*. 1956). The amide III vibrations are observed between 1200 and 1340 cm^{-1}, with contributions from NH in-plane bending and CN stretch modes (Chen *et al*. 1995; Mirkin & Krimm 1991; Miyazawa *et al*. 1956). Other contributing vibrations in the 1800 – 1200 cm^{-1} spectral region include CH wagging, twisting and deformations and COH deformations.

The amide I vibration is dependent upon protonation and deprotonation between the carbonyl and NH groups and interactions with the solvent and, with this in mind the amide I vibrational wavenumber can decrease with increased hydrogen bonding

between the solvent molecules making it a commonly used vibrational mode for the determination of secondary structure (Wang *et al.* 1991; Torii *et al*. 1998; Mirkin & Krimm 1996; Schweitzer-Stenner *et al*. 1998). For the Raman spectrum of *N*-acetyl-L-Asp there is a band observed at ~1720 cm^{-1} which we have assigned to the amide I vibration. There is also a band observed at 1649 cm^{-1} which suggests the amide vibration is interacting with the acetyl side chain. For the corresponding sample analyzed in D_2O the amide I vibration is observed at ~1721 cm^{-1} with a band observed at ~1635 cm^{-1}, which is further supporting evidence of vibrational mixing between the acetyl and amide groups. However, the intensity for the deuterated amide vibration is higher than for the sample analyzed in H_2O which is suggestive of a higher level of H/D exchange. The Raman spectrum for L-Asp displays a broad band at ~1632 cm^{-1} which we have assigned to C=O stretching with the sample in D_2O displaying a very weak band in the same area. In the ROA spectrum for *N*-acetyl-L-Asp there is an observed couplet where a positive signed band is observed at ~1731 cm^{-1} and a negative signed weak band at ~1681 cm^{-1}. For the sample analyzed in D_2O there is a positive band is observed at ~1723 cm^{-1} with a negative weak band at ~1672 cm^{-1}. The spectrum for L-Asp displays a weak positive band at ~1630 cm^{-1} which we have assigned to C=O stretching. The Raman and ROA spectra of L-Glu and *N*-acetyl-L-Glu share similarities with those observed for L-Asp and *N*-acetyl-L-Asp. The C=O vibration for L-Glu in H_2O generates a broad Raman band observed at ~1634 cm^{-1} with the deuterated sample displaying no vibration which suggests the C=O stretch for the samples analyzed in H_2O interact with the solvent. *N*-acetyl-L-Glu there are two Raman split bands observed at ~1712 cm^{-1} and ~1647 cm^{-1}, with the deuterated sample displaying corresponding bands at ~1723 cm^{-1} and ~1635 cm^{-1}. In the ROA spectrum of *N*-acetyl-L-Glu in H_2O we

observe a positive band at ~1726 cm^{-1} and a negative weak band at ~1676 cm^{-1}, for the corresponding deuterated spectrum for *N*-acetyl-L-Glu there is a positive band at ~1711 cm^{-1} and a very weak band with negative sign at ~1672 cm^{-1}. These observations suggest that the acetylated amino acids have strong intermolecular interactions with the side chain which leads to an intense profile for the amide I vibration.

The amide II vibration is considered to be less sensitive to secondary structure than the amide I vibration and is not often used for structural analysis in Raman spectroscopy (Barron *et al*. 2000). The Raman spectra for all four amino acids, deuterated samples follow this typical trend where there is no strong amide II band observed.

CH_2 deformations are observed in the spectral region close to the amide II vibrations and with this in mind ROA bands observed for L-Glu analysed in D_2O, at ~1544 cm^{-1} and bands observed at ~1500 cm^{-1} and for L-Asp analyzed in H_2O and D_2O have been assigned to CH_2 deformations.

The amide III vibration is a complex vibration because it contains contributions from different bond stretches and deformations. For the Raman spectra of *N*-acetyl-L-Asp and *N*-acetyl-L-Glu the bands observed are very similar in the amide III region in H_2O and D_2O and therefore are not distinguishable. The ROA spectra of these two acetylated amino acids reveals a positive negative couplet in the amide III region and the non-actylated samples also display a positive and negative couple in the same region which do not differ significantly from the acetylated sampes in intensity or

position, however, there is a shift for *N*-acetyl-L-Glu and L-Glu of 30 cm^{-1}which we have attributed to the extra CH_2 group for these amino acids. We have therefore only one observable difference between both the acetylated and non-acetylated amino acids in this spectral region.

CH_3 deformations and CH_2 scissoring vibrations are generally observed around ~1420 cm^{-1} and for the Raman spectra of L-Asp and L-Glu in both D_2O and H_2O there is a band observed at ~1423 cm^{-1}, for the acetylated amino acids the bands observed in this region display a much lower intensity than the monoamino acids which suggests the acetylated amino acids interact with the solvent more in this part of the molecule. This therefore also appears to be a good Raman marker band for acetylation. For the ROA spectra for L-Asp and L-Glu there is a negative band at ~1434 and ~1424 cm^{-1}, respectively. For the acetylated amino acids this negatively signed band shifts downwards for *N*-acetyl-L-Asp to ~1407 cm^{-1} and upwards to ~1452 cm^{-1} for *N*-acetyl-L-Glu, whereas for the samples analyzed in D_2O *N*-acetyl-L-Asp further shifts to ~1380 cm^{-1} which suggests a higher level of H/D exchange. The intensities of this ROA band for the acetylated amino acids in H_2O are weaker when compared to the non acetylated samples indicating increased interaction with the solvent molecules, however, for the amino acids in D_2O all four samples display similar intensities. Thus, this ROA feature provides a good level of sensitivity for the distinction between the two acetylated amino acids where a downward shift is observed for *N*-acetyl-L-Asp.

5.4.3 Backbone skeletal stretch region.

A wide range of skeletal vibrations span the 1200 – 700 cm^{-1} spectral region and the band assignments are assigned cautiously for this reason. Density function theory (DFT) calculations help assign vibrations in this region, however, accurate DFT calculations are difficult to achieve for monoamino acids due to their high levels of conformational freedom as they lack rigidity without a peptide bond (Ruud & Thorvaldsen 2009). DFT calculations were performed on a single conformer each of *N*-acetyl-L-Asp and *N*-acetyl-L-Glu by Prof. T.J. Dines (University of Dundee), but both the computed Raman and ROA spectra were in poor agreement with the experimental results presented here. See supplementary information for a comparison of calculated and experimental spectra for *N*-acetyl-L-Asp and *N*-acetyl-L-Glu (Fig. S5.2). Further calculations need to be performed in order to improve the level of agreement with experimental, but this will require a more extensive consideration of conformational averaging for each species which is likely to be a time consuming process. Therefore, analysis of the spectra for these amino acids is restricted to empirical means in this chapter.

In the ROA spectrum for *N*-acetyl-L-Asp there are three bands at ~981 (+), 946 (-) and 886 (+) cm^{-1}, and for the ROA spectrum of *N*-acetyl-L-Glu these band are observed at 1014 (-), 895 (+) and 832 (-) cm^{-1}. We have assigned these bands to CC and CN stretching, which would agree with observations by Dhamelincourt & Ramirez, who observed three CC stretching bands at 850, 918 and 970 cm^{-1} in the IR spectra of L-Asp and L-Glu (Dhamelincourt & Ramirez 1991) and Kausar *et al*. who predicted these bands to be a combination of CC(CH_3) and CN amide stretching

modes (Kausar *et al*. 2009a). The observation of the opposite signs suggests that the vibrational modes with respect to the chiral centre have opposite stereochemistry, which presents us with another distinctive ROA marker band for differentiating between the two acetylated amino acids.

In the ROA spectrum for L-Asp there is a strong band with negative sign at 804 cm^{-1} which has shifted downwards for the deuterated sample to 777 cm^{-1}. This band is not observed in the ROA spectrum of *N*-acetyl-L-Asp or for L-Glu or *N*-acetyl-L-Glu. From the DFT calculations for *N*-acetyl-L-Asp carried out by Kausar *et al*., the authors predicted a band at 773 cm^{-1} and assigned this to C=O wagging modes, (Kausar *et al*. 2009a), we therefore, tentatively assign this same vibrational mode as the origin of the band observed for L-Asp at 804 cm^{-1} in the ROA spectrum.

5.4.4 Spectral Region below 700 cm^{-1}

Bands observed in this region are primarily due to CC=O, CC, CNC bending vibrations, OCO rocking motions, CO out-of-plane deformations, CC stretching, NC twisting motions, C=O wagging and CC deformations (Kausar *et al* 2009a). Vibrations below 700 cm^{-1} are difficult to interpret as the OH bending motions in H_2O interact with the analyte which can therefore present weak bands in this region (Kapitán *et al*. 2006). The Raman and ROA spectra for L-Asp and L-Glu follow this typical trend. However, in the ROA spectrum of *N*-acetyl-L-Glu there is a negative band at ~544 cm^{-1} and a positive band observed at ~485 cm^{-1} which we have assigned to C(CO) deformations, C=O rocking and CC deformations, which were predicted at 531 and 471 cm^{-1} by Kausar *et. al.* (Kausar *et al*. 2009a). For *N*-acetyl-

L-Asp there is a negative band at ~534 cm^{-1} and a positive band at ~471 cm^{-1} which we have assigned to C=O wagging, OH twisting and NH wagging motions which were predicted at ~526 and ~489 cm^{-1} by Kausar *et al*. (Kausar *et al*. 2009a). The deuterated samples display similar profiles, however, the intensity for the positive sign band at ~471 cm^{-1} of *N*-acetyl-L-Asp decreases in the deuterated spectrum which is indicative of these vibrations being influenced by the solvent for *N*-acetyl-L-Asp in H_2O.

5.5 Conclusions

We have shown here that ROA provides more detailed information about amino acid conformations than Raman spectra does and although some band assignments are still to be confirmed the overall assignments are characteristic of the structural elements present in these amino acids. For this study we looked for bands that were specific to each amino acid and also bands that showed site-specific acetylation. In the amide I region for the acetylated amino acids for the Raman and ROA spectra the split band observed in H_2O and D_2O samples provides a clear marker of acetylation. We have also shown that vibrational mixing potentially occurs in acetylated amino acids for the amide I vibration where the acetyl group interacts with the amide group. In the region between 800-1000 cm^{-1} in the ROA spectra of *N*-acetyl-L-Asp and *N*-acetyl-L-Glu there exist several bands of opposite sign which originate from CC and CN stretching. This suggests that the vibrational modes generating ROA bands in this spectral region experience opposite stereochemistry and as this is not observed for the nonacetylated samples this give us a good site-specific marker of acetylation. From this study we have ascertained that ROA can provide more structural

information than Raman and more specifically allows us to detect modification of amino acids which is essential for future studies relating to transcriptional regulation and cancer therapies.

5.6 References

Atkins P.W. & Barron L.D. (1969) *Molecular Physics* **16**: 452-456

Ashton L.A., Dusting J., Imomoh E., Balabani S. & Blanch E.W. (2009) *BIOPHYS. J.* **96**: 4231–4236

Audtoe J.V., Wade L. & Olson E.J. (1966) *J. Neurochem.* **13**: 1149-1155

Bandekar J. (1992) *J. Biochim. Biophys. Acta* **1120**: 123-143

Bárány M., Spigos D.G., Mok E., Venkatasubramanian P.N., Wilbur A.C. & Langer B.G. (1987) *Magn. Reson. Imaging* **5**(5): 393-398

Barron L.D., Bogaard M.P. & Buckingham A.D. (1972) *J. Am. Chem. Soc.* **95**: 603-605

Bartalini G., Margollicci M., Balestri P., Farnetani M.A., Cioni M. & Fois A. (1992) *Childs Nerv Syst.* **8**(8): 468-470

Blanch E.W., McColl I.H., Hecht L., Nielsen K. & Barron L.D (2004) *Vib. Spectrosc.* **35**:87–92

Bowen M. & Brunger A.T. (2006) *PNAS.* **103**: 8378-8383

Burgess A.W. & Scheraga H.A. (1973) *Biopolymers* **12**: 2177-2183

Briggs S.D., Xiao T., Sun Z.W. & Caldwell J.A. (2002) *Nature* **418**: 498

Brown R. & Strathdee G. (2002) *Trends Mol. Med.* **8**: S43–S48

Chen X.-G., Schweitzer-Stenner R., Asher S.A., Mirkin N. & Krimm S. (1995) *J. Phys. Chem.* **99**: 3074-3083

Chin W., Piuzzi F., Dimicoli I. & Mons M. (2006) *Phys. Chem. Chem. Phys.* **8**: 1033-1048

Chitra M., Sharma S.M., Karmakar S. & Sikka S.K. (2003) *Physica B* **339**: 23-30

Choi J.H., Hahn S. & Cho M. (2005) *Int. J. Quantum Chem.* **104**: 616-634

Danielsen E.R. & Ross B. (1999) *Magnetic Resonance Spectroscopy Diagnosis of Neurological Diseases* Marcel Dekker: New York

Deng Z., Polavarapu P.L., Ford S.J. Hecht L., Barron L.D., Ewing C.S. & Jalkanen K. (1996) *J. Phys. Chem.* **100**: 2025-2034

Dhamelincourt P. & Ramirez F.J. (1991) *J. Raman Spectrosc.* **22:**577-582

Divry P. & Mathieu M. (1989) *Am. J. Med. Genet.* 32(4): 550-551

Dougherty D.A. (2006). *Modern Physical Organic Chemistry*. Sausalito, CA: University Science Books

Edler J. & Hamm P. *J.* (2002) *Chem. Phys.* **117**: 2415-2424

Fan T.W., Higashi R.M., Lane A.N. & Jardetzky O. (1986) *Biochim Biophys Acta.* **882**(2): 154-67

Fang C., Wang J., Charnley A.K., Barber-Armstrong W., Smith A.B., Decatur S.M. & Hochstrasser R.M. (2003) *Chem. Phys. Lett.* **2003**: 586-592

Garner W.H. & Spector A. (1978) *PNAS* **75**: 3618-3620

Grdadolnik J., Grdadolnik S.G. & Avbelj F. (2008) *J. Phys. Chem.* **112**: 2712-2718

Hudson D.B., Valcana T., Bean G.S. & Timiras P.S. (1976) *Neurochem. Res.* **1**: 73-81

Hug W. & Hangartner G. (1999) *J. Raman Spectrosc.* **30**: 841-852

Johnstone R.W. (2002) *Nat. Rev. Drug Discov.* **1**: 287–299

Jungmann R.A. & Schweppe J.S (1972) *J. Biol. Chem.* **247**:5535-5542

Kausar N., Alexander B.D., Dines T.J., Withnall R. & Chowdhry B.Z. (2009A) *J. Raman Spectrosc.* **40**(6): 670-678

Krimm S. & Bandekar J. (1986) *Adv. Protein Chem.* **38**: 181-364

Libby P.R. (1972) *Biochem. J.* **130**: 663-669

Maeda S., Eguchi S. & Sasaki H. (1958) *J. Home Econ.* **9**: 163-167

Masman M.F., Zamora M.A., Rodríguez A.M., Fidanza2 N.G. Peruchena N.M., Matalon R., Michals K., Sebesta D., Deanching M., Gashkoff P. & Casanova J (1988) *Am. J. Med. Genet.* **29**(2):463-471

Matsres P.M., Bada J.L. & Zigler Jr. J.S. (1978) *PNAS.* **75**:1204-1208

Maxfield F.R., Leach S.J., Stimson E.R., Powers S.P. & Scheraga H.A. **(**1979) *Biopolymers* **18**: 2507-2521

Mirkin N. & Krimm S. (1991) *J. Am. Chem.* **113**: 9742-9747

Mirkin N. & Krimm S. (1996) *J. Am. Chem.* **377**: 219-234

Mizzen C.A. & Allis C.D. (2000) *Science* **289**: 2290–2291

Miyazawa T., Shimanouchi T. & Mizushima S.I. (1956) *J. Chem. Phys.***24**: 408-418

Narita K. (1958) *Biochim. Biophys. Acta* **28**: 184-191

Pawlukojc A., Leciejewicz J., Tomkinson J. & Parker S.F. (2002) *Spectrochimica. Acta, Part A* **58**: 2897-2904

Pozo Ramajo A., Petty S.A. & Volk M. (2006) *Chem. Phys.* **323**: 11-20

Schweitzer-Stenner R., Sieler G., Mirkin N. & Krimm S. (1998) *J. Phys. Chem.* **102**: 118-127

Shigeto S. & DlottD.D. (2007) *Chem. Phys. Lett.* **447**: 134-139

Székely Z., Konya Z., Becskei A., Goldring W.P.D., Perczel A., Penke B., Molnár J., Michejda C.J., Aszalós A. & Csizmadia I.G., (1996) *J. Mol. Struct. (Theochem)* **367**: 159

Taday P.F., Bradley I.V. & Arnone (2003) *J. Biol. Phys.* **29**: 109-115

Takekiyo T., Imai T., Kato M. & Taniguchi Y. (2004) *Biopolymers* **73**: 283-290

Tallan H.H., Moore S. & Stein W.H. (1956) *J Biol Chem.* **219**(1): 257-264

Tallan H.H. (1957) *J Biol. Chem*. (1957) **224**(1): 41-45

Torii H., Tasumi T., Kanazawa T. & Tasumi M. (1998) *J. Phys. Chem.* **102**: 309-314

Wang Y., Purello R., Georgiou S. & Spiro T.G. (1991) *J. Am. Chem. Soc.* **113**: 6370-6400

Watson T.M. & Hirst J.D. (2002) *J. Phys. Chem.* **106**: 7858-7867

Watson T.M. & Hirst J.D. (2004) *Phys. Chem. Chem. Phys.* **6**: 998-1005

.

5.7 Supplementary Information

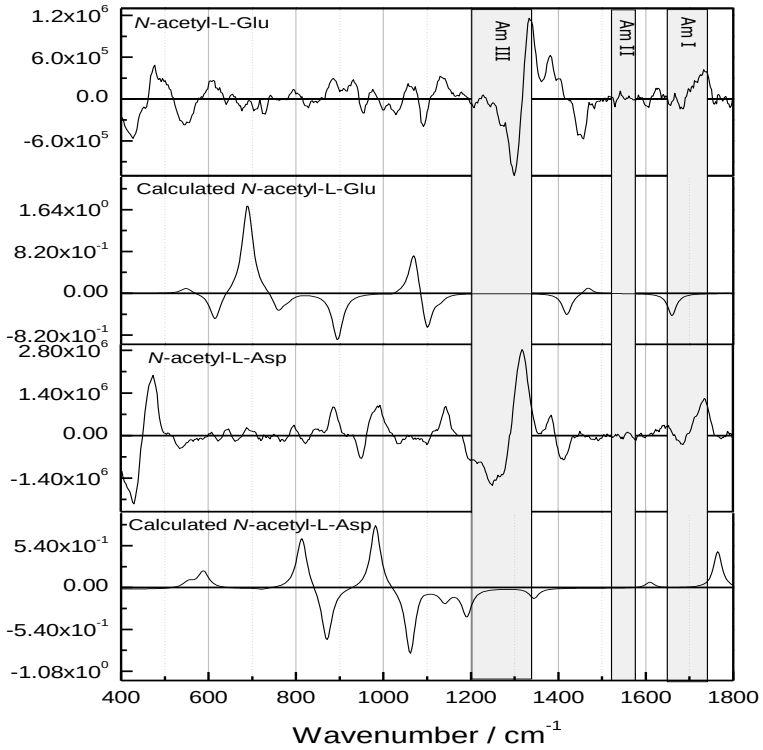


Fig. S5.1 Calculated spectra and experimental spectra of *N*-acetyl-L-Asp and *N*-acetyl-L-Glu which demonstrates the difficulty of obtaining a good level of agreement between DFT calculations and experimental data for monoamino acids. It is clearly apparent that there is a poor level of agreement due to the absence of conformational averaging in the calculations, particularly for the important extended amide III region.

Chapter 6

6.1 Conclusion

This thesis presents a combination of experimental and computational studies in the advancement of Raman spectroscopies. The aim of this research was to work towards optimization of experimental conditions for colloidal surface-enhanced Raman scattering (SERS) detection of pharmaceutical compounds and to investigate consequences of acetylation in biomedically important amino acids using Raman optical activity (ROA).

As part of the process of optimization of experimental conditions for SERS, samples were first measured to determine whether the preparation process had any effect on the SERS measurements (see chapter 2). The initial studies revealed that the time taken from preparation of the final solution to the measurement of SERS for propranolol was important as some time was required for aggregation, although this was only observed when using citrate reduction of $AgNO_3$ and a chloride aggregating agent as all the other colloidal systems aggregated immediately regardless of the aggregating agent used. It was also noted that if the measurement was delayed too long then the colloid would drop out of the solution, particularly when using citrate reduction of $AuCl_4$ with a sulfate aggregating agent. This was an important step in the optimization process and great care was taken to optimize the sample protocols before optimizing the experimental conditions and unless these steps are checked first then a sub-optimal or even zero SERS signal may be observed.

To explore the search space for the experimental design the parameters chosen reflected the parameters that have the most effect on the SERS signal, however, the concentration of the analyte was fixed as it was felt that it was important to determine the experimental conditions before moving on to varying the analyte concentration. Once the search space was determined 315 experiments out of the full search space of 7,785 were performed to establish that the MOEA could find the optimal result from the fitness function for enhancement. Once this was established the replicate experimental data was input into the MOEA and the parameters of the MOEA were developed in order to find the best values which would allow the MOEA to find acceptable solutions to the MOP. From the results observed it was clear that citrate reduction of gold and a chloride aggregating agent exceeded the other colloidal sols in producing both a suitable and acceptable enhancement and reproducibility factor for colloidal SERS of propranolol. By weighting both objectives equally it gave the operator the opportunity to decide whether reproducibility of the SERS signal was more important than absolute enhancement or vice versa. For this research problem a combination of both a good enhancement factor and reproducibility were important, and thus the results observed in the middle of the Pareto front allowed for a suitable compromise of both objectives and therefore an acceptable solution to the problem posed to the MOEA. By performing a practical LOD assay we were able to establish whether the results we observed from the MOEA led to advancement in the resultant SERS and this was shown in the detection limit achieved of 2.6 ng/mL, which was 24.5 times lower than found in previously published literature. This result clearly demonstrated that application of the MOEA to SERS spectra not only efficiently identified optimal experimental

parameters in less than 1% of the time required to perform an exhaustive search, but also led to a significant improvement in detection limits.

We then extended this technique to investigate the detection of propranolol in biofluids using the experimental conditions that were observed in the middle of the Pareto front. Vital medical information is often secreted through biofluids and a reliable and rapid technique is required to meet patient and commercial demand. During the initial studies of sensing of propranolol in plasma it was observed that the SERS signal of propranolol was not detected, however, upon further investigation it was realised that the colloid was no longer viable and a new batch was prepared. Detecting the characteristic bands of propranolol at concentrations lower than 130 ng/mL (8.79×10^{-6} M) proved difficult, however, this was a consideration we had already taken into account as blood plasma is a far more complex solvent than the water used in the MOEA development and with the added difficulty of propranolol binding to glycoproteins and albumin present in plasma effectively reducing the concentration of the free drug in the sample (Alvan *et al*. 1983). However, the detection limit achieved was still within physiological concentrations that have previously been published. Optimization of the experimental conditions for SERS of propranolol in biofluids requires further work and we have demonstrated that optimizing experimental conditions for an analyte in a different solvent will not always lead to the same behaviour for dH_2O. Despite only being able to measure propranolol in plasma at a higher concentration than in the preceding chapter with water this work still serves as a proof-of-concept for this approach and with the use of a separation technique to determine the effective free drug in the presence of

plasma and with further optimization on plasma samples we believe that the MOEA for experimental design is a significant move towards the optimization of colloidal SERS detection of pharmaceuticals in biofluids.

The combination of Raman and ROA spectroscopies is more beneficial than using either of the techniques in isolation especially when investigating chemical modifications of amino acids. Such an approach is important for understanding both the structures and conformational dynamics of amino acids and short peptides subject to chemical modifications. Both Raman and ROA spectra exhibit small differences that are specific to modification, however chirality can only be determined by ROA. This is illustrated in the spectra obtained for the DAAPs where it is clear that the Raman spectra cannot distinguish between the constitutional isomers whereas ROA is sensitive to these structural differences. However, the Raman spectra do show bands specific to acetylation when compared to non-acetylated mono and diamino acids. ROA does provide more structural information for site-specific acetylation, for example in the spectra for α- and β-*N*-acetyl-L-Asp-L-Glu marker bands for the acetyl group are observed, however, for β-*N*-acetyl-L-Asp-L-Glu the band is broader which suggests there is overlap from the carboxyl band at 1442 cm^{-1} whereas the carboxyl band for α-*N*-acetyl-L-Asp-L-Glu is observed separately at 1421 cm^{-1}. During the study of the monoamino acids it was observed that *N*-protected amino acids have stronger intramolecular interactions with the side chain which leads to a strong profile for the amide I vibration, for the diamino acids there is more interaction with the solvent and the band structure in that region is not as rich. DFT calculations performed by collaborators provided

predictions based on the electron density of the molecules investigated with an important factor being to predict the lowest energy conformers in order to predict their wavefunctions (page 112). However, there can be problems predicting the lowest energy conformers for molecules with a relatively flat energy landscape and calculations do not always agree with experimental results, although they can still be extremely useful for assigning spectral features. Reasonable agreement between the predicted spectra and the experimental results were observed for the DAAPs. The DFT calculations for the monoamino acids were not in good agreement, however this is a known consideration for monoamino acids as they can adopt a far greater range of conformations making their computational modelling a difficult challenge (Ruud & Thorvaldsen 2009). Nevertheless the spectra measured here will serve as a robust test for training future ROA calculations.

6.2 Future Work

Colloidal SERS is still a widely used detection technique for obtaining an analyte fingerprint in solution and although there are recent advances in producing stable SERS-active plasmonic materials, the process is complicated with expensive methodologies and instrumentation required, whereas the colloidal sols used in this study are inexpensive, easy to prepare and label free. Colloidal SERS is a complex vibrational technique and when investigating samples it is important that the optimal experimental conditions for the analyte are chosen. Any scientific problem can produce a number of solutions that are much smaller than the number of experiments required to identify them and it is considerably faster and more efficient to search a

combinatorial landscape *in silico*. This can be done in a step-by-step process; however, for a MOP it is not ideal and there is no indication what effect changing one parameter at a time is having on the other parameters? Thus, the introduction of a MOEA, which takes into account all parameters, can search a landscape quickly and reduce the number of experiments required in finding acceptable solutions to a MOP. As this thesis has shown by optimizing the experimental conditions for an analyte a LOD has been found that is considerably lower than previously published values. It is therefore important, and a significant advancement, for colloidal SERS that the experimental conditions are optimized and all parameters are considered simultaneously. Future studies on a wide range of biologically relevant analytes can be investigated using this method where further parameters can be introduced into the MOEA, such as additional aggregating agents, additional metal types and the concentration of the analyte. Although aggregation of colloids is difficult to stabilize in solution due to the dynamics of the particles, colloidal SERS is still widely used and optimizing the experimental conditions is the way forward for this extremely useful technique. Another complication of colloidal SERS is the stability of the colloid over time, which was demonstrated during the studies for plasma spiked with propranolol. It is therefore recommended that the colloids are checked before initial use by measuring a known sample to determine whether the colloid is still viable. Although optimizing experimental conditions for an analyte has shown a significant reduction in concentrations required for detection it has to be noted that changing the solvent can have a significant effect on the SERS signal and therefore future optimization must take into account the solvent as well as the analyte (Chang & Chen 1996). A separation technique is also required to determine the effective concentration of free drug in the presence of a biofluid. Further development of this

MOEA approach will evolve new experimental protocols for a range of analytes. This approach can be extended to surface-enhanced Raman optical activity and other vibrational techniques that require solution structure or detection of analytes, particularly in biofluids to meet commercial and patient demand.

The identification of chemical modification in amino acids is an essential step forward for biological studies in relation to gene control, cancer therapies and transcriptional regulation (Briggs *et al*. 2002 Mizzen & Allis 2000 Brown & Strathdee 2002; Johnstone 2002). New approaches are required to detect and characterize chemical modifications such as acetylation, cyclization, phosphorylation, methylation and glycosylation to name but a few. Conventional, and particularly atomic resolution, methods are currently exploited for this purpose and will continue to be so, however the success of ROA has extended its use in many laboratories giving it an important role to play in providing new and detailed structural information. As the role of ROA advances so will its versatility for the analysis of structure and behaviour of most if not all biological molecules through further instrument developments and machine learning. ROA spectra provide a great deal of information concerning differences in structure, conformation and modification and with the combination of DFT calculations these spectral features can be assigned. However, calculations of ROA spectra are complex due to the computational complexity of molecular properties required for predicting a ROA spectrum and the limited experimental activity using this spectroscopic tool (Crawford 2006). As ROA increases in use and advances are made in linear and nonlinear response theories there are an increasing number of *ab initio* programs that

exist in order to calculate ROA spectra (Ruud & Thorvaldsen 2009). With further advancements in these *ab initio* programs the agreement between DFT calculations and experimental results will improve.

6.3 References

Alvan G., Bergstrom K., Iselius L. & Perdersen N. (1983) *Br. J. Clin. Pharmacol.* **25**: 437-441

Briggs S.D., Xiao T., Sun Z.W. & Caldwell J.A. (2002) *Nature* **418**: 498

Brown R. & Strathdee G. (2002) *Trends Mol. Med.* **8**: S43–S48

Chang Y.C. & Chen T.T. (1996) *Chinese Journal of Physics* **34(6)**: 1352-1362

Crawford T.D. (2006) *Theor. Chem. Acc.* **115**: 227-245

Mizzen C.A. & Allis C.D. (2000) *Science.***289**: 2290–2291

Ruud K. & Thorvaldsen A.J. (2009) *Chirality* **21**(1E): E54-E67