DEVELOPMENT OF NANO-BASED BIOSENSORS FOR THE DETECTION AND DIFFERENTIATION OF BACTERIAL PATHOGENS USING SURFACE ENHANCED RAMAN SPECTROSCOPY

by

Omar Eduardo Rivera

(Under the Direction of Richard A. Dluhy)

ABSTRACT

Rapid and sensitive detection of pathogens is an essential strategy for intervention of a possible disease outbreak. The unique ability of SERS to provide analyte specific response has allowed for the development of this method for rapid biomedical diagnosis needed for routine analyses of different species and strains of pathogens.

The specific agents we examined in the first part of this project are lipophilic extracts containing mycolic acids isolated from Tuberculosis (MTB) and non-tuberculosis (NTM) strains using chromatography, mass spectrometry (MS), nuclear magnetic resonance (NMR), and Raman spectroscopy. Surface-enhanced Raman (SERS) spectra were obtained from the mycolic acids extracted from the bacterial cell envelopes of the MTB or NTM mycobacterial species.

The Raman spectral profiles were used to develop a classification method based on chemometrics for identification of the mycobacterial species. Multivariate statistical analysis methods, including principal component analysis (PCA), hierarchical cluster analysis (HCA),and partial least squares discriminant analysis (PLS-DA) of the SERS spectra enabled differentiation of NTM mycobacteria from one another with 100% accuracy. These methods are also sensitive enough to differentiate clinically-isolated MTB strains that differed only by the presence or absence of a single extracytoplasmic sigma factor.

We examined as well *Mycoplasma pneumoniae* which is a major cause of respiratory disease in humans and accounts for as much as 20% of all community-acquired pneumonia. There is a critical need to develop a new platform for mycoplasma detection that has high sensitivity, specificity, and expediency. Here we report three different layer-by-layer (LBL) encapsulation procedures of *M. pneumoniae* and mycoplasma commensals cells with Ag nanoparticles in polyelectrolyte matrixes. We evaluated nanoparticle encapsulated mycoplasma cells as a platform for the differentiation of *M. pneumoniae* mycoplasma commensal strains using surface enhanced Raman scattering (SERS) combined with multivariate statistical analysis. Since the LBL-SERS has inherent biochemical specificity, we analyzed a panel of 11 human commensal from a pathogen strain of mycoplasma to demonstrate that this platform could distinguish M129 from its clinically relevant phylogenetic relatives.. The feature selection information was used to perform PLS-DA and SVM-DA models. These results suggest that SERS along with multivariate statistics can be used as an accurate and sensitive method for species and strain discrimination in mycobacteria and mycoplasma.

INDEX WORDS: *Mycobacterium tuberculosis*, mycolic acids, *Mycoplasma pneumoniae*, commensal mycoplasma strains, Surface-enhanced Raman Spectroscopy (SERS), Layer-by-layer technique, multivariate statistical methods modeling, chemometrics.

DEVELOPMENT OF NANO-BASED BIOSENSORS FOR THE DETECTION AND DIFFERENTIATION OF BACTERIAL PATHOGENS USING SURFACE ENHANCED RAMAN SPECTROSCOPY

by

Omar Eduardo Rivera

B.S.,University of Puerto Rico-Mayagüez, 2006

M.Sc.,University of Puerto Rico-Mayagüez, 2009

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2014

DEVELOPMENT OF NANO-BASED BIOSENSORS FOR THE DETECTION AND DIFFERENTIATION OF BACTERIAL PATHOGENS USING SURFACE ENHANCED RAMAN SPECTROSCOPY

by

Omar Eduardo Rivera

Major Professor:	Richard A. Dluhy
Committee:	Jason Locklin
	Jeffrey Urbauer

Electronic Version Approved:

Julie Coffield
Interim Dean of the Graduate School
The University of Georgia
December 2014

ACKNOWLEDGEMENTS

First of all I would like to thank my parents Doris and Edgardo for all your love and support. Without your words of love, wisdom and motivation I wouldn't have been really hard to finish this degree. I would also like to thank my supporting team, my stepmother Acacia, my grandmother Myrna and my four brothers Rafael, Eric, Guillermo and Juan. My significant other Natalie, ever since you came into my life you have turned it around. Without your unconditional love and support I don't know if I would've been able to continue through the tough times.

I would like thank to my advisor Dr. Richard Dluhy, for all his support, supervision and guidance. Throughout these four years in his research lab I have learned much more than just being a research scientist. I would like to acknowledge as well my committee members for Dr. Urbauer and Dr. Locklin for their guidance, suggestions and constructive criticism that helped me improve my research work.

In addition to my committee and family I would like to thank Sharon Martin and Dr. Pierre Negri for helping me getting my foot in the door when I started to work at the Dluhy lab and for being such great friends and colleagues. I would like to thank as well Dr. Duncan Krause, Dr. Sue Hennigan and Dr. Kelley C. Henderson for all your help with the mycoplasma project. I would like to acknowledge as well Dr. Fred Quinn and Dr. Russ Karls for all your help with the mycobacteria project.

TABLE OF CONTENTS

Page

CHAPTER 3

CHAPTER 4

CHAPTER 5

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

Literature Review

The purpose of this chapter is to introduce the reader in retrospect of the history and background of the main families of bacteria that we worked on.

1.1 Mycobacterium Tuberculosis: Fatty acid cell components, clinical syndrome and current treatment diagnostics

The bacterium *Mycobacterium tuberculosis*(TB), the primary causative agent of tuberculosis, infects one-third of the world's population.[1] A challenge in treating *M. Tuberculosis* infection is the evolution of strains that are resistant to first and second line drugs that are key components of TB treatment.[2] Infection with TB often involves decades of persistence in the host, during which time the organism sets up supply lines for importing host nutrients. Robert Koch identified *Mycobacterium tuberculosis* sometimes called the tubercle bacillus, as the cause of tuberculosis. This rod shaped obligate aerobe has an unusual waxy cell wall which affects many of its properties, including its slow growth and unusual staining properties. Special staining techniques, called acid fast stains are used to detect its cells in clinical specimens. Its waxy wall also allows it to survive prolonged drying: *M. tuberculosis* cells that enter the air when a tuberculosis patient coughs can remain infectious for up to eight months. The mycolic acid components of the waxy wall also add to the microorganism's disease causing potential-they protect *M. tuberculosis* against the potentially lytic enzymes and oxidants within the phagocytes that engulf them. Iron is an essential micronutrient for pathogenic bacteria. Chelation therapy is used to treat conditions that arise due to undesirable or excess quantities within the body. Treatment for the removal of excess iron is the most common type of chelation therapy and is routinely

carried out for the treatment of acute iron poisoning and chronic iron overload such as hemochromatosis.[3]

To acquire iron from their environment, mycobacteria develop high affinity iron scavenging molecules via the synthesis of two types of siderophores: mycobactins, which as lipophilic ligands that remain cell wall-associated, and carboximycobactins that are hydrophilic which are released to the extracellular medium.[4 5] Both siderophore types are hexadentate chelators that include identical hydroxamate binding sites and differ only in lateral chain characteristics, such that a lipophilic character is conferred on mycobactins and a hydrophilic character on carboxymycobactins.[4 6] In some mycobacteria, mycobactins and carboxymycobactins possess identical iron (III) binding sites. The lipophilic character of mycobactins make them good candidates to enter cells and deplete intracellular iron pools of infected host cells.[7]

Restricting the availability of iron is an important strategy for defense against bacterial infection.[8 9] *Mycobacterium tuberculosis* survives within the phagosomes of macrophages; consequently, iron acquisition is particularly difficult due to the phagosomal membrane acting as an additional barrier for iron access.[10] The purpose of this project is not to describe the mechanism in which lipophilic mycobactins transport iron through the membrane barrier in which *M. tuberculosis* is harbored in the phagosome inside of a host cell macrophage.[11] The way in which mycobacteria acquire endogenous iron is through the accumulation on mycobactin-metal complexes with selectivity in macrophage lipid droplets that come in contact with the phagosome[12] as shown in Figure 1.1.[13,14] Many groups, view the process of iron acquisition as one of the representative possible pathways that can be successfully targeted by novel therapeutic tools.[15 16]

Immunological mechanisms against *M. tuberculosis* have been mainly investigated in murine models. Often a low dose of bacilli are introduced through an aerosol, to resemble human infection. The immune resistance to *M. tuberculosis* relies mainly on cellular immunity, which is stimulated by proteins secreted by the bacilli. Mycobacteria produce a wide variety of glycolipids that are located in the outermost layers of the cell as shown in Figure 1.2. The unique cell wall of *M. tuberculosis* contributes to the success of this pathogen. The cell wall contains three covalently-linked layers outside the cytoplasmic membrane: a peptidoglycan matrix, an arabinogalactan matrix, and an outer membrane consisting of long lipids, such as the mycolic acids. Their structures are based on very characteristic methyl-branched long chains and alcohols as shown in Figure 1.3.

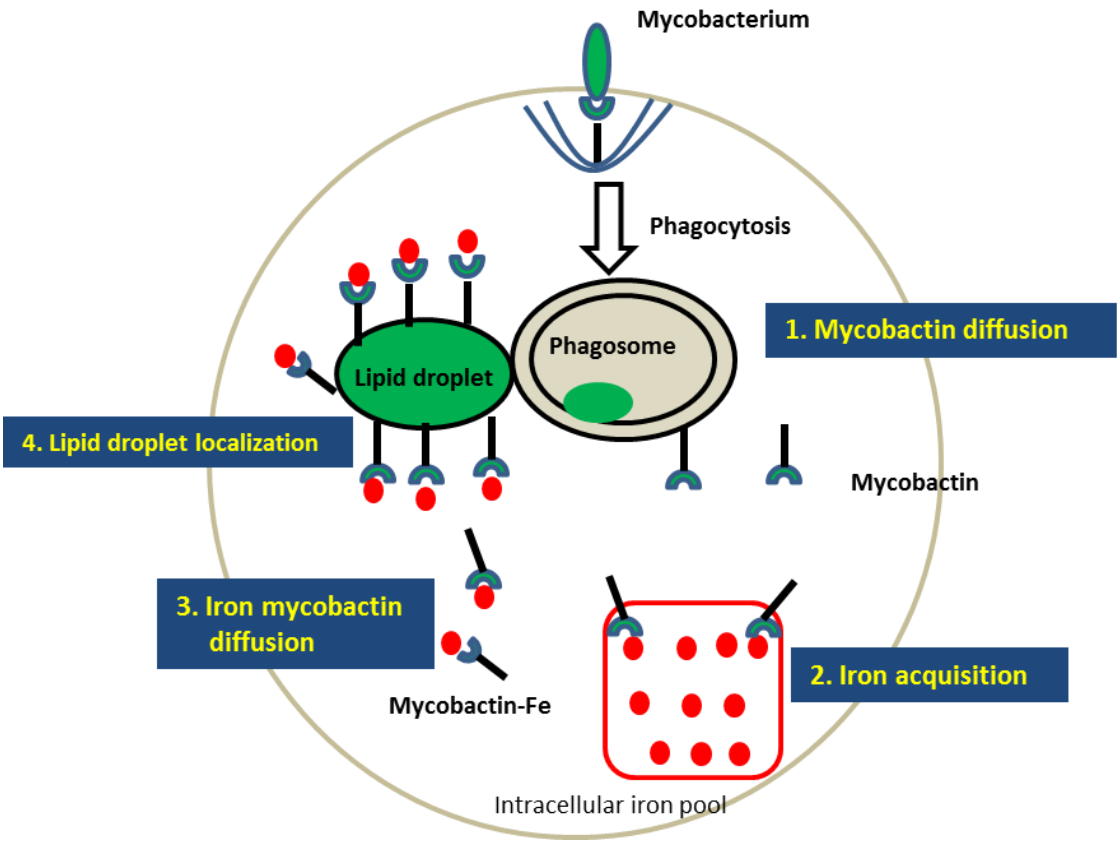

Figure 1.1 Mycobactin iron acquisition pathway for macrophage-niched mycobacteria .

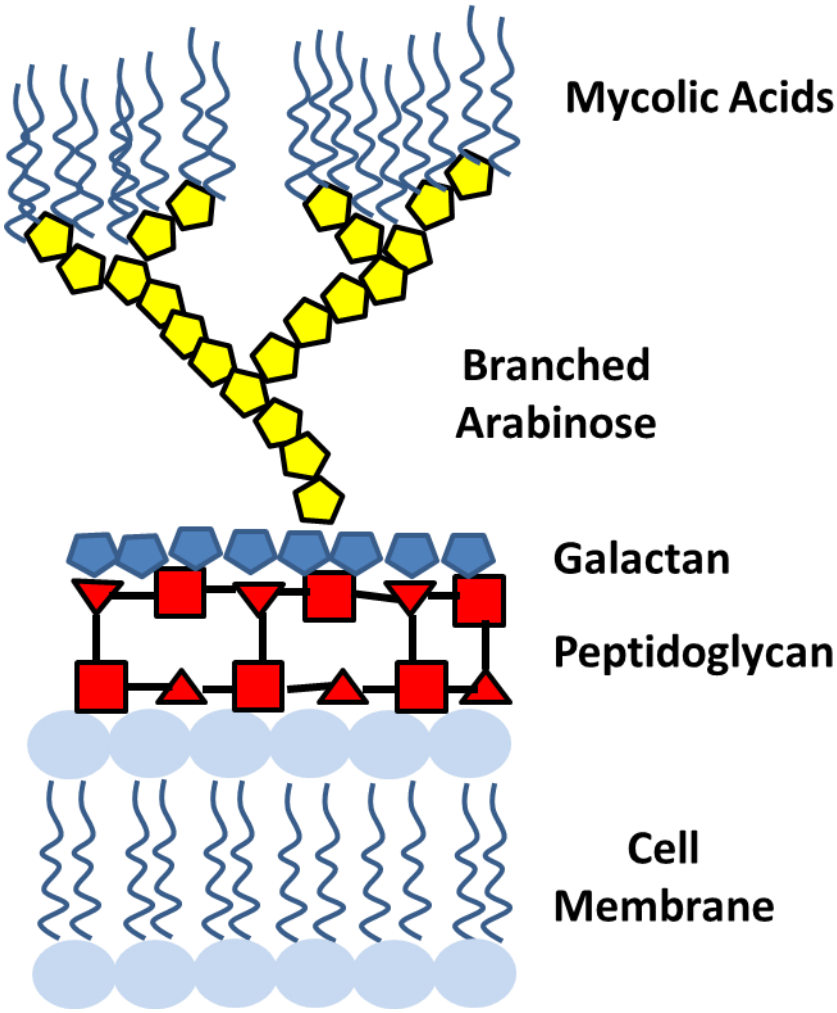

Figure 1.2 Illustration of the cell wall of *Mycobacteria tuberculosis.*

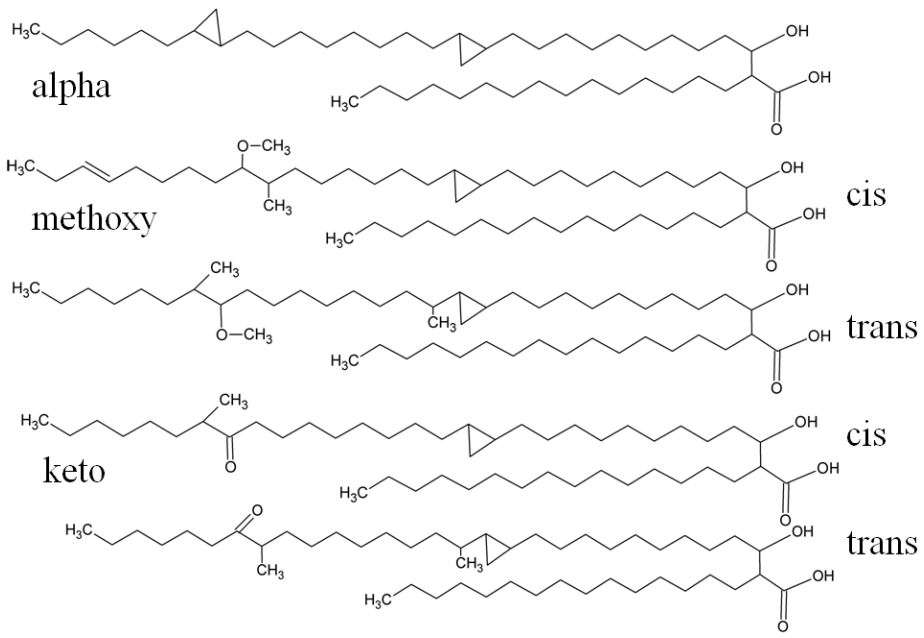

Figure 1.3 Mycolic acid sub-structures commonly found in mycobacteria.

Mycolic acids are long-chain α-alkyl-branched, β-hydroxylated fatty acids present in mycobacteria and related micro-organisms.[17] The extremely-low fluidity of the inner leaflet of the outer membrane contributes to the exceptionally-low permeability of the mycobacterial cell wall to polar agents and natural resistance of mycobacteria to antibiotics effective against other bacteria.[18] [19] Mycobacterial lipids are known to facilitate persistence during infection. The biosynthesis of mycolic acids has been investigated for a number of reasons, but primarily because of the presumption that enzymes involved in the synthesis of these lipids are attractive targets for the development of novel therapeutic agents. Two of the frontline *M. tuberculosis* drugs (e.g Ethambutol and Isoniazid) target the mycobacterial cell wall. Isoniazid directly impacts expression of mycolic acids. Ethambutol targets arabinosyl transferases required for the arabinogalactan synthesis.[20]

The first MA was isolated from the tubercle bacilli in 1929 by Anderson. The name mycolic acid was proposed in 1935 for a portion of the lipid fraction from *M. tuberculosis* by Stodola. Structural characterization followed in 1950 by Asselineau.

Mycolic acids are described as a group of long chain α-alkyl, β-hydroxy fatty acids. MA has a largely asymmetric structure around the hydrophilic headgroup. They have a general structure of R1-CH(OH)-CH(R2)-COOH. R1 is a meromycolate chain (50-60 carbons) and R2 is a shorter α-branch (22-26 carbons). They are the major constituents of the inner leaflet of the lipid bilayer of the mycobacterial cell wall, where they form an effective impermeable barrier to protect the mycobacteria from antimicrobial agents.[21] In mycobacterial cells, MA are predominantly covalently bound to the arabinogalactan polysaccharide. Some are found as trehalose monomycolates (TMM) and dimycolates (TDM).[22] [23] TMM is involved in the transfer of MA into the cell wall.

α-Branched MA as used in adjuvant emulsions may expose extensive hydrophobic domains onto which proteins might adsorb. Wedge shaped hydrophobic domains are formed because of the large ratio of the areas between the hydrophobic chains (folded up) and the hydrophilic head group. This folding up of the MA prevents formation of cylindrical micelles, which are unable to absorb proteins. Some have suggested through the measurement of limiting molecular areas, MA can be presented as double or triple chain molecules.[24] The folding of the mycolic acids might be important for recognition by CD1 molecules. The hydrophilic headgroup is then presented to T cells to result in various immune functions including antibody formation to MA.

We are discussing tuberculosis because it is usually transmitted by respiratory droplets and the lungs are most often affected. Primary infection usually begins from an infected person's respiratory secretions. Primary infections produces tubercles or granulomas that become calcified (Ghon complexes) in patients who overcome infection. Bacili in old tubercles may escape to cause a secondary, or reactivation, infection. The clinical syndrome includes fever, fatigue, weight loss, and a chronic cough. There is a vaccine (BCG), but it's not reliable. The United States relies on a tuberculin skin testing program. Although usually treatable with current antibiotic therapies, inadequate means of *M. tuberculosis* diagnosis remains one of the major obstacles to the global control of this disease, especially in resource limited areas, and is on the increase even in countries where the disease had previously been controlled.[25] The estimated number of new TB cases worldwide was 9.4 million in 2009, the highest number ever recorded.[26]

Unfortunately, the diagnosis of MTB still relies on sub-optimal methods that-in some areas of the world-are not far removed from the methods Robert Koch used to discover *M.*

tuberculosis in 1882. For example sputum smear microscopy remains the primary means of *M. tuberculosis* diagnosis in developing countries, yet smear microscopy is insensitive as it only detects ~ 50% of all active cases of TB.[27] Culturing of *M. tuberculosis* bacilli is considered the gold standard in TB diagnostics, and is more sensitive than smear microscopy; however *M. tuberculosis* is an extremely slow growing bacterium and test results may not be available for weeks, if not months, placing a heavy burden on the patient and potentially providing a continuing source of infection in the community.

Molecular-based methods of TB diagnosis have been developed, but each method suffers from one or more drawbacks that limit its applicability and usefulness.[28] Because *M. tuberculosis* infections produce a primarily cellular immune response, serological tests are considered to be of minimal diagnostic value.[29] Assays based on phage amplification or detection of interferon γ have been developed, but these methods suffer from i) a lack of sensitivity/specificity, ii) a significant delay in obtaining results due to sometimes long supply chains , iii) a need for significant personnel training and first world equipment, iv) the use of needles or other invasive collection tools , or v) the need for high BSL-2+ or BSL-3 laboratory facilities.[30] Test based on amplification of MTB DNA have been reported, and the most promising PCR-based commercial test (GeneXpert),[31,32] received CE certification in 2009. However this test has been criticized, and faces obstacles of logistics, sensitivity, and cost.[33] Notwithstanding the progress made to date, a simple, non-invasive, reliable, rapid, inexpensive, point-of-care test for active tuberculosis that can compete with conventional smear microscopy is not available.

1.2 Historic Overview of Mycoplasmas: current treatments and clinical applications

Artificial classification schemes for prokaryotes are still being use. The most popular of these is the one published in Bergey's Manual under auspices of the American Society for Microbiology which takes a highly practical approach to classifying bacteria. According to the manual bacteria are separated into four different divisions, the mycoplasma are separated into their division due to their lack of a rigid cell wall.

Mycoplasmas share a number of properties: all are parasites of humans, animals, or plants; almost all obligate fermenters (they ferment in the presence of oxygen). Mycoplasmas, the smallest self-replicating organisms, are parasitic species that lack a cell wall and periplasmic space, have reduced genomes, and limited metabolic activity.[34-36] *Mycoplasma pneumoniae* cells have an elongated shape that is approximately 1-2 μm in length and 0.1–0.2 μm in width as seen on Figure 1.4.[37] The layer of carbohydrate located outside of the cytoplasmic membrane is probably responsible for maintaining such shapes. The extremely small cell size means they are incapable of being examined by light microscopy; a stereomicroscope is required for viewing the morphology of *M. pneumoniae* colonies, which are usually less than 100 μm in length.[34] Unlike other prokaryotes mycoplasmas have sterols in their cytoplasmic membrane, and these lipids strengthen the membrane slightly, and possess more genes that encode for membrane lipoprotein variations than other mycoplasmas.[35] The inability to synthesize a peptidoglycan cell wall is due to the absence of genes encoding its formation and results in an increased importance in maintenance of osmotic stability to avoid desiccation. The lack of a cell wall also calls for increased support of the cell membrane, which includes a rigid cytoskeleton composed of an intricate protein network and, potentially, an extracellular capsule to facilitate

adherence to the host cell. When growth conditions are suboptimal, mycoplasma cells become distorted forming long strands that resemble fungi (thus accounting for their name: myco means “fungus”). Being wall-less, mycoplasmas can squeeze through small holes, even through the pores in membrane filters used to sterilize liquids. . As a result, tissue cultures of animal cells cannot be reliably sterilized by filtration; antibiotics, usually penicillin and streptomycin are added to the media to suppress growth of contaminating mycoplasmas.

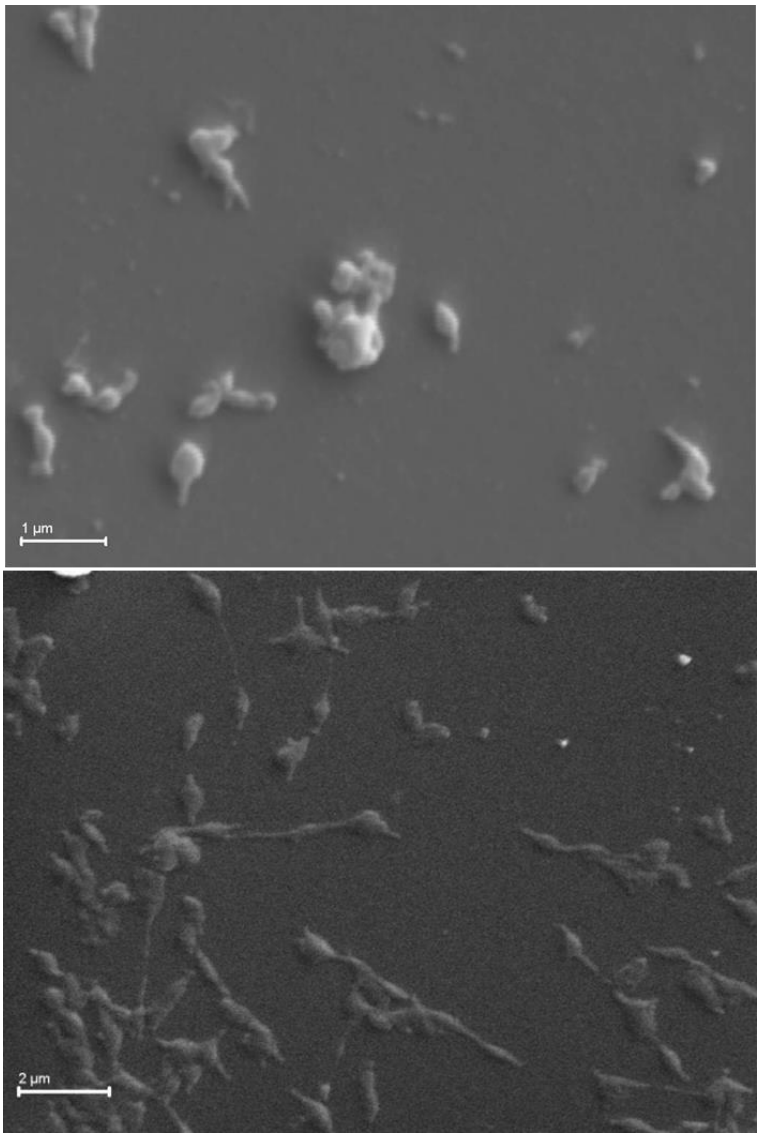

Figure 1.4 SEM images showing the typical morphology of *M. pneumoniae*.

Mycoplasmas were mistakenly described as viruses for nearly twenty years due to their ability to pass through a filter that excluded frequently encountered bacteria. In the process of elucidation, mycoplasmas were described as pleuropneumonia-like organisms (PPLO) and then as L-forms of bacteria, that is, bacteria that had lost their cell walls. Finally in the 1960's, DNA homology techniques allowed the designation of *M. pneumoniae* as Eaton's agent and subsequent taxonomic descriptor *M. pneumoniae* was assigned in 1963,[38] settling the debate over its status and identifying it as a new species of bacterium and not a cell-wall-less variant of bacteria that normally possess cell walls. Its characteristic lack of a cell wall and small size lend understanding to the initial confusion regarding the nature of Eaton's agent. Additionally, reductive evolution resulted in its biosynthetic limitations and necessitated a parasitic lifestyle, while confounding efforts to reliably culture the organism in clinical settings.

The main cytotoxic effect of *M. pneumoniae* is local disruption of tissue and cell structure along the respiratory tract epithelium due to its close proximity to host cells.[34] Attachment of the bacteria to host cells can result in loss of cilia, a reduction in metabolism, biosynthesis, and import of macromolecules, and, eventually, infected cells may be shed from the epithelial lining. The organism is not known to produce any exotoxins, but formation of hydrogen peroxide is a key virulence factor in *M. pneumoniae* infections. Attachment of *M. pneumoniae* to erythrocytes permits diffusion of hydrogen peroxide from the bacteria to the host cell without detoxification by catalase or peroxidase, which can injure the host cell by reducing glutathione, damaging lipid membranes and causing protein denaturation.[39] Local damage may also be a result of lactoferrin acquisition and subsequent hydroxyl radical, superoxide anion and peroxide formation. Local inflammation and

hyperresponsiveness by infection induced cytokine production has been associated with chronic conditions such as bronchial asthma and has also been linked to progression of symptoms in individuals with cystic fibrosis and COPD.[10] Accounting for approximately 20% to 40% of all community-acquired pneumonia (CAP) cases, and the leading cause of CAP infection in older children and young adults.[39-42] In adults alone the annual economic burden of CAP exceeds $17 billion, and the incidence of infection in the very young and elderly is on the rise.[42] Furthermore, extra-pulmonary sequelae occur in up to 25% of cases, and chronic *M. pneumoniae* infection may play a contributing role in the onset, exacerbation, and recurrence of asthma.

M. pneumoniae infection is acquired through respiratory secretions and spreads efficiently within close living quarters, with incubation periods as long as three weeks.[43] Symptoms tend to be nondescript and the disease often has complex and variable presentations, making definitive diagnosis challenging.[39,42,44] As a result, diagnosis of *M. pneumoniae* is often presumptive and relies heavily on the combination of physical findings and elimination of other possible causes.[41,43] Historically, serologic testing has long been considered the foundation for the diagnosis of *M. pneumoniae* infection but has severe limitations in sensitivity and specificity, a high tendency for false negatives, and often must be paired with another diagnostic method. At present, the gold standard for detection is qPCR. This method can exhibit high sensitivity and allow for detection in the early stages of infection, but the cost, complexity, and expertise it requires limit the practicality of widespread use in hospitals and reference laboratories. These limitations create a critical barrier to the accurate and timely diagnosis of *M. pneumoniae* infection,

and a rapid, simple, reliable diagnostic platform would greatly improve the control of *M. pneumoniae* disease.

One notable aspect of *M. pneumoniae* infection is the periodicity of outbreak epidemics shown to occur in regular patterns every 4 years.[45,46] Little is known regarding the significance of or factors involved with driving these epidemic cycles, as difficulties in the diagnosis of M. *pneumoniae* have impeded the ability to obtain consistent or comprehensive epidemiological data.[41,45] The incidence of disease does not appear be related to season or geography, however infection tends to occur more frequently during the summer and fall months when other respiratory pathogens are less prevalent. Also of interest in *M. pneumoniae* infection is the role of strain genotype in pathogenesis and disease epidemiology. Reinfection and epidemic cycling is thought to be a result of P1 adhesin subtype variation and is also an important virulence factor for *M. pneumoniae* infection.[47-49] Genetic diversity among *M. pneumoniae* is limited, and as such is generally categorized into one of two groups, type I (strain M 129) or type II (strain FH) based on variations within the sequence of the *P1* gene.[46] Type-switching between the two variants seems to regularly occur in 4-7 year cycles, though it is unknown whether there is a link between the observed trends in type-switching and outbreak periodicity.[48] *M. pneumoniae* genotyping is currently done by restriction length fragment polymorphism or by qPCR in combination with high resolution melt analysis.[41,44,46,50,51] While the application of these PCR-based typing assays has greatly facilitated the study of *M. pneumoniae* epidemiology, both require a great deal of cost, technical expertise, and additional testing beyond clinical detection that limit their practicality for widespread, point-of-care use.[52] Therefore, a biosensing platform with the ability to simultaneously detect and type clinical specimens

in a single assay would be of great value from both diagnostic and epidemiological standpoints.

CHAPTER 2

General Experimental Concepts

The number of methods for the preparation of nanostructures with SERS activity is numerous and synthesis continues to be a very active field of research in the development of nano-based sensors. SERS can be used to try to solve a vast range of chemical and biophysical related problems. Consequently, it has been accepted as mayor analytical tool in many research areas. For quantitative analysis there has been an urgency to develop homogenous, stable and reproducible SERS substrates fit for various needs. The purpose of this chapter is to introduce the reader to the theory and basic principles of Raman spectroscopy, the vibrational spectroscopic technique employed in this work. Also included within this chapter are the experimental techniques used, and an overview of the theoretical concepts will be shown.

2.1 Optical properties of metallic nanostructures

The optical properties of metallic nanoparticles depend on the shape and size and also of the dielectric constant and that of the surrounding medium. Spherical metallic nanoparticles exhibit a single surface plasmon (SP) band attributed to the coherent oscillation of the conduction electrons caused by oscillating electric field when they are irradiated by light.[53] Under the irradiation of light, the conduction electrons in a metallic nanostructure are driven by an electric field as shown in Figure 2.1, to collectively oscillate at resonance frequency relative to the lattice of positive ions. Depending on the metallic nanostructure, they will possess different surface plasmons resonances, that are dependent on size, shape and dielectric constant of the metal analyzed.

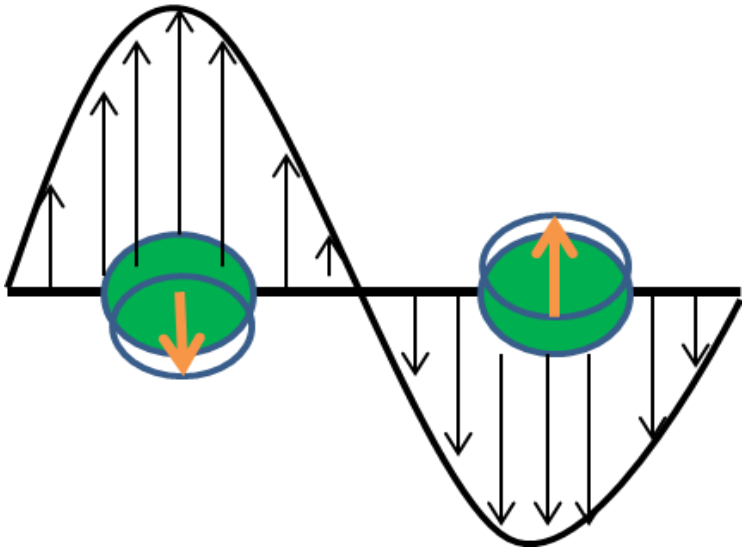

Figure 2.1 Schematic representation of a surface plasmon.

2.2 Surface Enhanced Raman Spectroscopy

In 1928, Sir C.V. Raman described the first evidence that light is inelastically scattered.[54] The Raman scattering cross-section is typically 10^{-29} cm^2 per molecule, a much smaller component than UV and IR cross-sections, which are 10^{-18} cm^2 and 10^{-21} cm^2 per molecule respectively;[55,56] this small cross-section has limited its application to analytical problems and favored the use of UV and IR for analysis of chemical samples in the past. It has been shown that the intensity of Raman scattering is inversely proportional to the fourth power of the incident light's wavelength, according to the Tyndall effect.[57]

Raman spectroscopy is a popular nondestructive, ambient probing tool used to characterize molecular structures and usually imposes very little constraint on the substrate size. Figure 2.2 illustrates the main events that occur when a light quantum $h\nu_0$ impinges on a surface. The elastic scattering process (Rayleigh scattering) of quanta with energy $h\nu_o$ occurs without energy loss.[58] This process has the highest probability among other competing scattering phenomena.

However, there are inelastic processes in which the vibrational energy is altered by $h\nu_s$ also prevail with much lower probability of occurrence. These inelastic processes are called Raman Scattering and quanta of energy $h\nu_o \pm h\nu_s$ are emitted. According to Boltzmann's law, at ambient temperature vibration of molecules in the excited state is much less probable than that of the ground state molecules. Thus, it is more efficient to excite ground-state molecules to a vibrationally excited state than to receive the radiative decay energy from the vibrating molecules.

Hence, the emitted quanta having energy of $h\nu_o - h\nu_s$ are more prevalent than the emitted quanta with energy of $h\nu_o + h\nu_s$. The Raman lines corresponding to the quanta with energy

of ($h\nu_o - h\nu_s$) are referred to as the Stokes lines whereas the higher energy lines ($h\nu_0 + h\nu_s$) are called the anti-Stokes lines. As the intensities of the anti-Stokes lines are lower, only the Stokes lines are usually recorded in a typical Raman spectrum. The light scattering processes are illustrated in Figure 2.2. Raman spectra are described by the observed wavenumber shift in (cm^{-1}) as a result of excitation by an external source. Raman spectra collected at low excitation energies, such as a 785nm laser, are usually not complicated by fluorescence contamination due to the lower state of excitation energy used.

Two other forms of vibrational spectroscopy, normal Raman and IR, have been explored for use in bacterial detection. IR was initially favored for its usefulness in describing the chemical natures of samples. However, IR suffers from a large water and CO_2 effect, which tends to be present in every biological sample. Furthermore, a larger sample size is required as sensitivity is limited with IR detection, so identification of small concentrations of a compound or small numbers of biological agents in a complex matrix is not consistently possible. Normal non-SERS Raman scattering signals of biological samples, considered a bulk sampling technique, have been used for bacterial detection in the past[59 60] but are typically limited due to the smallness of the cross section. Furthermore, protocols must be standardized and stringent to render the Raman data useful and reliable for classification.

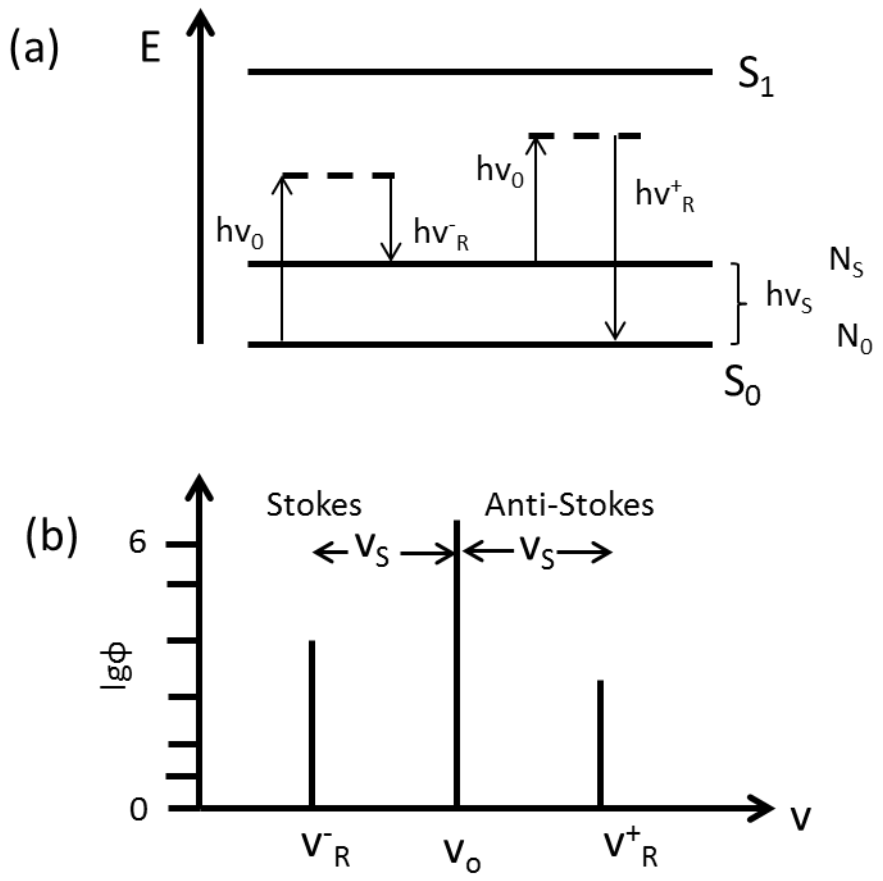

Figure 2.2 Schematic diagram illustrating the principle of Raman scattering : (a) term diagram; (b) Raman spectra. Because vibration of atoms in the excited state is much less than that of the ground state atoms, the Stokes line is stronger than anti-Stokes line.

2.3. SERS Mechanism on Coinage and Transition Metals

Elucidation of the mechanisms that enable the SERS phenomena is always an important area of research in SERS. This active area of research has attracted much interest from the communities of surface science, spectroscopy, condensed phase physics and nanoscience. Two well established factors that contribute to the SERS mechanisms include the Electromagnetic Enhancement and the Chemical Enhancement.

In order to understand the Electromagnetic (EM) enhancement,[61,62] one must consider the size, shape and material of the nanosurface. Nanoscale roughness features affect the electromagnetic field enhancement of Ag, Au, and Cu and is considered to mainly come from geometrically defined local surface plasmon resonances (LSPRs) at metal nanostructures. If the correct wavelength of light strikes a metallic roughness feature, the plasma of conduction electrons will oscillate collectively. The surface Raman enhancement arising from the LSPR of a spherical nanoparticle can be estimated using the following equation:

$$G(\omega)\alpha\left|\frac{\varepsilon_i-\varepsilon_0}{\varepsilon_i+2\varepsilon_0}\right|^2.\left|\frac{\varepsilon_s-\varepsilon_0}{\varepsilon_i-2\varepsilon_0}\right|^2.\left(\frac{r}{r+d}\right)^{12} \quad (1)$$

where ε_0 is the dielectric constant of the medium, ε_i is the wavelength-dependent complex dielectric constant of the metal nanoparticles, r is the radius of the nanoparticles, and d is the distance of the point to be measured to the center of the nanoparticle.[63-65] The EM enhancement mechanism contributes more than 10^4 times enhancement over normal Raman scattering. The LSPR allows the resonant wavelength to be absorbed and scattered, creating large electromagnetic fields around the roughness features. When the incident

light is resonant with the collective modes of oscillation of the metal conduction band electrons, there is an enhancement of the electromagnetic field at the surface at both the incident and Raman scattered wavelengths. This enhancement in turn induces a change in polarizability of the molecule leading to intense Raman signals. For the electromagnetic enhancement effect to be operative, it is not necessary to be in contact with a metal surface.[66] Considering that in most of SERS studies the medium is water and $\varepsilon 0 = 1.77$, when the real part of ε_i approaches -3.54 and the imaginary part approaches 0, the enhancement reaches the maximum. Some free electron metals, such as Ag, Au, and Cu, can meet this condition and therefore produce very high enhancement. When the shape of the nanoparticles deviates from the spherical nanoparticles, such as ellipsoids or rods, the lightning-rod effect should be considered due to the existence of the high curvature points. Furthermore, more and more experimental and theoretical studies demonstrate that the coupling between nanoparticles can effectively increase the enhancement effect. These general conclusions can also be borrowed to analyze the case of transition metal systems. Most transition metals show a large imaginary part of the dielectric constants and therefore are not effective in generating high SERS activity over the visible light region.[63]

Chemical enhancement, according to its definition, includes any enhancement of the Raman intensity of surface species resulting from its chemical interaction with the surface or other surface species and is clearly correlated to the charge transfer between the probed molecule and surface or other surface species.[63] One should note the following three types of charge-transfer process that contribute to the chemical enhancement: (1) when a molecule interacts with the surface or other surface species, the electron distribution and therefore the polarizability of molecule will change. Such a change may cause different

enhancements for different vibrational modes; (2) metal ion, probed molecule, and electrolyte ion may form a surface complex, leading to the change of the polarizability of the molecule. Some surface complexes may even create a new electronic level in resonance with the incident laser energy, similar to the resonance Raman Effect of a complex; (3) the third type is the most complicated one and called photon-driven charge transfer process or often simplified as charge transfer. This process occurs when the incident laser energy matches the energy difference between the surface molecules' HOMO or LUMO and Fermi level or surface state of the metal substrate. This process could also be associated with the excited state of the whole molecule/metal system and with the charge transfer between the molecule and the metal surface (or surface ad-clusters). This could result in a considerable increase in the Raman intensity of probe molecules. The photon-driven charge-transfer mechanism has been used to explain the change of intensity–potential profile with the changing excitation line and is the most important type of chemical enhancement mechanism. The chemical enhancement mechanism, now thought to contribute an enhancement factor of 10^2, states that a charge-transfer state is created between the metal and adsorbate molecules.[63]

In systems where both the chemical and electromagnetic enhancements come into action, the mechanisms are multiplicative.[67] Estimations regarding the overall enhancement per molecule for the contribution of both chemical and electromagnetic enhancement mechanisms, scales on the order of $10^5 - 10^6$.[68,69] It is commonly thought that the EM enhancement contributes the greater extent (~10^4-10^6) to the signal enhancement observed in SERS while the CE enhancement is thought to contribute to a smaller extent (~10-10^2)

to the overall enhancement.[70,71] Taking these parameters into consideration, the SERS surface enhancement factor (SEF) may be determined using the following expression:

$$SEF = \frac{\frac{I_{SERS}}{\mu_M \mu_S A_M}}{\frac{I_{RS}}{C_{RS} H_{eff}}}$$

In the above expression, I_{SERS} and I_{RS} are the SERS and standard Raman intensities, respectively. μ_M is the surface density of the individual nanostructures, μ_s is the surface density of molecules on the substrate surface, A_M is the surface area of the metallic surface, C_{RS} is the concentration of the solution used, and H_{eff} is the effective height of the scattering volume. Using the above expression, enhancement factors of up to 10^{14} have been reported for single molecule detection.[72,73] This interaction between analyte and metal spurred activity to produce this electromagnetic effect in stable colloids of silver or gold particles and thus serve as sensors for amplifying the signal from the analyte of interest.[74] Problems arose as these colloidal solutions, while easy to make, were variable and unstable in their production of the SERS effect. The search for "hot spots" became the prevailing drive in research for a time. Around the same time, the development of nano-structured surfaces, on a scale of 10^{-9} meter, was emerging.

2.4 Silver Nanorod Arrays- Fabrication through the OAD method and functionality as a SERS substrate

As previously discussed, for thin noble metallic films, when electromagnetic (EM) waves interacts with the metallic surface, it will generate charge density wave, and at certain wavelength or incident geometry, geometry, when the horizontal component of the incident wavevector matches with the intrinsic wavevector of the charge density wave, a resonance absorbance will occur, which is referred to as SPR. For individual metallic nanostructures, similar resonance absorbance due to confined collective oscillation of electrons in the nanostructure can be observed, and is referred to as LSPR. The LSPR depends on the material, size, shape, and also the surrounding dielectric environment,[75-77] and has been widely applied for chemical and biological sensing.[78,79]

A great deal of efforts have been devoted to the fabrication of various metallic nanostructures for plasmonic applications. Many of them require nanostructures supported on substrates. This imposes a great challenge to the fabrication techniques. Methods that can directly fabricate nanostructures on substrates are usually used, such as lithography methods,[80,81] self-assembly methods,[82] and physical vapor deposition (PVD) methods.[83,84] Self-assembled metallic colloid is another popular way to produce ordered nanostructures.[82] Physical vapor deposition methods, such as electron-beam deposition, sputtering growth, and thermal evaporation, are commonly used to fabricate large-scale thin films. With a normal vapor incidence with respect to the substrates, nanoparticles or thin films are generally formed on the substrates. When the vapor is incident with a large angle ($\theta>70^{\circ}$) with respect to the surface normal of substrates, nanocolumns are usually

formed. This kind of physical vapor deposition with large incident angles is referred to as oblique angle deposition (OAD).[85]

The principle of OAD is shown in Figure 2.3a. The incident vapor flux is arriving at the substrate at an angle θ, which can be tuned by changing the orientation of substrates through a stepper motor. A quartz crystal microbalance (QCM) is usually used to monitor the deposition thickness, which is referred to as QCM thickness t_{QCM}. In OAD, the geometric shadowing effect and surface diffusion of the adatoms account for the formation of the nanostructures.[85,86] These conditions result in preferential growth of cylindrical, irregularly shaped rods that are randomly, yet uniformly, distributed on the surface. The growth of the nanorods arises from initial metal nucleation sites in the direction of vapor depositon. The nanorods produced by OAD are cylindrical in shape but encompass a variety of random protrusions and irregularities.SERS has become a powerful technique for chemical and biological sensing applications due to its great sensitivity. The application of SERS usually requires the substrates with good sensitivity and reproducibility.

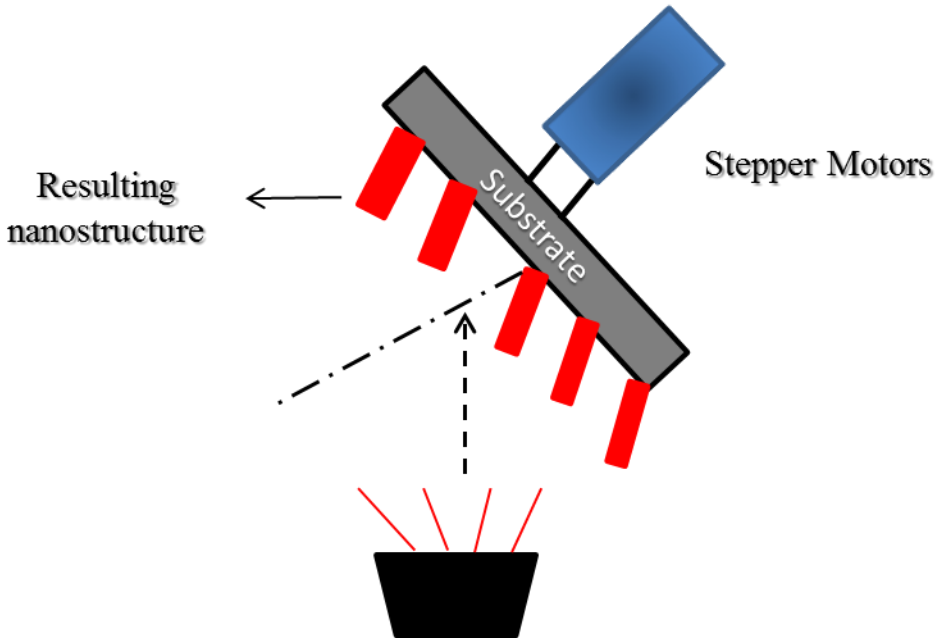

Figure 2.3 The schematics of the development of the silver nanorods substrate through OAD

A variety of Ag nanostructures fabricated by OAD and GLAD techniques have been explored[83,87-90] and some of them have demonstrated as excellent SERS substrates. Among these Ag nanostructures, tilted Ag nanorod(Ag NR arrays are simple to prepare by OAD, and have exhibited very good sensitivity and reproducibility.[83,91-94] This high SERS signal enhancement observed with silver nanorod arrays results from the chemical and/or chemical electromagnetic enhancement mechanisms previously discussed within this chapter. The SERS response of these AgNR substrates is characterized by measuring the Raman spectrum of trans-1,2-bis(4-pyridyl) ethylene (BPE) molecules adsorbed on AgNRs. The enhancement factor (EF), defined as the enhancement in SERS signal per molecule relative to the bulk Raman signal, is evaluated to be as high as 5×10^8. The uniformity and reproducibility of the AgNR substrates are also estimated by static measurement:[95] the point-to-point uniformity (relative standard deviation) in BPE SERS intensity obtained on a single substrate is about 10%; the substrate-to-substrate reproducibility within one batch ranges from 6% to 13%; and the batch-to-batch reproducibility is less than 15%.[96] Enhancements associated with high curvature regions have been deemed the "lightning rod effect" and contribute to the high enhancement factor observed for molecules at the tips of the nanorods or in pores. The overlap of both the longitudinal and lateral plasmon bands is primarily responsible for the strong electromagnetic (EM) enhancement at the tips of the nanorods. As a result, the maximum SERS intensity observed for these Ag nanorod arrays is in the polarization direction perpendicular to the long axis of the nanorods.[96,97] These nanostructures have been extensively used for a variety of sensing applications. Specifically, silver nanorod arrays fabricated by OAD have been used as SERS-active substrates for detection of a variety of

analytes. In part of this work we have employed silver nanorod (AgNR)-based SERS to detect the chemical signatures found in lipids extracted from the cell envelopes of tuberculous (MTB) and non-tuberculosis (NTB) mycobacterial species.

2.5 Historical perspective on layer-by-layer self-assembly

Novel materials have always been sought and the employment of surface modification at the molecular level realized this goal. Surface modification resulted in a multitude of new properties that were previously not associated with the native material. These changes include modifications of the electrical, optical, magnetic, physicochemical and biological properties of the material in question. As a consequence, several disciplines of natural science have experienced the impact of surface modifications, changing the fundamental properties of materials at the building-block-level. Blodgett expanded the Langmuir film technique to produce multilayer coatings known as the Langmuir–Blodgett (LB) technique.[98]

The group of Kuhn then explored the possibility to adsorb different oppositely charged dyes with the LB technique, discovering the potential of layer thickness and energy transfer. Iler[99] however observed that oppositely charged colloids could be alternately assembled onto glass substrates and work by Nicolau and colleagues illustrated successive layering of substrates with oppositely charged metal ions to produce polycrystalline coatings[100] and successive polymerizations steps in situ to produce alternating polymer coats onto a substrate and successive polymerizations steps in situ to produce alternating polymer coats onto a substrate.[101]

The work of Iler and Nicolau probably inspired the seminal breakthrough made by the group of Decher, who used synthetic polyelectrolytes i.e. polymers with ionizable surface

groups to form polyions that were successively layered onto a substrate by electrostatic interaction. This method is robust, simple, does not require sophisticated equipment and precise stoichiometry, nor does it rely on complicated chemical reactions to deposit successive layers. Layer-by-layer self-assembly (LBL) is still seen as the true alternative to the LB technique.

Since the late 1990s, work has focused on development of multilayer composites based on interactions other than electrostatic interactions such as hydrogen bonding.[102-104] Controllable polymerization reactions also resulted in novel approaches to assemble layer-by- layer constructs through successive polymerization.[105] Polyelectrolyte multilayers (PEMs) are furthermore of fundamental physical interest, since they form two-dimensionally stratified layers, which are growing step-by-step into the third dimension. This leads to a behaviour being dominated by internal interfaces, and differing largely from the corresponding volume material properties.

The layer-by-layer (LBL) self-assembly of multiple polyelectrolytes and other particles resulted in the production of multifunctional hybrid carrier systems[106,107] for dyes,[108] sensors,[109] enzymes,[110] drugs,[111] multiple components,[111] and cells.[112] Additionally, nanocoated substrates provide a surface platform for the attachment of targeting molecules, i.e. folic acid,[113] antibodies,[114] or a variety of surface functional groups such as hydroxyl, carboxyl and thiol groups.[115]

2.5.1 Principles of the layer-by-layer technique

The buildup of LBL mutlilayers is driven by the electrostatic attraction between the oppositely charged constituents.[99] However, hydrogen bonding,[116,117] hydrophobic interactions,[118] and van der Waals forces[119] may be exploited to assemble LBL systems or

influence the stability, morphology and thickness of the films, particle/molecule depositions and permeation properties of the film.

Generally, LBL self-assembly proceeds as follows: (1) a charged substrate is immersed in a solution of an oppositely-charged colloid to adsorb the first monolayer, (2) a washing cycle follows to remove unbound material and preclude contamination of the subsequent oppositely-charged colloid, (3) in which the coated substrate is submerged to deposit a second layer and the multilayered structure is formed. [105] Some LBL processes require no washing cycles thus shortens the duration of the assembly process.[120] A generalized scheme of the LBL process is show in Figure 2.4.

The polyelectrolytes or colloids, which exhibit a high linear surface charge density, are utilized in excess to prime the substrate. Therefore, a non-stoichiometric excess of charge is absorbed after each step relative to the preceding layer.[121,122] This surplus of charge provides the step-wise mechanism for the reversal of the surface charge polarity, facilitating a favorable surface for the adsorption of the subsequent layer. The LBL self-assembly methods have advantages compared to the more conventional coating methods, including (1) the simplicity of the LBL process and equipment, (2) its suitability to coating most surfaces, (3) the availability of an abundance of natural and synthetic colloids, (4) the flexible application to objects with irregular shapes and sizes, (5) the formation of stabilizing coats and (6) control over the required multilayer thickness.[123-125]

Polyelectrolytes and nanoparticles can be utilized to form the ultrathin multilayer structures using the LBL self –assembly technique. Polyelectrolytes are classified according to their origin. Standard synthetic polyelectrolytes include poly(styrene sulfonate) (PSS), poly (dimethyldiallylammonium chloride) (PDDA), poly(ethylenimine) (PEI), poly(N-

isopropyl acrylamide (PNIPAM), poly(acrylic acid) (PAA), poly(methacrylic acid) (PMA), poly(vinyl sulfate) (PVS) and poly(allylamine)(PAH).[123] Natural polyelectrolytes include nucleic acids, proteins and polysaccharides of which alginic acid, chondroitin sulfate, DNA heparin, chitosan, cellulose sulfate, dextran sulfate and carboxymethylcellulose are most common.[126,127] The prerequisite for successful LBL coating is the presence of a minimal surface charge, which is one of the few disadvantages of the technique. However, charge can be induced to still facilitate LBL.[128] Most commonly glass, quartz, silicon wafers, mica, gold-coated substrates are coated. The type of substrate that is encapsulated depends primarily on the colloids assembled into PEMs and analytical monitoring techniques for the coating steps.[123]

Surface charge is not the only factor that may affect the multilayer adhesion. The surface texture could also affect the adhesion properties. Pretreatment of a substrate by annealing with sodium chloride smoothened the surface of the substrate, resulting in more intimate contact between the substrate and colloid to produce higher quality coats

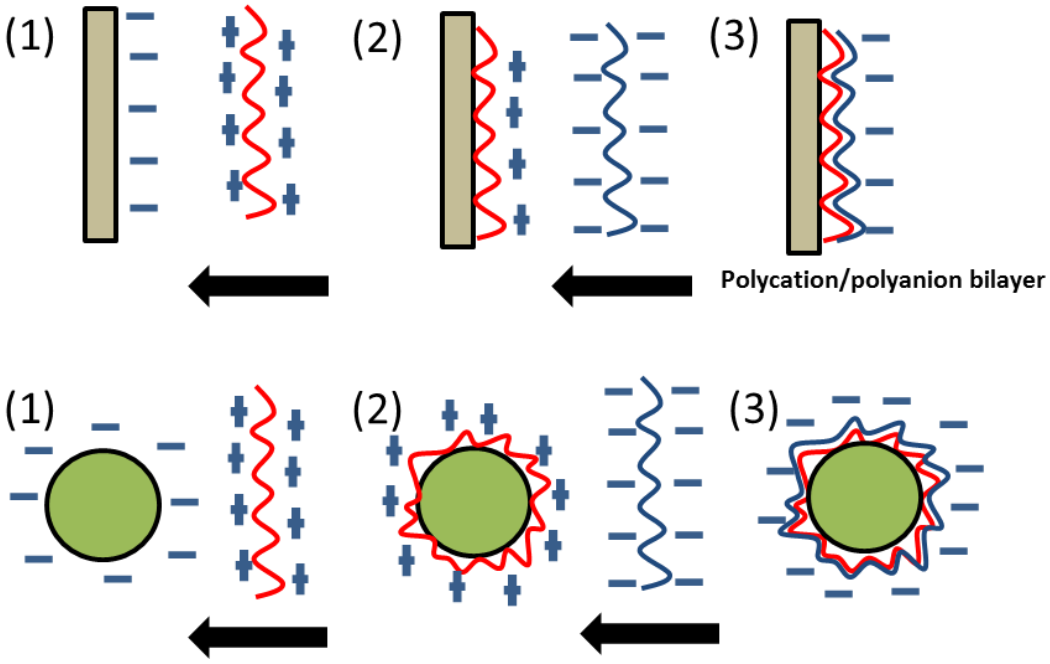

Figure 2.4 Schematic illustration showing how LBL assembly is achieved by observing the absorption of oppositely charged polyelectrolytes, from reference 105.

2.5.2 Fundamental Experimental parameters of LBL adsorption

From a practical point of view, LBL assembly only requires the exposure of a charged substrate in a solution of a polyelectrolyte with a charge opposite to that of the substrate. The procedure is repeated a number of times alternating between polycations and polyanions, including water or salt washings in between layers, until a film of the desired thickness is obtained.[121,129]

The formation of polyelectrolyte multilayer self-assembly is usually reliant on the electrostatic adsorption between the substrate and subsequent layers.[99,121] A two-stage process is envisioned by which (1) an initial anchoring of the coating material to the surface is followed by (2) a slow relaxation to form a densely-packed structure on the surface. Though the time scales vary largely, a common the two-step process, suggests that a fast first order adsorption process is followed by slow rearrangements at the surface: the first kinetic step involves the transport of chains to the surface by diffusion, so that adsorption due to electrostatic forces can take place. The second growth step is diffusion-controlled since the level of coverage in the coated domain is saturated with time.[130] Subsequent slow chain rearrangements enable the diffusion of segments into the inner regions of the previously deposited layer. By mixing of positive and negative segments, finally the irreversible complexation of charges occurs.[131]

The adsorption of PEMs from salt solutions of varying electrolyte concentration was the first approach to control layer thickness over a wide range.[129] Rather different power laws of the adsorbed amount or layer thickness in dependence of the salt concentration were found: while some authors reported a linear dependence on c_{salt}, others found the thickness

to scale with c^{α}_{salt} with α=0.5, or the adsorbed amount scaling with an exponent of 0.05–0.15.[105]

The practical procedure applied by the different groups working in LBL assembly varies, regarding the number of washings between the deposition of layers, the use of NaCl or water for the washings, assembly time, pH, and the ionic strength of the polyelectrolyte solution. This parameters can have a strong impact on the quality of the assembled PEMs. For instance, although polyelectrolyte conformation is determined by the ionic strength in bulk solution, the thickness of an LBL film will be notably influenced by the ionic strength at which the assembly has been performed. If polyelectrolyte assembly takes place at low ionic strength, the polymer chains are normally extended, resulting in a thin film. Increasing the ionic strength results in the coiling of the chains, which become less extended but increase in volume.[132] This, in turn, results in an increase in layer thickness. Standard values of ionic strength for PEM fabrication range between 0.1 and 1 M NaCl due to considerations of layer thickness and packing of polyelectrolyte chains.[133]

Polyelectrolyte assembly can, nevertheless, be performed at higher ionic strengths even though the electrostatic interaction between subsequent layers is reduced because of charge screening. Although electrostatic interactions become weaker, the assembly at high ionic strengths is still favored by the release of counterions during layer assembly. In general, PEMs do not show significant changes in thickness with the ionic strength. It has been reported that an increase in the ionic strength makes PEM swell since the interaction between oppositely charged polyelectrolytes is weakened and the fi lm has the possibility to swell.[134]

2.6. Bioanalytical Sensing and Pathogen Detection by SERS

The ability to rapidly detect and identify bacterial cells in human body fluids at relatively low cost and in point-of-care settings is a continuing need for health care providers worldwide. Traditional methods of bacterial identification are phenotypic based approaches that require a cell growth period and are consequentially slow (24 – 48 hours or longer). Furthermore, distinguishing closely related strains may be difficult via traditional methods and it is not a point of care technique. The best current methods are molecular diagnostic approaches that utilize specific primers or probes for particular gene targets, such as "real time" polymerase chain reaction methods (PCR)[135] which are increasingly finding use in clinical settings. If no culturing is required PCR time frames are typically in the 2 – 6 hour framework. However, PCR is not without some limitations or liabilities such as sample contamination, infectious mixture resolution, need for required primer sets, speed, cost and point-of-care capabilities.[136] In general these methods are generally very sensitive and selective, but exceedingly time consuming due to the fact that they rely on several enrichment steps.[89,137] Other current diagnostic methods available for the detection of pathogens rely on immunofluorescence tests and antigen-capture immunoassays such as enzyme-linked immunosorbent assay (ELISA),[138]hemi-nested multiplex RT-PCR,[139,140] or hemadsorption.[141] Many of these assay techniques suffer from lack of sensitivity and reproducibility or require the use of synthetic labels, species-specific reagents (genotyping primers). Consequently, new diagnostic approaches for rapid and sensitive means of detecting pathogens both in the laboratory and in the field are urgently needed to successfully curb potential disease outbreaks.

For the past few years there have been developing techniques for optical approach for rapid, sensitive and specific bacterial and biomedical diagnostics of infectious diseases based on surface enhanced Raman spectroscopy (SERS).[60,142,143] SERS is a well-known spontaneous light scattering technique, discovered ~35 years ago which results in the 10^5 – 10^8 effective enhancement of the Raman scattering intensity of some molecular vibrational modes of molecules that are close (≤ ~5 nm) to nanostructured metal surfaces.[69,144] This effect is predominantly attributed to the plasmonically enhanced local electric fields that become concentrated near nanosized structures that are coincident with the surface plasmon resonances of these nanomaterials. These plasmon resonances are usually in the NIR to visible for the most commonly employed metals, Ag or Au. However, until the development of stable, reproducible nanomaterials which could enhance the Raman signal reproducibly (SERS), this approach was not practical. The ultrasensitive, rapid, and label-free attributes of SERS seem to satisfy all criteria needed for routine analysis of different species and strains of pathogens. As a label-free, molecularly specific detection method, SERS allows detection of pathogens with a high degree of sensitivity and specificity. In contrast with the other diagnostic methods previously noted, SERS offers several advantages, including the ability to provide molecular vibrational information for both in-vivo and in-vitro applications, hence allowing discrimination of subtle structural differences and classification of pathogen types and species.[145,146] The main reasons for the interests in SERS are its high sensitivity, intrinsic selectivity due to the spectroscopic finger-print, simple and fast preparation and nondestructive data acquisition in aqueous environment.

The assignment of individual peaks in spectral data to specific biological input has also been prevalent in the literature. Vast spectral libraries have been compiled from classic analytical chemistry studies i.e. spectral analysis of a purified single component. Comparison of bacterial fingerprints to these pure components reveals small fluctuations in wavenumber due to the matrix effect, but none-the-less can be helpful in elucidating the contents of the sample. For example, proteins,[147] nucleic acids,[148] carbohydrates,[149] and lipids[59] have been extensively studied and have representative spectra in bacterial samples. Over the last several years, Ag nanorod arrays fabricated by OAD have been evaluated as SERS-active substrates for a variety of biosensing applications. The unique signal enhancement offered by Ag nanorod substrates allows nondestructive detection and discrimination of pathogens at the strain level with minimal sample or culture preparation. Ag nanorod arrays have been employed for the SERS-based detection of a variety of bacterial[83,90,92,150,151] and viral[41,94,152] species to allow classification of bacterial and viral types as well as discrimination between pathogen strains and pathogens having single gene mutations. In addition, oligonucleotide-modified Ag nanorod arrays have been used to directly probe sequence information and detect binding of complementary targets. The following sections summarize recent applications of the use of these SERS substrates for pathogen detection and provide an outlook on the future of Ag nanorod arrays for bioanalytical and diagnostic applications.

A number of studies have been performed using SERS with a variety of substrates: in situ silver[153,154] or gold colloids,[142] silver or gold colloids[155] added post cell growth, silver coatings,[153] silver nanoparticles with antibodies,[156] and biofilms. Culha and co-workers used a gold SERS-active substrate with 785 nm excitation to investigate the individual

components of bacteria and concentrated their analysis on the study of the spectral features of nucleic acids, amino acids, and peptides. However, they stated that SERS spectra obtained from bacteria under these conditions will take their major contribution from cell surface biochemistry but omitted to analyze the carbohydrate components associated with the outer cell membrane present in the SERS spectra.

As mentioned previously, simple mixing e.g. metal nanoparticles with bacteria is a common approach for SERS detection of biological samples. However, the generated mixture is not always homogeneous.[145,157] There are no specific interactions between the bacteria and NPs. Additionally, the capping reagent employed for stabilizing the NPs prohibits intimate contact with bacteria. The inefficient interaction of NPs with the bacterial cell wall resulted in a very limited reproducibility of bacterial SERS spectra. The inefficient interaction of NPs with the bacterial cell wall resulted in a very limited reproducibility of bacterial SERS spectra. This limitation can be overcome either by increasing the NP concentration[158,159] or by the "convective assembly" method,[160] but these methods often have poor reproducibility due to the difficulty posed by efficient, precise control of the NP aggregation process on the bacteria. Thus, the main goal is that the NPs could come into contact with the bacteria surface at as many points and close as possible.[158] Recently, an electrostatic attraction force strategy for efficient, precise self-assembly NPs on bacteria has been developed.[161-163] The fact is based on that the bacteria wall is negatively charged for the presence of either teichoic acid in Gram-positive bacteria,[162] or the outer membrane lipopolysaccharides in Gram-negative bacteria. The positive charged poly (L-lysine) coated AuNPs,[164] cetyltrimethylammonium bromide (CTAB)- terminated nanorods and poly(allylamine hydrochloride) (PAH)/AuNPs/PAH layer by layer (LbL)

structures[161] can allow the highly efficient and precise deposit onto the bacterial cell wall through electrostatic interaction.

Recent research has suggested a faster approach for developing LBL structures for SERS analysis. Zhou's group[165] recently assumed that traditional LBL structures are difficult to prepare and that SERS spectra of the bacteria analyzed are prone to interfering signals of the capping agents. The later group stated that Efrima's group were the first to develop a method to directly produce NPs in the presence of the bacteria.[153,154,166,167] They directed the production of the NPs to either intimate contact at external (cell wall) or interior components of the bacteria. Although the procedure ensures an intimate contact of the colloids with the cell, the authors obtained very similar spectra for different species of silver-coated bacteria, they claim that this makes the method unsuitable for bacteria discrimination. So they developed an in situ synthesis of AgNP coating on the cell wall of bacteria for SERS-based label free detection of bacteria in drinking water, which is a more efficient way to achieve intimate contact between NPs and the bacteria cell in comparison to previous methods. They first soaked bacteria in a silver nitrate solution and then used hydroxylamine hydrochloride as a reducing agent. Finally, a "colloid deposit" was formed on the cell wall of the bacteria, in which they addressed in following as "Bacteria@AgNP". SERS substrates fabricated by OAD and the LBL structures that will be presented within this project have been assessed as a potential analytical sensor for rapid pathogenic bacteria detection. The main challenge associated with the detection and identification of bacteria lies in distinguishing the protein and carbohydrate components present along the outside of the bacterial cell, which is characteristic to a specific cell line and different between bacterial strains. This change in composition has been investigated by SERS in various

reports. Different studies demonstrate that the spectral patterns can be altered by a variety of experimental conditions such as matrix effect, the background contents, as well as sample handling. This variation has not confounded attempts to use spectral data for species and strain level differentiation.[60,168] In spectral data, a large number of wavenumbers are collected for each sample, and many of these display collinearity with one another, that is, one or more wavenumbers are linearly dependent in their response to the incident energy, a condition which could cause problems in statistical analysis. Chemometric software programs, by optimizing the similarity of the data structure, can extract a large amount of information out of the spectra. A brief description of principal component analysis (PCA) as well as other multivariate statistical techniques are provided in Chapter 3. These multivariate statistical methods demonstrate the ability to distinguish different bacterial species.

CHAPTER 3

Multivariate Statistical Analysis in Chemistry

The purpose of this chapter is to introduce to the reader to commonly used multivariate statistical methods employed for spectral data analysis in vibrational spectroscopy.

3.1 Pre-processing in vibrational spectroscopy

Many users of chemometric software simply accept without much insight the results of PCA: yet interpretation depends critically on how the original data have been handled. Data preprocessing or scaling can have a significant influence on the outcome. For the application of multivariate data analysis to work optimally, it is vital to pre-process the data in a correct manner. If this is not done, there will be a mix-up between the information which is sought and the noise which one is not interested in. Noise does not only consist of random deviations in the measurements themselves. It can also contain systematic variations in the samples which are not of interest to the analyst.[169,170] One such variation is the light scattering; created by particles which are illuminated. This effect is nearly non-existent for liquid samples (although suspensions will show scattering), while solid samples are prone to show scattering.

A natural pre-processing technique to apply to raw spectral data would be the application of a derivative. Most often a first or a second order derivative is used. The first derivative will remove any offset difference between the data and the second derivative will furthermore remove any slope effect on the data. There are two typical ways of estimating the derivative: Norris–Williams derivation (NW)[171] and Savitzky–Golay derivation (SG).[172] The former is in many ways a simplification of the latter. In NW the smoothing performed to the data is according to a 0th order polynomial (the average only), while for

SG this smoothing function can be set to any polynomial order (two is though the most common). The second parameter which should be set for NW is the gap size. This truly involves is a bit of a curiosity specially when there is nothing clear in the spectra which should indicate that you need a gap size for anything. The larger the gap-size, the more numbers are used in the estimation of the derivative, and thus a greater smoothing effect is achieved. The total smoothing effect is thus a combination of the window size and the gap size used. The smoothing effect can, though, never be higher than the window size. SG on the other hand estimates a polynomial on the window size which is used for the smoothing of the data. The derivative is subsequently estimated from this fitted polynomial. The window size thus has a direct and straight forward effect on the estimated derivative.

The SG derivatization in general is a common spectral processing method which alleviates issues from spectra-to-spectra baseline variations. The SG method as well improves the S/N ratio, reduces noise, sharpens peaks and eliminates the need for manual baseline correction of the spectra. Each derivative spectrum is then normalized with respect to its maximum value such that the values in the vertical axis are ranged from 0 to +1. Normalization also accounts for slight differences in the enhancement factors provided by each substrate. Mean centering is used so that unsupervised chemometric analyses such as principal components can be informative.

Principal component analysis (PCA) has been performed directly on the raw spectral data, something statisticians in other disciplines very rarely do. It is important to be very careful when using packages that have been designed primarily by statisticians, on chemical data. Traditionally, what is mainly interesting to statisticians is deviation around a mean, for example, how do the mean characteristics of a forged banknote vary? What is an ‘average’

banknote? In chemistry we are often (but by no means exclusively) interested in deviation above a baseline, such as in spectroscopy. It is, though, possible to mean center the columns in a data matrix, it will be noticeable that the sum of each column will be 0. Almost all traditional statistical packages perform this operation prior to PCA whether desired or not. Note that it is also possible to mean center the rows, and also double mean center data both simultaneously down the columns and along the row however, this is rarely done in chemometrics.

Standardization is another common method for data scaling and first requires mean centering: in addition, each variable is also divided by its standard deviation. In other literature this is also referred to autoscaling. Prior to performing PCA autoscaling which will make each column have the same 'size' so that all variables have an equal opportunity of being modelled. Autoscaling means that from each variable, the mean value is subtracted and then the variable is divided by its standard deviation.

3.2 Unsupervised and Supervised Pattern Recognition

A more formal method of treating samples is unsupervised pattern recognition, often called cluster analysis. Many methods have their origins in numerical taxonomy. Biologists measure features in different organisms, for example various body length parameters. Using a couple of dozen features, it is possible to see which species are most similar and draw a picture of these similarities, such as a dendrogram, phylogram or cladogram, in which more closely related species are closer to each other. The main branches can represent bigger divisions, such as subspecies, species, genera and families.

These principles can be directly applied to chemistry. Unsupervised pattern recognition differs from exploratory data analysis in that the aim of the methods are to detect

similarities, whereas using exploratory data analysis EDA there is no particular prejudice as to whether or how many groups will be found.

There are many reasons for supervised pattern recognition, mostly aimed at classification. Multivariate statisticians have developed a large number of discriminant functions, many of direct interest to chemists. Supervised techniques require a training set of known groupings to be available in advance, and try to answer a precise question as to the class of an unknown sample. It is, of course, always first necessary to establish whether chemical measurements are actually good enough to fit into the predetermined groups. However, spectroscopic or chromatographic methods for diagnosis are often much cheaper than expensive medical tests, and provide a valuable first diagnosis. In many cases chemical pattern recognition can be performed as a form of screening, with doubtful samples being subjected to more sophisticated tests.

3.3 Principal Component Analysis (PCA)

PCA is probably the most widespread multivariate statistical technique used in chemometrics, and because of the importance of multivariate measurements in chemistry, it is regarded by many as the technique that most significantly changed the chemist's view of data analysis. Natural scientists of all disciplines, from biologists, geologists and chemists have caught on to these approaches over the past few decades. Within the chemical community the first major applications of PCA were reported in the 1970s, and form the foundation of many modern chemometric methods.[173]

The aims of PCA are to determine underlying information from multivariate raw data. There are two principal needs in chemistry.[174] The first is to interpret the principal components (PCs) often in a quantitative manner. The number of significant PCs. The

second are the characteristics of each PC, usually the scores relating to the objects or samples and the loadings relating to the variables or measurements.

Principal component analysis (PCA) is a method for the visualization of complex data by dimension reduction. The high number of variables in multivariate data leads to three main problems: (1) graphical representation of the data is not possible for more than three variables; (2) high correlation between the variables makes it impossible to apply many statistical methods; and (3) many variables contain only very few information. PCA is able to avoid all of these problems by transforming the original variables into a smaller set of latent variables which are uncorrelated. Data transformed in such a way can then be used by other methods. The latent variables with the highest concentration of information form lower dimensional data which can be visualized graphically as seen in Figure 3.1.[174] Noise is separated from important information. PCA deconstructs the data and reduces the dimensionality by finding the vector in multivariate space which captures the maximum variance and redrawing the space with that component as the new x-axis. Then, the next dimension is found with the second highest variance captured, this vector being orthogonal, at 90° to, and independent from the first vector. Principal components are ordered by decreasing eigenvalues, which reflect this reduction in captured variance.[175] PCA has been applied to the bacterial classification problem in the past by many investigators.[60,89,145] Since PCA is described in lower dimensions and visualized in 2 dimensions (usually by the first 2 eigenvalues), hierarchical cluster analysis (HCA) can be more robust in classification than PCA by including more dimensions in the multivariate space.[176] A distance tree can be constructed in Euclidean space or with Mahalanobis distance. In this manner, more of the variance can be used to describe the classes.

3.4 Partial Least Squares Discriminant Analysis (PLSDA)

Multivariate classification methods are chemometric techniques aimed at finding mathematical models able to recognize the membership of each sample to its appropriate class, on the basis of a set of measurements. Once a classification model has been calibrated, the membership of unknown samples to one of the defined classes can be predicted. Therefore, classification techniques (also known as supervised pattern recognition) handle qualitative responses, that is, they define mathematical relationships between a set of descriptive variables (e.g. chemical measurements) and a qualitative variable, i.e. the membership to a defined category.

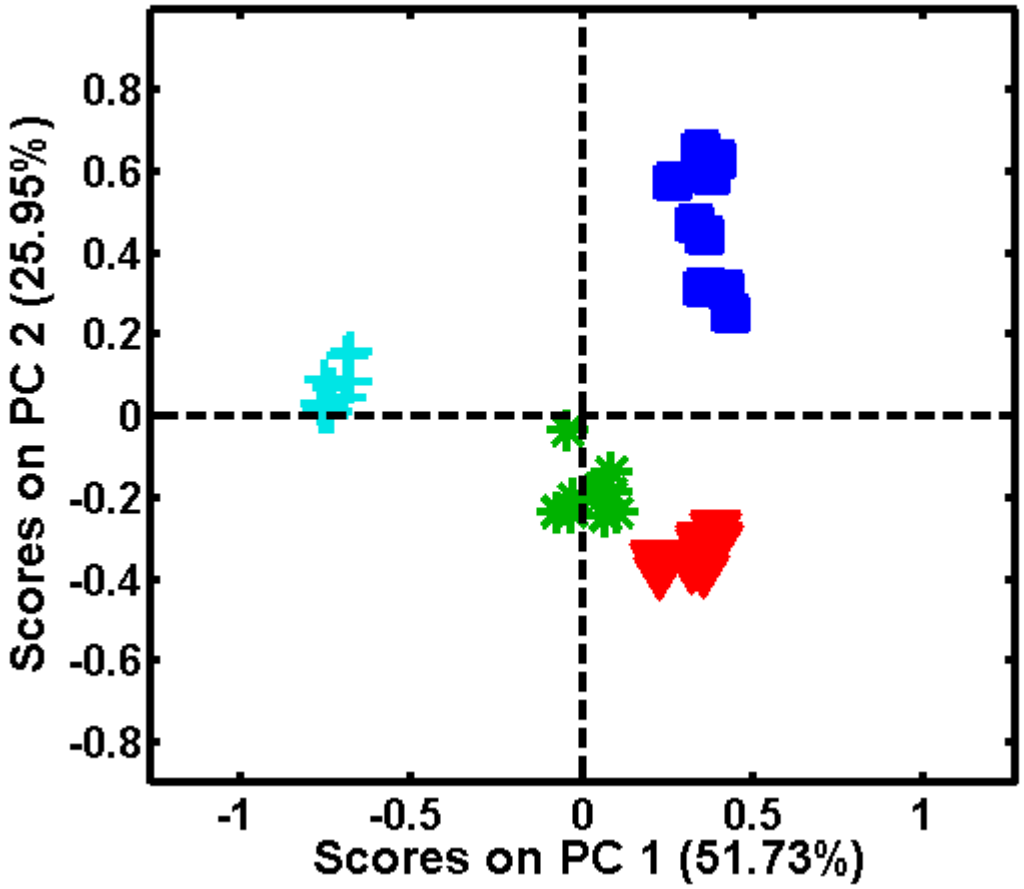

Figure 3.1 Example of a scores plot of components 2 vs 1 of a series of bacterial strains.

However, when *a priori* knowledge is available for the dataset, such as known classes, then a more robust approach is to use partial least squares for discriminatory analysis (PLS-DA).[177] An often used data analysis tool for classification in the metabolomics area,[178,179] chemical analytical technology as LCMS, GCMS and NMR.[180,181] PLSDA is a linear classification method that combines the properties of partial least squares regression with the discrimination power of a classification technique. PLS-DA is based on the PLS regression algorithm(PLS1 when dealing with one dependent Y variable and PLS2 in the presence of several dependent Y variables), which searches for latent variables with a maximum covariance with the Y-variables.[175,182] The main advantage of PLS-DA is that the relevant sources of data variability are modelled by the so-called Latent Variables (LVs), which are linear combinations of the original variables, and, consequently, it allows graphical visualization and understanding of the different data patterns and relations by LV scores and loadings. Loadings are the coefficients of variables in the linear combinations which determine the LVs and therefore they can be interpreted as the influence of each variable on each LV, while scores represent the coordinates of samples in the LV projection hyperspace.

Cross validation is often used for validation of a classification model due to the low number of samples available as separation into training, validation and test set is often not possible, cross validation makes better use of the data. However, cross validation only gives a reliable error rate when the complete modelling procedure is cross validated. The object that is predicted should in no way be used in the development of the model.[183-185] For a proper cross validation, the total data should be divided into a training set, a validation (sometimes called optimization) set and a test set. Using the validation and training set a

model is developed and optimized. The test set is only used to test the models performance. By repeating the procedure in a way that each sample appears once and only once in the test set, the prediction error is representative for new samples. For a complete independent test set, it should also not be used in data pre-treatment and pre-processing, scaling etc.[182] Each time, the model is calibrated on the remaining training samples and then used to predict samples of the cross validation group. Samples are usually divided in cross validation groups on the basis of two procedures: venetian blinds or contiguous blocks. When dealing with classification, the choice of the suitable type of cross validation groups basically depends on how classes are distributed among samples. In this manner, a model's usefulness can be assessed by comparing root mean squared cross validated errors (RMSECV) to other models.[186]

When cross validating models, another relevant issue is the choice of a reasonable number of cross validation groups. Another one of the common practices is to perform the leave-one-out (LOO) validation, which corresponds to the selection of a number of cross validation groups equal to the number of samples, that is, just one sample at a time is included in the validation set.[187] However, LOO usually overestimates the predictive power of a model and thus cannot return a reliable number of latent variables. Consequently, good estimates provided by LOO appear to be necessary but not a sufficient condition to have a high predictive power. When the dataset comprises a few samples, a higher number of groups could be preferred in order to get most of the samples in the training set and a few samples selected in each validation group; thus, a small perturbation is produced on the model.

3.5 Support Vector Machines (SVM)

An alternative chemometric approach is the use of support vector machine algorithms. These are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis.[188,189] Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other, making it a non-probabilistic binary linear classifier. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on. In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using what is called the kernel trick,[190] implicitly mapping their inputs into high-dimensional feature spaces. As compared to other machine learning approaches, SVM has several advantages. Firstly, for data classification it obtains unique global solution rather than local. Secondly, it maximizes the generalization capability by minimizing the training error of data samples.[191]

A support vector machine constructs a hyperplane or set of hyperplanes in a high- or infinite-dimensional space, which can be used for classification, regression, or other tasks.[192] Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training data point of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier as shown in Figure 3.2.

Whereas the original problem may be stated in a finite dimensional space, it often happens that the sets to discriminate are not linearly separable in that space. For this reason, it was proposed that the original finite-dimensional space be mapped into a much higher-dimensional space, presumably making the separation easier in that space. To keep the computational load reasonable, the mappings used by SVM schemes are designed to ensure that dot products may be computed easily in terms of the variables in the original space, by defining them in terms of a kernel function ***k(x,y)*** selected to suit the problem.[191] The hyperplanes in the higher-dimensional space are defined as the set of points whose dot product with a vector in that space is constant. The vectors defining the hyperplanes can be chosen to be linear combinations with parameters α_i of images of feature vectors that occur in the data base. With this choice of a hyperplane, the points *x* in the feature space that are mapped into the hyperplane are defined by the relation:

$$\sum_i \alpha_i k(x_i x) = constant$$

Note that if ***k(x,y)*** becomes small as ***y*** grows further away from ***x***, each term in the sum measures the degree of closeness of the test point ***x*** to the corresponding data base point $\boldsymbol{x}_i$. In this way, the sum of kernels above can be used to measure the relative nearness of each test point to the data points originating in one or the other of the sets to be discriminated. Note the fact that the set of points ***x*** mapped into any hyperplane can be quite convoluted as a result, allowing much more complex discrimination between sets which are not convex at all in the original space.

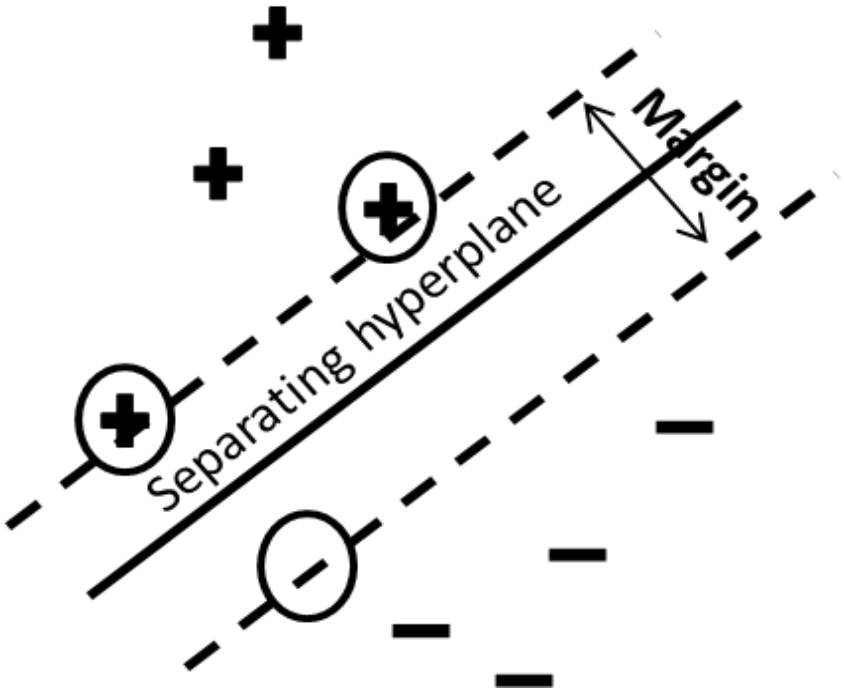

Figure 3.2 Maximum-margin hyperplane and margins for an SVM trained with samples from two classes. Samples on the margin are called the support vectors.

Similar to the PLSDA models, training sets are used to calibrate the SVM-DA classification models. Validation sets are usually used in the final steps to evaluate the true predictive ability of the calibrated model. A requisite of the validation set is that it should sufficiently represent the entire dataset in order to provide reliable estimates of the true model predictive ability. In most cases, this is accomplished by rationally-guided algorithms able to select those samples that are most representative of all the regions of the dataspace.

3.6 Calibration and Validation

For any spectroscopy technique, such as Raman, SERS and NIR spectroscopy, multivariate calibration (MVC) is defined as "A process for creating a model 'f' that relates sample properties 'y' to the intensities or absorbance 'X' at more than one wavelength or frequency of a set of known reference samples". [193] The development of the regression models comprises of the following three stages: (i) the calibration model is built and validated using a training set (X0, y0) and a validation set (X1, y1); the result is an error of validation having an associated standard error of validation (SEV), that is used to configure the model, (ii) both (X0, y0) and (X1, y1) are used to compute the standard error of calibration (SEC) of the model, and (iii) an independent test set (Xp, yp) is used to evaluate the model's performance with an indicator criterion, namely the error of prediction, where the standard error of prediction (SEP) is utilized.

Generally, the first and second steps are merged together using the cross-validation technique (e.g., leave-one-out (LOO) method, contiguous blocks, randomization, venetian blinds or the bootstrap), so the standard error of calibration (SEC) and the standard error

of validation (SEV) are computed simultaneously. In this case, Spectra **X** and the related sample properties "y" are split into calibration and prediction subsets separately. The calibration data usually comprised of between 50 and 75% of the total data set and include the smallest and largest "y", with the remaining data partitioned randomly into the calibration and prediction sets. The efficiency of a model approximation for a set of calibration and prediction samples can be reported as standard error of calibration (SEC), the root mean square error of cross-validation (RMSECV), the correlation coefficient (r) and standard error of prediction (SEP). These coefficients

are computed as follows:

$$SEC = \sqrt{\frac{1}{I_c - 1 - h}\sum_{i=1}^{I_c}(\hat{y}_i - y_i)^2}$$

$$SEP = \sqrt{\frac{1}{I_p - 1}\sum_{k=1}^{I_p}(\hat{y}_k - y_k)^2}$$

$$RMSECV = \sqrt{\frac{1}{I_c}\sum_{i=1}^{I_c}(\hat{y}_i - y_i)^2}$$

$$r^2 = 1 - \frac{\sum_{g=1}^{n}(y_g - \hat{y}_g)^2}{\sum_{g=1}^{n}(y_g - \bar{y})^2}$$

To evaluate the error of each calibration model, root mean square error of cross-validation (RMSECV) method is used.

3.7 Statistical procedures to measure spectral variability

3.7.1 Spectral Variability

Aiming to assess within- and among-species spectral variability, the metric D,[194] should be calculated as follows:

$$\boldsymbol{D} = \left[\frac{\mathbf{1}}{\boldsymbol{\lambda_b} - \boldsymbol{\lambda_a}} \int_{\boldsymbol{\lambda_a}}^{\boldsymbol{\lambda_b}} [\boldsymbol{S_1}(\boldsymbol{\lambda}) - \boldsymbol{S_2}(\boldsymbol{\lambda})]^2 \boldsymbol{d\lambda}\right]^{\mathbf{1/2}}$$

where D corresponds to the root mean square difference between a pair of spectra (S_1 and S_2), averaged over the spectral interval (λ_a to λ_b). The Raman wavenumber interval used to calculate this metric for our experiments was 1800 to 400cm^{-1} as will be showed in a later chapter.

3.7.2 Tukey's Range Test

Tukey's range test is a single-step multiple comparison procedure and statistical test. It is used in conjunction with an ANOVA to find means that are significantly different from each other. For spectroscopy purposes, One-way ANOVA is used to verify the statistical difference between species means in each waveband. The ANOVA tested the following hypothesis:

$$H_0 = \mu_1 = \mu_2 = \cdots = \mu_n$$

$$H_1 = Not\ all\ \mu_n\ are\ equal$$

where μ_n represents the mean reflectance of the *n*'th species (n = 1, 2,…, 7) and *i* denotes the waveband. Rejection of the null hypothesis (H_0) indicated the wavebands, at a 99% (p-value <0.01) confidence level, in which the species differ statistically. H_0 rejection was followed by pairwise multiple comparisons with the post-hoc Tukey HSD test. The total number of pair combinations is calculated as $n(n-1)/2$, where n is the number of species. By counting the number of pairs that is statistically significantly different on each waveband, it is possible to identify the spectral regions where the species most differ.[194] Applying Tukey's test compares all possible pairs of means, and is based on a studentized range distribution (q) (this distribution is similar to the distribution of t from the t-test). The

Tukey HSD tests should not be confused with the Tukey Mean Difference tests (also known as the Bland-Altman Test).

Tukey's test compares the means of every treatment to the means of every other treatment; that is, it applies simultaneously to the set of all pairwise comparisons

$$\mu_i - \mu_j$$

and identifies any difference between two means that is greater than the expected standard error. The confidence coefficient for the set, when all sample sizes are equal, is exactly $1 - \alpha$. For unequal sample sizes, the confidence coefficient is greater than $1 - \alpha$. In other words, the Tukey method is conservative when there are unequal sample sizes. This method assumes; (1) the observations being tested are independent; and (2) there is equal within-group variance across the groups associated with each mean in the test (homogeneity of variance).

3.7.3. Feature Selection

Variable and feature selection have become the focus of much research in areas of application for which datasets with tens or hundreds of thousands of variables are available. These areas include text processing of internet documents, gene expression array analysis, and combinatorial chemistry. The complexity of the spectra obtained from a biological sample makes extracting relevant information and interpreting the data challenging. The most primitive data analysis procedures used for Raman studies such as peak-by-peak analysis and peak deconvolution do not allow for extensive data extraction and are a major limitation, often involving tedious, error-prone manual analysis procedures. Such methods do not allow for complete data extraction, often only using a very limited subset of data.[195] Data mining and machine learning techniques are able to circumvent these pitfalls by

optimizing data extraction, exposing obscured correlations and reducing classification error. Computational experiments indicate that supervised classification algorithms such as Support Vector Machines (SVM) and Linear Discriminant Analysis (LDA) are able to separate the high dimensional spectral data with high accuracy.[196] [197,198] These algorithms can be used in combination with dimensionality reduction or feature selection techniques in order to identify the critical regions and bands of the spectrum that allow for discrimination and classification of cell types, cellular processes and cell response to various stimuli based variations in biochemical composition.

In machine learning and statistics, feature selection, is the process of selecting a subset of relevant features for use in model construction. The central assumption when using a feature selection technique is that the data contains many redundant or irrelevant features. Redundant features are those which provide no more information than the currently selected features, and irrelevant features provide no useful information in any context. Feature selection techniques are a subset of the more general field of feature extraction. Feature extraction creates new features from functions of the original features, whereas feature selection returns a subset of the features. Viewed from a multivariate calibration perspective, feature or variable selection attempts to identify and remove the variables that penalize the performance of a model, since they are useless, noisy and redundant or correlated by chance.

A fundamental problem of machine learning is to approximate the functional relationship $f(\,)$ between an input $X= \{x_1, x_2, \ldots, x_m\}$ and an output Y, based on a memory of data points, $\{X_i, Y_i\}$, i = 1, ..., N, usually the X_i's are vectors of reals and the Yi's are real numbers. Sometimes the output Y is not determined by the complete set of the input features

$\{x_1, x_2, \ldots, x_M\}$, instead, it is decided only by a subset of them $\{x_{(1)}, x_{(2)}, \ldots, x_{(m)}\}$, where $m<M$. With sufficient data and time, it is fine to use all the input features, including those irrelevant features, to approximate the underlying function between the input and the output. But in practice, there are two problems which may be evoked by the irrelevant features involved in the learning process.[199] The irrelevant input features will induce greater computational cost. Apparently, with more features, the computational cost for predictions will increase polynomially; especially when there are a large number of such predictions, the computational cost will increase immensely. The irrelevant input features may lead to overfitting.

CHAPTER 4

Goals of this dissertation

4.1 Spectroscopic Analyses of Lipid Profiles of Mycobacteria

The bacterium *Mycobacterium tuberculosis*(TB), the primary causative agent of tuberculosis, infects one-third of the world's population[1]. Mycobacterial lipids are known to facilitate persistence during infection within the host. The biosynthesis of mycolic acids has been investigated for a number of reasons, but primarily because of the presumption that enzymes involved in the synthesis of these lipids are attractive targets for the development of novel therapeutic agents.[17-19] This work will examine lipophilic extract containing mycolic acids isolated from tuberculosis (MTB) and non-Tuberculosis (NTM) mycobacterial strains using chromatography, Mass spectrometry (MS), nuclear magnetic resonance(NMR), and Raman spectroscopy.[92] ^{1}H-NMR shows the presence of cis and trans cyclopropane rings within the different mycolic acid types. SERS studies with our OAD fabricated Ag nanorods arrays are going to be used to differentiate the saponified mycolic acid profiles in the bacterial envelope of tuberculous and non-tuberculous species. Multivariate statistical methods such HCA, PCA and PLS-DA analyses of the SERS spectra of the mycolic acid profiles should enable rapid differentiation from each species with high accuracy. These methods could be used as a safe, rapid and accurate method for mycobacterial species and strain discrimination, these results will be shown in chapter 5.

4.2 Layer-by-Layer Encapsulation of *Mycoplasma pneumoniae* for Enhanced Raman Detection

M. pneumoniae is the leading cause of pneumonia in older children and adults. The main cytotoxic effect of *M. pneumoniae* is local disruption of tissue and cell structure along the respiratory tract epithelium due to its close proximity to host cells. The organism is not known to produce any exotoxins, but formation of hydrogen peroxide is a key virulence factor in *M. pneumoniae* infections. Diagnosis of acute infections remains difficult; therefore, early recognition of outbreaks has been problematic. Basic diagnostic strategy in clinical practice includes serology and standard polymerase chain reaction (PCR). New diagnostic techniques (PCR-related methods) may enable more rapid diagnosis. PCR can exhibit high sensitivity and positive detection, but is prone to false negatives from reaction inhibitors. Here we will develop a new detection platform with high sensitivity, specificity and expediency to detect mycoplasma. As proof method studies we report the layer-by-layer (LBL) coating of *M. pneumoniae* cells with polyelectrolytes poly(allylamine hydrochloride) (PAH) and poly(styrene sulfonate)(PSS) as our first study[37] discussed in chapter 6. Followed by a separate study where we coated mycoplasma cells with poly(diallyldimethylammonium chloride)(PDADMAC) and PSS shown in chapter 7. In both studies silver nanoparticles were deposited after the consecutive placement of the polyelectrolytes on the surface of the mycoplasma cells. The SERS spectra paired with chemometric methods with the platform described should be able to detect and differentiate *M. pneumoniae* strains. The technique shows great promise in its potential to improve diagnosis of *M. pneumoniae* infections, through the detection of real clinical specimens at high sensitivity.

Based on the last remark this led us to the last project, in which we employed both LBL techniques shown in chapter 6 and chapter 7 on human commensal clinical strains. Our LBL SERS analyses of our first studies with mycoplasma cells yielded around 85-100% specificity and sensitivity due to when we tried PLSDA on limited sample size(e.g. 4 mycoplasma strains and a control) . We anticipated the same results when a series of 13 human commensal strains and type I M129 strain SERS spectra were analyzed for sensitivity using our chemometric techniques that have been employed in analyzes for years. For this project we dug a bit further with our chemometric analyzes. We employed such statistical techniques aiming to assess the spectral variability of the different mycoplasma species. We were interested in identifying regions of the Raman spectrum in which the species most differ from each other. For this purpose, we first conducted a one-way analysis of variance (ANOVA) followed by the post-hoc Tukey Honestly Significance Difference (HSD) test across each wavenumber of the SERS spectra of each species, in order to verify differences among species means. We extract the information obtained from these methods and compare the PLSDA results obtained with those of obtained using the wavenumbers of the whole SERS spectrums of the commensals strains so we finally compare which technique gave us the optimal results.

CHAPTER 5

Identification of Mycobacteria Based on Spectroscopic Analyses of Mycolic Acid Profiles

Rivera-Betancourt, O.E., Karls, R., Grosse-Siestrup,B., Helms,S., Quinn,F.,Dluhy,R.A., 2013, *Analyst*, 138, 6774-6785

Reprinted here with permission of The Royal Society of Chemistry

5.1 Abstract

This report examines lipophilic extracts containing mycolic acids isolated from tuberculosis (MTB) and non-tuberculosis (NTM) mycobacterial strains using chromatography, mass spectrometry (MS), nuclear magnetic resonance (NMR), and Raman spectroscopy. Gas chromatography-MS was used to identify major fatty acid mycolate components, while proton NMR confirmed the presence of characteristic *cis*- and *trans*-cyclopropane rings within different mycolic acid sub-types. Surface-enhanced Raman (SERS) spectra were obtained from the mycolic acids extracted from the bacterial cell envelopes of the MTB or NTM mycobacterial species. The Raman spectral profiles were used to develop a classification method based on chemometrics for identification of the mycobacterial species. Multivariate statistical analysis methods, including principal component analysis (PCA), hierarchical cluster analysis (HCA), and partial least squares discriminant analysis (PLS-DA) of the SERS spectra enabled differentiation of NTM mycobacteria from one another with 100% accuracy. These methods are also sensitive enough to differentiate clinically-isolated MTB strains that differed only by the presence or absence of a single extracytoplasmic sigma factor with 83 – 100% sensitivity and 80 – 100% specificity. The current work is the first report on discrimination of mycobacteria strains based on the SERS spectra of the constituent mycolic acids in lipophilic extracts. These results suggest that SERS can be used as an accurate and sensitive method for species and strain discrimination in mycobacteria.

5.2 Introduction

Mycobacterium tuberculosis (MTB), the bacterial agent responsible for tuberculosis (TB), claimed approximately 1.4 million lives in 2012.[25] The problem is exacerbated by the development of multiple-drug resistant strains of *M. tuberculosis*. The unique cell wall of *M. tuberculosis* contributes to the malignancy of this pathogen. The mycobacterial cell wall contains three covalently-linked layers outside the cytoplasmic membrane: a: peptidoglycan matrix, an arabinogalactan matrix, and an outer membrane consisting of long lipids, such as mycolic acids.[200]

Routine identification and diagnosis of MTB relies on a combination of methods, including such insensitive methods as sputum smear microscopy and time-consuming culturing of the MTB bacilli.[201] Molecular-based methods of TB diagnosis, including tests based on amplification of MTB DNA, have been reported.[202] However, these tests have been criticized, and face obstacles of logistics, sensitivity, and cost.[203] To date, high performance liquid chromatography (HPLC) of extracted mycolic acids remains a standard method for identification of MTB species.[204]

Mycolic acids are long-chain α-alkyl-branched, β-hydroxylated fatty acids present in mycobacteria and related micro-organisms. While components of the inner leaflet of the outer membrane are covalently-bonded to the arabinogalactan polymer, outer leaflet components such as mycolates are extractable with organic solvents.[205]

Chemical side groups on the main meromycolic chains on the mycolic acids differentiate the mycolate subclasses. The R-mycolic acid type, which is present in all mycobacteria, is composed of C_{70}-C_{76} fatty acids and contains two sites of saturation (either *cis* cyclopropane rings or one or both of *cis* and *trans* double bonds), and no additional

oxygenated groups.[206] Subtle variations in the mycolic acid structure have profound effects on the physiology and virulence of the different mycobacterial species.

While HPLC has proven to be a valuable method of NTM and MTB identification, it suffers from the need for fluorescent labeling of the extracted mycolic acids for UV detection and lengthy chromatographic run times. Raman spectroscopy is a potentially fast and sensitive test for the detection of lipids unique to *M. tuberculosis*. The high molecular specificity provided by the Raman spectrum provides a means to identify molecular fingerprints unique to the pathogen. We have previously shown that surface enhanced Raman spectroscopy (SERS) enables rapid differentiation and classification of both viral and bacterial pathogenic species and strains.[207]

In this work we employed silver nanorod (AgNR)-based SERS[208] to detect the chemical signatures found in the mycolic acids extracted from the cell envelopes of tuberculous (MTB) and non-tuberculosis (NTB) mycobacterial species. GC-MS and NMR confirmed the identity of the mycolic acid species isolated from the different mycobacterial species. Multivariate statistical analysis methods, including principal component analysis (PCA), hierarchical cluster analysis (HCA), and partial least squares discriminant analysis (PLS-DA) were employed to identify and classify the MTB and NTB profiles based on the SERS spectra of their lipophilic extracts. In addition, we showed the SERS/PLS-DA methods are sensitive enough to differentiate MTB strains based on the presence or absence of a single gene encoding the extracytoplasmic function sigma factor (*sigC*).

5.3 Materials and Methods

Bacterial strains and growth conditions

The cultures examined in this study included the following non-tuberculosis mycobacteria: 1) *Mycobacterium smegmatis* strain mc^2155, 2) *Mycobacterium bovis* BCG, and 3) *Mycobacterium avium* subspecies *avium* serovar #8. The tuberculous strains used here included two separate types of *M. tuberculosis* strains. The first three MTB strains were derived from a lab-passaged clinical strain and include: 1) H37Rv, 2) RvΔsigC, H37Rv with an internal in-frame *Eco*NI-*Xho*I deletion within *sigC*(17), and 3) RvΔCcomp, RvΔsigC complemented with *sigC* encoded on chromosome-integrating plasmid pMV306. An additional three MTB strains were derived from a clinical isolate, and include 1) Erdman, 2) EΔsigC, Erdman with an internal in-frame *Eco*NI-*Xho*I deletion within *sigC* (Grosse-Siestrup, unpublished), and 3) EΔCcomp, Erdman complemented with *sigC* encoded on plasmid pMV306.

Each *Mycobacterium* species was cultured on Lowenstein-Jensen medium at 37°C and 5% CO_2 to obtain equivalent-sized biomass colonies. For fast-growing species, *M. avium* was cultured for 1 day, and *M. smegmatis* cultures were incubated for 3-4 days. For slow-growing species *M. bovis* BCG was incubated for 14 days, while the *M. tuberculosis* H37Rv and Erdman strains both were incubated for 28-30 days. In each case, a loopful of cells (approximately 0.1 g) was scraped from the culture plate for lipid analysis.

Isolation of the lipophilic extracts from mycobacteria

Lipids were extracted using established procedures.[209] Briefly, 2.0 ml of a 20% methanolic KOH solution was added to a new test tube. The bacteria were added to the tube and mixed vigorously for at least 20 sec. The tubes were placed in a water bath at 80°C for 30 minutes.

Afterwards the tubes were autoclaved for 30 minutes at 121°C and then cooled to room temperature. Next, 2.0 ml of chloroform was added to the tubes followed by the addition of a solution of 50% HCl. The tubes were centrifuged for 5 minutes at 320×g for proper phase separation. The organic phase containing the mycolic acids was transferred to a new test tube.

Preparation of the mycolic acids methyl esters (MAME) for GC-MS analysis

The chloroform phase from the lipid extracts was removed by evaporation with the aid of a heating block at 80°C. After cooling to room temperature, mild methanolysis was performed as previously described.[210] Briefly, the sample in a screw-capped glass test tube was dissolved in 0.2 ml of toluene. To the solution, 1.5 ml of methanol and 0.3 ml of an 8.0% HCl solution were added. The sample was mixed by vortexing and incubated at 45°C overnight (14 h or longer) for mild methanolysis/methylation. After cooling to room temperature, 1.0 ml of hexane and 1.0 ml of water were added for extraction of the MAME. After mixing by vortexing, the hexane layer was recovered and analyzed by GC-MS.

No commercial synthetic standards are available for mycolic acids. However, a heterogeneous preparation of mycolic acid methyl esters were obtained from BEI Resources, NIAID, NIH (Manassas, VA). These MAME's were purified from the mycolyl-arabinogalactan peptidoglycan (mAGP) complex of *M. tuberculosis* and consist of three fractions representing the α-, methoxy-, and keto-mycolic acid methyl esters. This sample was used in the GC-MS experiments to verify retention times and m/z peak values for the mycolic acids contained in the lipophilic extracts isolated from the MTB and NTM samples.

GC-MS parameters

Analyses were performed in the splitless mode by injecting 1 µL of the hexane extract into a Hewlett-Packard (Agilent Technologies, Santa Clara, CA) 5890 gas chromatograph equipped with an EZ5 Econocap capillary column (30 m × 0.25 mm id, 0.25 µm film thickness, Altech, Deerfield, IL). The conditions were similar to those previously described,[211] with a slight modification. The injector temperature was held constant at 250°C for the entire run time. The initial GC oven temperature was held at 50°C for 1 min after injection, followed by an increase of 10°C min^{-1} to 240°C. The oven temperature was then increased at a rate of 20°C min^{-1} to a final temperature of 290°C and maintained for 7 min. Helium was used as carrier gas and the pressure programmed such that the helium flow was kept constant at 1.2 ml min^{-1}. Detection was via an HP 5971A mass spectrometer equipped with an electron ionization source and a quadrupole mass selector. Detection was achieved by using MS detection in full scan mode (m/z 50-500). For library matching, the acquired GC-MS spectra were searched against the 2008 Wiley/NIST EI GC-MS database.

Structural analysis using ^{1}H-NMR spectroscopy

^{1}H-NMR spectra of the underivatized mycolic acids were acquired on a Varian (Agilent Technologies, Santa Clara, CA) Unity Inova 500 MHz NMR spectrometer in $CDCl_3$. Chemical shifts were referenced to the chloroform peak.

Fabrication of SERS-active substrates

Aligned Ag nanorods used as SERS-active substrates were prepared by an oblique-angle vapor deposition (OAD) technique. Ag nanorod substrates were fabricated using a custom-designed electron-beam/sputtering evaporation system; Ag deposition and substrate preparation steps were identical to those used in previously published procedures.[208,212,213]

SERS measurements of mycobacterial lipids

SERS spectra were acquired using a Renishaw (Hoffman Estates, IL) inVia confocal Raman microscope system using a 785 nm near-IR diode laser as the excitation source. Radiation from the diode laser was attenuated to ~15 mW using a series of neutral density filters and focused onto the sample using a 20× microscope objective. Spectra were collected between 3200 – 500 cm^{-1} and integrated for 10 s per scan with 3 scans per spectrum. The SERS spectrum of each mycolic acid were collected applying a 1 µl sample droplet to an OAD fabricated substrate and allowed to dry prior to the acquisition of the spectra. A minimum of twelve spectra was collected for each bacterial strain from different locations on each individual substrate as well as on different substrates.

Classification using chemometric analysis

Raman spectra were imported into GRAMS AI (Version 8.0 Thermo Electron Corp, Waltham, MA) for spectral averaging and baseline correction. Chemometric analysis was carried out with MATLAB version 7.2 (The Mathworks, Inc., Natick, MA), using the PLS Toolbox version 4.1 (Eigenvector Research Inc., Wenatchee, WA). SERS spectra in the range 1700 – 600 cm^{-1} were used for classification. Prior to analysis, first derivatives of the SERS spectra were calculated using the Savitzky-Golay method with a 2^{nd} order polynomial and a fifteen-point window. Each data set was then vector normalized and mean centered. Multivariate statistical analysis of the extracted mycobacterial lipids was performed using principal components analysis (PCA), hierarchical cluster analysis (HCA), and partial least squares discriminant analysis (PLS-DA) using the PLS Toolbox software. The calculated principal components were used as inputs to the HCA algorithm, which used the *k*-nearest neighbor (KNN) algorithm to evaluate inter-cluster distances.

5.4 Results and Discussion

GC-MS Analysis of the mycolic acid methyl esters

A total of nine *Mycobacterium* strains belonging to four species were analyzed. Three of the *Mycobacterium* strains were non-tuberculous mycobacteria species, *i.e. M. smegmatis*, *M. bovis* BCG, and *M. avium* subspecies *avium*. Three of the *M. tuberculosis* strains were related to the lab-passaged clinical strain H37Rv, and included i) the H37Rv strain, ii) RvΔsigC – the H37Rv strain with an in-frame deletion within the *sigC* transcription factor, and iii) RvΔCcomp – the H37Rv deletion mutant strain complemented with *sigC*. The final three *M. tuberculosis* strains were related to the clinical Erdman strain, and included i) the Erdman strain, ii) EΔsigC – the Erdman strain with an in-frame deletion within the *sigC* transcription factor, and iii) EΔCcomp – the Erdman deletion mutant strain complemented with *sigC*. Lipids were extracted from each culture and mycolic acid methyl esters (MAME) prepared as described in the Methods section.

Each sample was subjected to GC-MS analysis. Fatty acid cleavage products from the mycolic acids are formed in the injection port of the gas chromatograph at temperatures higher than 235°C,[210,214]resulting in clear differences in the chromatograms. The mass spectra generated were identified using spectral library matching based on GC retention times, MS m/z values, and area percentage. If the overall match quality was above 90%, the mass spectra were identified as a fatty acid methyl ester. Representative GC-MS total ion chromatograms of the mycolic acids extracted from the four mycobacteria under study here are presented in figure A1.1 on Appendix 1.

Table 5.1 presents a list of the retention times of the most abundant compounds detected in the GC-MS total ion chromatograms of the lipids extracted from the mycobacterial

species. Similar to previous reports,[208,210,214,215] we observed that the most abundant fatty acids detected in the mycolic acids methyl esters of all mycobacterial species were hexadecanoic acid (C16:0) and oleic acid (C18:1). Another fatty acid that was also detected in all species was octadecanoic acid (C18:0), although this was lower in abundance than hexadecanoic or oleic acid. Other short chain fatty acids, such as tetradecanoic acid (C14:0), pentadecanoic acid (C15:0), heptadecanoic acid (C17:0) and nonadecanoic acid, were detected, although these were not present in every species. Long-chain fatty acid such as eicosanoic acid (C20:0), eicosatetranoic acid (C20:4), tetracosanoic acid (C24:0) and hexacosanoic acid (C26:0). Like the shorter chain fatty acids, the long chain fatty acids were detected in some but not all of the species.

^{1}H-NMR analysis

Mycolic acids derived from mycobacteria are characterized by high molecular mass 2-alkyl branched, 3-hydroxy fatty acids. A figure containing the common mycolic acid structures found in mycobacteria is included in Appendix 1 (figure A1.2). In the genus *Mycobacterium*, mycolic acids can be separated into two classes: 1) those containing oxygenated functional groups in the so called 'mero' chain, *i.e*. methoxy-, keto-, epoxy-, or carboxy-mycolates, and 2) those that do not, the α- and α'-mycolates.[216] Mycobacteria generally contain complex mixtures of α-mycolates, with two *cis* di-substituted cyclopropane rings, and oxygenated mycolic acids, with one di-substituted cyclopropane ring. Evidence suggests that both *cis* di-substituted cyclopropane α-mycolates and *trans* di-substituted oxygenated mycolic acids are implicated in some aspects of the tuberculosis pathogenesis.[217]

Table 5.1. Retention times of various lipid components found in the GC-MS chromatograms of the nine mycobacterial species/strains used in these studies.

Species		*M. smeg*	*BCG*	*M. avium*	*M. tuberculosis* H37Rv			*M. tuberculosis* Erdman		
Strains					Wild type	RvΔsigC	RvΔCcomp	Wild type	EΔsigC	EΔCcomp
Compound	m/z	Retention Times (min)								
C14:0	220			20.44						
C15:0	232								20.53	
C16:0	270	22.65	22.64	22.67	22.8	22.82	22.82	22.70	22.75	22.80
C16:1	237		22.45	22.44				22.49		
C17:0	284	23.66		23.11					23.75	
C18:0	298	24.62	24.62	24.63	24.75	24.77	24.77	24.67	24.70	24.75
C18:1	296	24.46	24.50	24.45	24.54	24.55	24.55	24.47	24.50	24.55
C18:2	298			24.38						
C19:0	312		25.04		25.09			25.00	25.03	
C20:0	326								26.20	
C20:4	318			25.74						
C24:0	382	28.79							28.88	
C26:0	410								30.70	

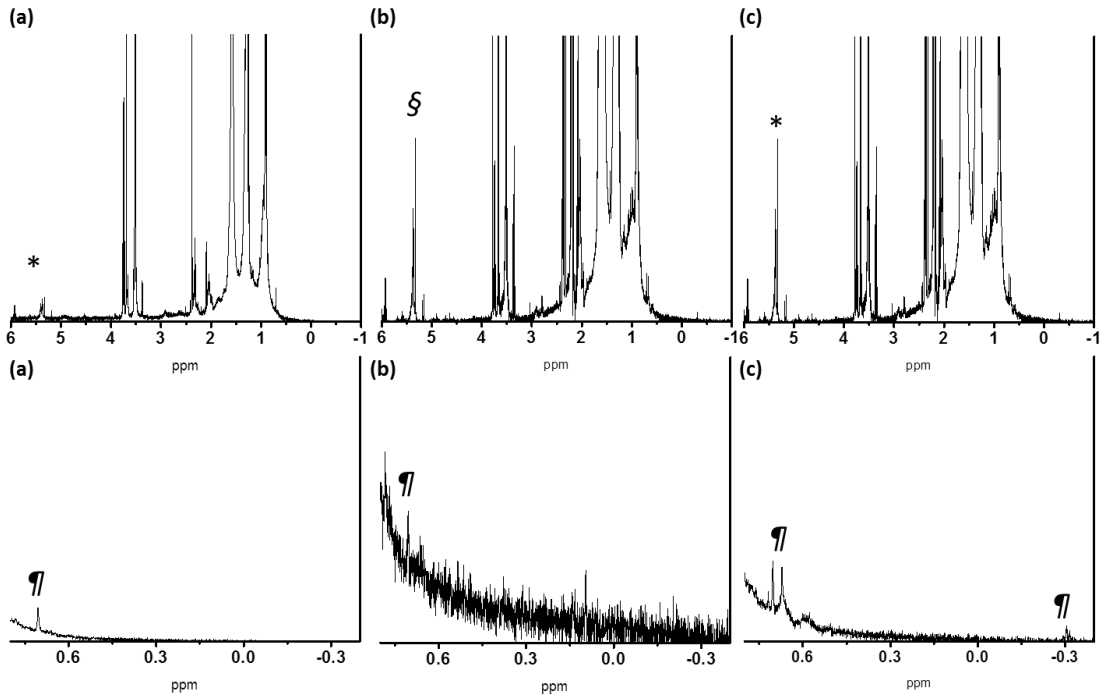

Figure 5.1 ^{1}H-NMR spectra of the NTM species. (a) *M. avium* ; (b) *M. smegmatis* ; (c) *M. bovis* BCG. Symbols indicate various resonances as follows: * = *trans* double bond; § = *cis* double bond; ¶ = *cis* cyclopropane ring.

Figure 5.1 shows the ^{1}H-NMR spectra of the NTM species, containing specific signals indicative of mycolic acid functional groups. All three species showed characteristic signals for either α- or keto-mycolates. All spectra show signals at 3.7 and 2.45 ppm, which correspond to a methoxyl from the carboxyl methyl ester and a methine, respectively. Chemical shifts arising from *trans* double bonds located at 5.39 and 5.35 ppm were found in the *M. avium* (figure. 5.1a) and *M. bovis* BCG (figure. 5.1c) spectra but not in the *M. smegmatis* spectra. In the *M. smegmatis* spectrum (figure 5.1b), a single peak at 5.3 ppm suggests the presence of a *cis*-double bond.

Figures 5.1a and 5.1b reflect only trace resonances of *cis*- and *trans*-cyclopropyl protons at 0.6 – 0.7 ppm from the *M. avium* and *M. smegmatis* spectra, indicating a predominance of keto-mycolate type 3 and 4 (k3, k4) subclasses with little α-mycolate structure. Figure 1c, the spectrum of BCG, shows increased signals at 0.66 ppm and -0.31 ppm, indicating the presence of *cis*-cyclopropane rings, and an α-mycolate type 1 (α1) subclass. These peaks, as well as the other *cis*-, *trans*-double bond peaks described above, also shows that BCG contains a keto-mycolate type 3 (k3) structure.[218]

A previous study used *M. tuberculosis* strains to identify an extra-cytoplasmic secondary sigma (*sigC*) transcription factor that coordinates expression of genes encoding functions that facilitate bacterial adaptation to stresses encountered inside a host.[215] Studies of knockout mutations of *sigC* in *M. tuberculosis* strains suggested that some sigma factors play significant roles in virulence, and that *sigC* mutants were less adept at survival inside guinea pig hosts.

We studied the NMR spectra of mycolic acids isolated from two separate MTB wild type strains, *i.e*. Erdman and H37Rv, as well as *sigC* knockout mutants and *sigC* complements

for both. Figure 5.2 presents ^{1}H-NMR spectra from the *M. tuberculosis* Erdman strain. Similar to the spectra of the NTB strains seen in figure.5.1, these spectra possess chemical shifts at 3.7 and 2.45 ppm, corresponding to a methoxyl and a methine. Chemical shifts of a *trans*-double bond located at 5.39 and 5.35 ppm were found in the EΔsigC *sigC* deletion mutant in figure.5.2c. The Erdman wild-type and EΔCcomp *sigC* complement spectra (figures. 5.2a and 5.2b) show a signal at 5.3 ppm suggesting the presence of a *cis*-double bond. The presence of *cis*-cyclopropyl (0.70 ppm and -0.31 ppm) and *trans*-cyclopropyl (0.65 and 0.10 ppm) resonances in figures. 5.2a–5.2c indicate that the MTB Erdman strains contain no α-mycolate structures.[219,220] The *trans*-cyclopropyl resonance at 0.10 ppm appears more intense in the wild-type Erdman (figure.5.2a) and EΔsigC *sigC* (figure.5.2b) mutant strains, while this resonance is much smaller in the EΔCcomp *sigC* complement (figures.5.2c).

We also obtained the ^{1}H-NMR spectra of *M. tuberculosis* H37Rv strains (figure.5.3). The spectra of the H37Rv wild-type and RvΔCcomp *sigC* complement in figuress. 5.3a and 5.3b show a *cis*-double bond signal at 5.3 ppm, while the double resonances at 5.39 and 5.35 ppm in the spectrum of the RvΔsigC *sigC* mutant (figure. 5.3c) indicate a *trans* double bond. All three spectra show peaks at 0.70 and -0.31 ppm indicative of *cis*-cyclopropyl protons, while only a weak resonance at 0.6 ppm indicates only small amounts of *trans*-cyclopropyl protons (figures. 5.3a – 5.3c). The *cis*- and *trans*-cyclopropyl protons resonances below 0.8 ppm indicate that these strains have no α-mycolate structures present. In summary, the NMR data confirmed the structure of mycolic acids extracted from NTM and MTB strains in these sample preparations. Only trace amounts of cyclopropane groups were found in the NTM species. The predominant mycolate classes found in both the NTM

and MTB strains were the keto-mycolates.[218] No proton resonances belonging to *trans*-epoxy or *cis*-epoxy mycolates were seen in either the Erdman or H37Rv strains.[220] The NMR spectra of both MTB mutant samples indicate that *cis*-cyclopropyl rings predominate; peaks attributable to *trans*-cycloproprane protons appear only in trace amounts.

Detailed assignments of the observed ^{1}H-NMR chemical shifts to specific mycolate structures are provided in Tables A.1, A.2, and A.3 in Appendix 1.

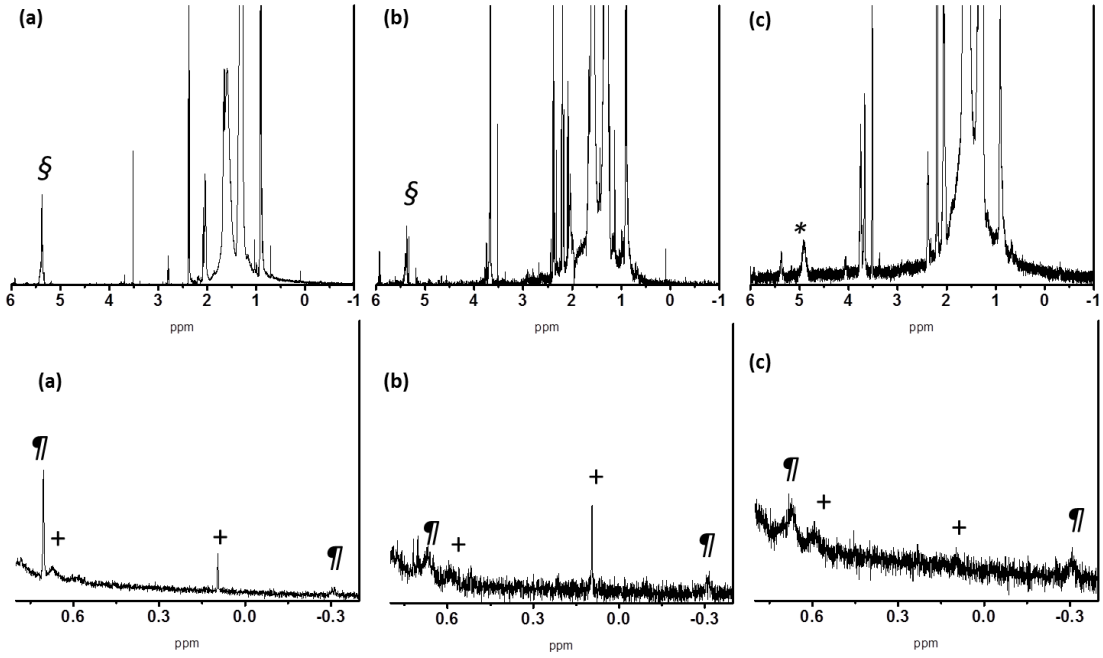

Figure 5.2 ^{1}H-NMR spectra of the *M. tuberculosis* Erdman strains. (a) Wild type; (b) EΔCcomp, *sigC* complement; (c) EΔsigC, *sigC* mutant. Top NMR spectra show the chemical shift region between -1 and 6 ppm; bottom NMR spectra show the details of the region between -0.4 and 0.8 ppm. Symbols indicate various resonances as follows: * = *trans* double bond; § = *cis* double bond; ¶ = *cis* cyclopropane ring; + = *trans* cyclopropane ring.

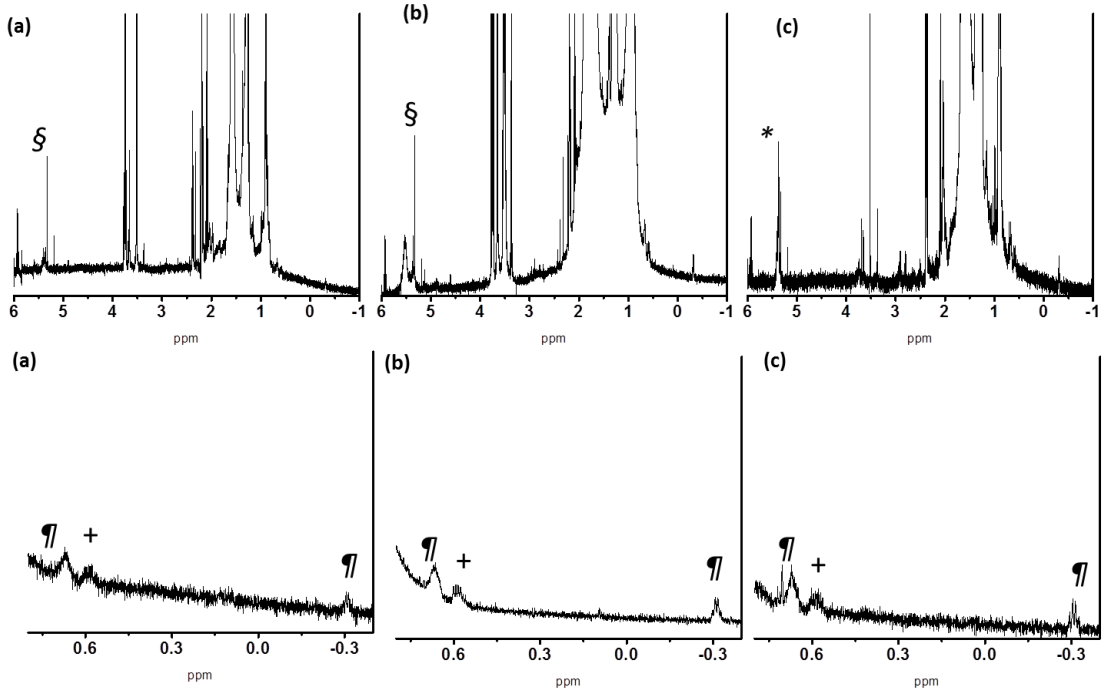

Figure 5.3 ^{1}H-NMR spectra of the *M. tuberculosis* H37Rv strains. (a) Wild type; (b) RvΔCcomp, *sigC* complement; (c) RvΔsigC *sigC* mutant. Top NMR spectra show the chemical shift region between -1 and 6 ppm; bottom NMR spectra show the details of the region between -0.4 and 0.8 ppm. Symbols indicate various resonances as follows: * = *trans* double bond; § = *cis* double bond; ¶ = *cis* cyclopropane ring; + = *trans* cyclopropane ring. Figure 3. ^{1}H-NMR spectra of the *M. tuberculosis* H37Rv strains. (a) Wild type; (b) RvΔCcomp, *sigC* complement; (c) RvΔsigC *sigC* mutant. Top NMR spectra show the chemical shift region between -1 and 6 ppm; bottom NMR spectra show the details of the region between -0.4 and 0.8 ppm. Symbols indicate various resonances as follows: * = *trans* double bond; § = *cis* double bond; ¶ = *cis* cyclopropane ring; + = *trans* cyclopropane ring.

SERS spectra of NTM and MTB species

We employed Ag nanorod-based surface-enhanced Raman scattering (SERS)[221] to detect the spectral differences in the profiles of mycolic acids extracted from the bacterial cell envelopes of NTM and MTB mycobacterial strains. Figure 5.4 illustrates the SERS spectra of the NTM species *M. avium* (figure. 5.4a), *M. bovis* BCG (figure. 5.4b), and *M. smegmatis* (figure. 5.4c). Spectra were acquired after 1.0 µl aliquot of each sample was applied to the SERS substrate, allowed to dry, and the signal integrated for 10 s with 3 co-added scans.

The spectra for these NTM species shown in figure.5.4 present common peaks at 1654, 1598, 1442, 1299, 1164, 1135 and 1001 cm^{-1}. In all three spectra the most intense bands were centered at 1442 and 1001 cm^{-1}, corresponding to CH_2 asymmetric bend and C-C-O out of phase stretching vibrations of a primary alcohol.[222-224] These are common functional groups found in the meromycolate chain.[205,218,225] The spectra of each individual NTM species shows characteristic signature bands below 1000 cm^{-1}, reflecting the different mycolate classes in each species.

The *M. tuberculosis* strains used in this study were the same as those used in the NMR studies described above, *i.e*., clinical isolates Erdman and H37Rv, the derivatives of each strain defective in production of sigma factor C, and the *sigC* mutants complemented with a wild-type copy of *sigC* to restore sigC production. The SERS spectra for the Erdman and H37Rv family of strains are shown in figures. 5.5 and 5.6, respectively. SERS spectra of the Erdman and H37Rv derivatives exhibited very similar features, such as the strong bands at ~1650, 1589, 1564, 1392, 1240, 1135, 1005 and 850 cm^{-1}. The pattern of these

vibrational bands is similar in all SERS spectra of the wild-type and *sigC* complemented derivatives of both Erdman and H37Rv strains.

Detailed assignments of the observed SERS bands to specific molecular vibrations are provided in Tables A.4, A.5 and A.6 in the ESI. In addition, a figure has been included in Appendix 1 of ten unprocessed SERS spectra of the MTB Erdman strain, illustrating the reproducibility of the raw data used in this analysis (Figure A1.3).

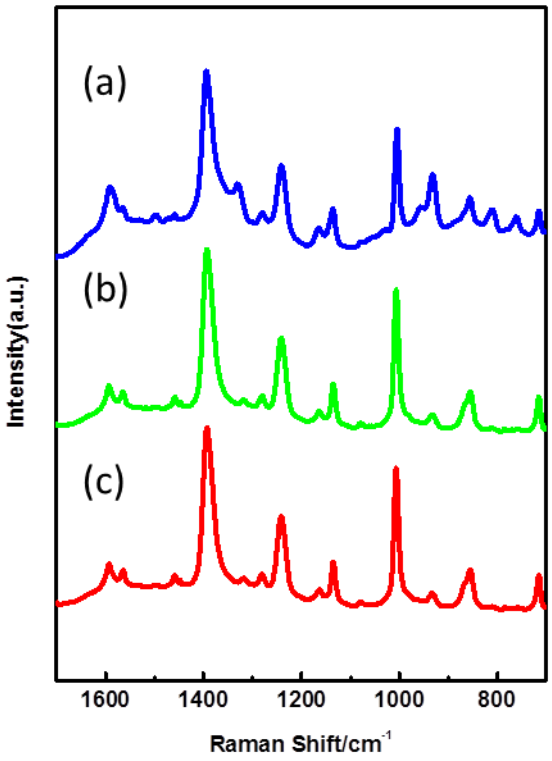

Figure 5.4 SERS spectra for the non-tuberculous mycobacteria samples. (a) *M. avium* ; (b) *M. bovis* BCG ; (c) *M. smegmatis*. Each spectrum is an average of 10 spectra obtained for the individual NTM species. The spectra have been baseline corrected and normalized for visualization.

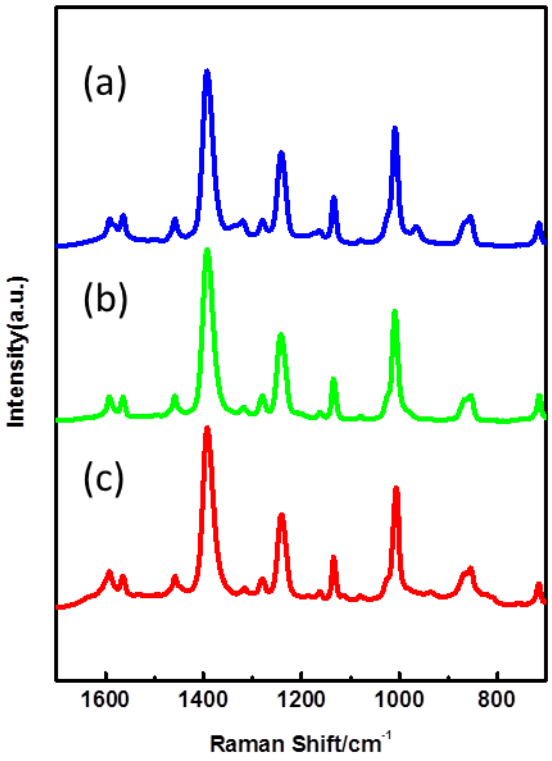

Figure 5.5 SERS spectra for the *M. tuberculosis* Erdman strain samples. (a) Wild type;(b) EΔCcomp, *sigC* complement; (c) EΔsigC, *sigC* mutant. Each spectrum is an average of 10 spectra obtained for the individual MTB Erdman strains. The spectra have been baseline corrected and normalized for visualization.

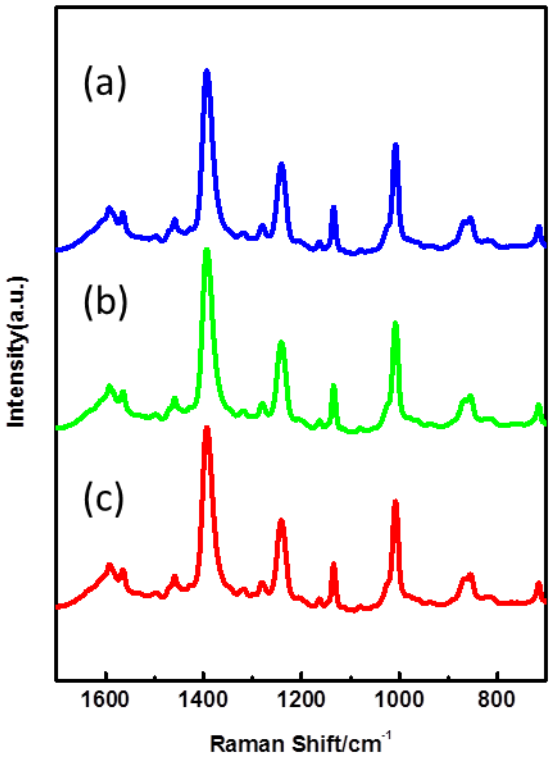

Figure 5.6 SERS spectra for the *M. tuberculosis* H37Rv strain samples. (a) Wild type; (b) RvΔCcomp, *sigC* complement ; (c) RvΔsigC, *sigC* mutant. Each spectrum is an average of 10 spectra obtained for the individual H37Rv strains. The spectra have been baseline corrected and normalized for visualization.

Classification of the Bacterial Species

The excellent spot-to-spot and substrate-to-substrate reproducibility of the SERS Raman spectra offered by Ag nanorod substrates allow for multivariate analysis as a method for classification and identification.[226] The statistical basis for the application of chemometric techniques to vibrational spectroscopy is well established.[227] Spectral interpretation was accompanied by unsupervised pattern recognition methods such as principal components analysis (PCA) and hierarchical component analysis (HCA), as well as the supervised method partial least squares discrimination analysis (PLS-DA), as previously described.[226,228-230]

There have been previous reports in which chemometric analysis has been applied to the normal Raman spectra of intact mycobacteria.[231] However, the current work is the first report on classification based on the SERS spectra of mycolic acid components contained in lipophilic extracts from mycobacteria. Figures 5.4 –5.6 demonstrate that the quality and reproducibility of the SERS spectra of mycolic acids make these suitable for subsequent chemometric analysis.

Classification of the Non-Tuberculous Mycobacterial Species

We used PCA, HCA, and PLS-DA methods to analyze these spectra. Principal components were calculated from the SERS spectra of the three NTM species *M. smegmatis*, *M. avium*, and *M. bovis* BCG. The PC model consisted of 36 raw spectra, 12 for each NTM species, and was calculated using the SERS spectra in the 1700 - 700 cm^{-1} range. PC scores for the NTM species were also used to compute a hierarchical cluster analysis. Both the PCA and HCA plots show distinct clusters, and that SERS is able to differentiate NTM species at

the species level for each of the samples analyzed. The PCA and HCA plots are presented in the ESI (figures A1.4 and A1.5).

We also used PLS-DA as an alternative chemometric method to determine statistically significant spectral differences among the SERS spectra in a supervised fashion. Unlike PCA and HCA, PLS-DA is a supervised method in which a reduced set of latent variables (LVs), in combination with an *a priori* knowledge of class membership, determines the best-fit mathematical relationship between a descriptor matrix, *i.e.*, the individual sample SERS spectra, and a class matrix of sample identities.[232,233] We previously used PLS-DA to analyze SERS spectra and showed that it is a robust statistical technique for identification, discrimination, and classification of pathogens and biomolecules.[234,235]

When initial PLS-DA calibration is performed on a calibration data matrix, a threshold value between 0 – 1, calculated using Bayes' Theorem to minimize total error, is calculated for each class.[236] A threshold value of 0 indicates 0% possibility of finding a spectrum in a particular class, while a value of 1 means there is 100% of finding the spectrum in that assigned class; the optimum predictive threshold value for most classes lies between these limiting values. The model then tests each individual spectrum to predict whether they belonged to a specific class. Spectra that have a predicted value greater than the Bayesian threshold are classified as belonging to a specific sample class, while those with predicted values below the threshold are excluded from that class.[237]

Figure 5.7 presents the results of a PLS-DA analysis of the NTM SERS spectra. The horizontal red line in figure.5.7 and subsequent figures represents the PLS-DA calculated threshold value of prediction. Figure 5.7 indicates that the PLS-DA model was able to

correctly classify each NTM spectrum in its class with 100% accuracy. A table of the quantitative statistics calculated from the PLS-DA model of the NTM species is presented in Appendix 1 (Table A1.7).

Classification of the Tuberculous Mycobacteria

SERS spectra of two clinical strains of *M. tuberculosis* (*i.e.* Erdman and H37Rv) were analyzed. The six strains included the wild type, the *sigC* knockout mutant, and the *sigC* complement for both the Erdman and H37Rv strains.

A series of PCA plots, HCA dendrograms, and PLS-DA models were calculated to determine if SERS combined with chemometrics could be used to presumptively identify the mycobacteria present in six strains of *M. tuberculosis*.

Principal components and cluster analysis were calculated for the SERS spectra of the MTB Erdman wild type, EΔsigC knockout, and EΔCcomp complement strains. PCA scores plots and HCA dendrograms derived from PC scores were able to classify the three *M. tuberculosis* Erdman strains. These plots resulted in three clusters, albeit with two of the clusters having significant overlap. The Erdman EΔsigC knockout mutant cluster was distinct from that of the wild type and the EΔCcomp complement. However, the wild type and EΔCcomp strains had significant overlap in both PCA and HCA. The PCA scores plots and HCA dendrograms for the MTB Erdman strains are provided in Appendix 1 (Figures A1.6 and A1.7).

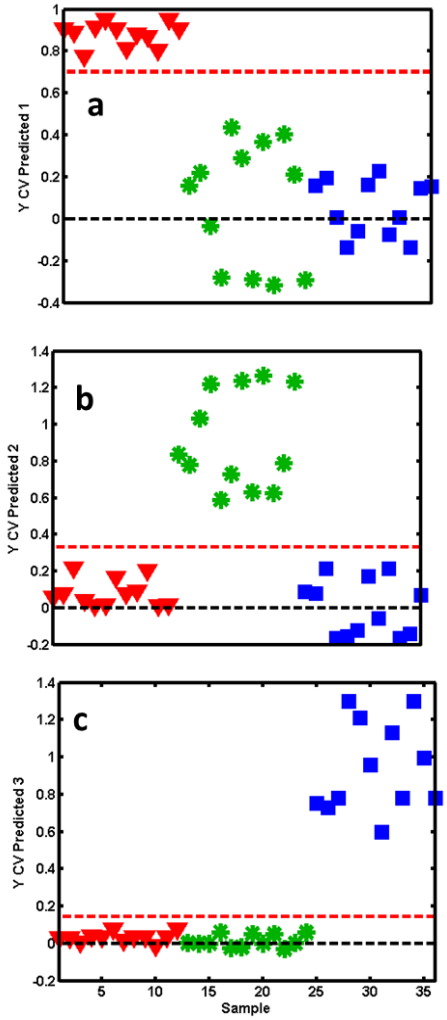

Figure 5.7 PLS-DA cross-validated prediction plot based on the SERS spectra of mycolic acids extracted from the non-tuberculosis species. Horizontal red line denotes calculated class prediction threshold level. (a) Predictions for *M. avium*. (⑤) (b) Predictions for *M. bovis* BCG. (☒) (c) Predictions for *M. smegmatis*. (⓿) Thirty six spectra, corresponding to 12 spectra in each sample category, are represented in this plot.

Figure 5.8 presents the results of a PLS-DA analysis of the *M. tuberculosis* Erdman strain SERS spectra. Unlike the PCA and HCA methods, the PLS-DA model was able to correctly classify not only the MTB Erdman knockout mutant from the wild-type strain, but also the knockout from the complement with 100% accuracy. This is remarkable classification sensitivity, as EΔCcomp consists of EΔsigC complemented with a wild-type copy of *sigC* to restore *sigC* production. Therefore, the distinction between the wild type and EΔCcomp strains likely come from other genetic information encoded on the complementing vector such as the kanamycin resistance or the integrase gene that are not known to impact mycolic acids. Yet a PLS-DA model based on the SERS spectra of the mycolic acids from these strains is able to distinguish them with 100% sensitivity and specificity. The quantitative statistics calculated from the PLS-DA model of the *M. tuberculosis* Erdman strains are presented in Appendix 1 (Table A1.8).

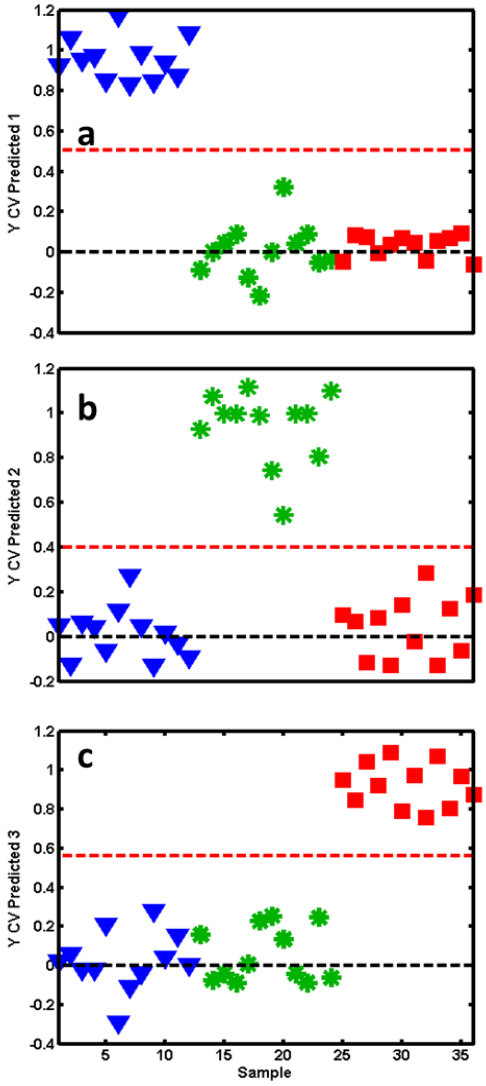

Figure 5.8 PLS-DA cross-validated prediction plot based on the SERS spectra of mycolic acids extracted from the *M. tuberculosis* Erdman strains. Horizontal red line denotes calculated class prediction threshold level. (a) Predictions for the wild type strain. (⑤) (b) Predictions for the EΔCsigC deletion mutant strain. (☒) (c) Predictions for the EΔCcomp complement strain. (⓿) Thirty six spectra, corresponding to 12 spectra in each sample category, are represented in this plot.

We further analyzed three additional *M. tuberculosis* strains related to the lab-passaged clinical strain H37Rv, which included i) the H37Rv wild type strain, ii) RvΔsigC – the H37Rv strain with an in-frame deletion within the *sigC* transcription factor, and iii) RvΔCcomp – the H37Rv deletion mutant strain complemented with *sigC*. Principal components and cluster analysis were calculated for the SERS spectra of the H37Rv wild-type, RvΔsigC knockout, and RvΔCcomp complement strains. Similar results were obtained to those of the MTB Erdman strain described above, *i.e.*, in the PCA scores plots and HCA dendrograms the RvΔsigC knockout mutant cluster was distinct from the clusters of the H37Rv wild type and the RvΔCcomp complement. However, the wild type and RvΔCcomp clusters had significant overlap. The PCA scores plots and HCA dendrograms for the MTB H37Rv strains are provided in Appendix 1 (figures A1.8 and A1.9).

Figure 5.9 presents the results of a PLS-DA analysis of the *M. tuberculosis* H37Rv strain SERS spectra. Similar to the MTB Erdman results, the PLS-DA model was able to quantitatively classify the wild type, knockout mutant, and the complement strains. However, in contrast with the MTB Erdman strain (figure. 5.8), figure.5.9 shows that there is not 100% accuracy in this classification model. This analysis model, which used 3 latent variables and a Venetian blinds cross-validation with 6 splits, resulted in the following classification statistics: i) 100%

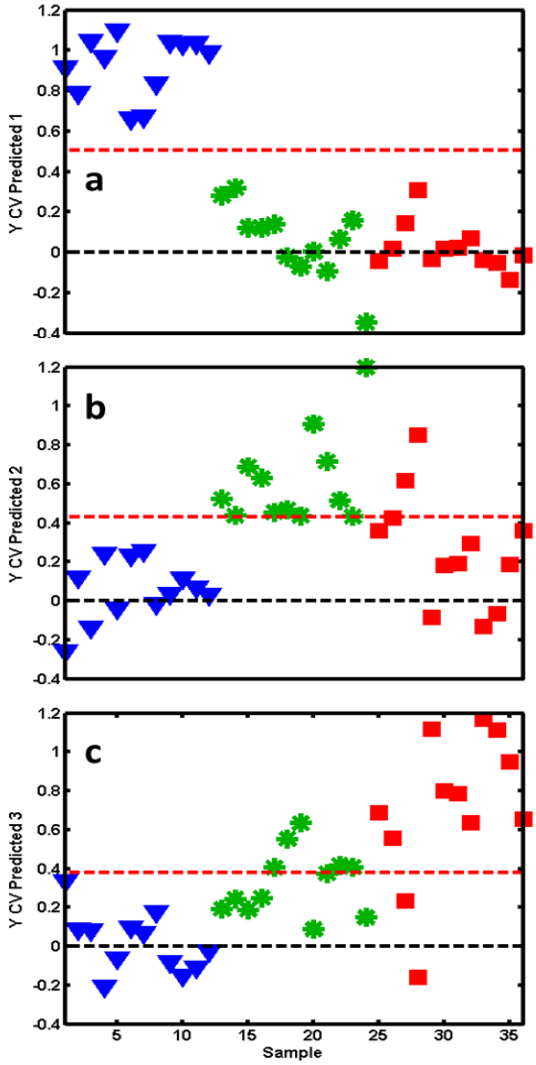

Figure 5.9 PLS-DA cross-validated prediction plot based on the SERS spectra of mycolic acids extracted from the *M. tuberculosis* H37Rv strains. Horizontal red line denotes calculated class prediction threshold level. (a) Predictions for the wild type strain. (⑤) (b) Predictions for the RvΔCsigC deletion mutant strain. (☒) (c) Predictions for the RvΔCcomp complement strain. (⓿) Thirty six spectra, corresponding to 12 spectra in each sample category, are represented in this plot.

sensitivity and specificity for the H37Rv strain, ii) 100% sensitivity and 91.7% specificity for the RvΔComp strain, and iii) 83.3% sensitivity and 79.2% specificity for the RvΔsigC mutant strain. The quantitative statistics calculated from the PLS-DA model of the *M. tuberculosis* H37Rv strains are presented in Appendix (Table A1.9).

Finally, we simultaneously analyzed all six *M. tuberculosis* strains (3 H37Rv and 3 Erdman) using a single PLS-DA model. These calculations used seventy-two spectra and demonstrated the ability of a single chemometric model to discriminate among all the closely related MTB strains used in this study. As shown in figure. 5.10, the classification results from all six MTB strains calculated simultaneously are remarkably similar to those calculated from the H37Rv and Erdman strains individually. In figure.5.10, each panel on the left side corresponds to the PLS-DA predictions made for one of the H37Rv strains, *i.e*. the RvΔsigC mutant (figure.5.10a), the RvΔCcomp complement (figure.5.10b), or the wild type (figure.5.10c). It is clear from these panels that the RvΔsigC mutant and wild type strains are easily distinguished. However, discrimination of the RvΔCcomp is not complete; the classification results of these spectra overlap with those of the wild type strain. This result is interesting, since the H37Rv complement should be biochemically identical to the wild type. These results are also consistent with the data obtained for the three H37Rv strains alone, shown in figure.5.9.

Each panel on the right side of figure.5.10 corresponds to the PLS-DA predictions made for one of the MTB Erdman strains, *i.e*. the EΔsigC mutant (figure.5.10d), the EΔCcomp complement (figure. 5.10e), or the wild type (figure. 5.10f). The level of class discrimination for each Erdman strain is quite high, and while it is not 100%, it is comparable to the results obtained for the three Erdman strains alone (figure.5.8).

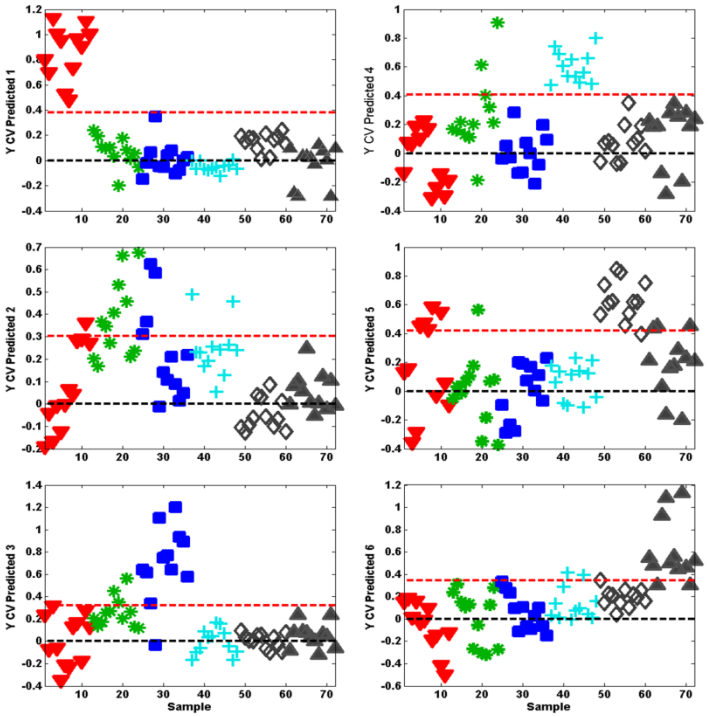

Figure 5.10 PLS-DA cross-validated prediction plots based on the SERS spectra of mycolic acids extracted from both *M. tuberculosis* H37Rv and Erdman strains. All spectra were analyzed simultaneously using the same PLS model. Horizontal red lines in the individual graphs denote calculated prediction threshold level for each individual class.

Left side of the plot, panels (a), (b) and (c), are predictions for the H37Rv strains: (a) the RvΔsigC mutant (⓿) (b) the RvΔCcomp complement (☒), (c) Wild Type (⑤). Right side of the plot, panels (d), (e) and (f) are predictions for the Erdman strains: (d) EΔsigC mutant (▯), (e) EΔCcomp complement (◇), (f) Wild type (▲). Seventy-two spectra, corresponding to 12 spectra in each sample category, are represented in this plot.

The quantitative statistics calculated from the single PLS-DA model of the six MTB strains are presented in Table5.2. For the six-class classification model, sensitivities are between 83 – 100% for all classes, while specificities are between 90 – 98%. The only class with a lower calculated

specificity is that of the H37Rv complement (58%). This lower value is actually due to the biochemical similarity of the RvΔCcomp strain with the wild type strain, such that the statistical regression algorithm had difficulty in separating the classes.

Table5.2 Quantitative statistics calculated from the joint PLS-DA model developed from the SERS spectra[a] of the three *M. tuberculosis* H37Rv and three Erdman strains

Modeled class[b]	Sensitivity[c]	Specificity	Class Error[d]	RMSEC[e]
H37Rv RvΔsigC	1.000	0.983	0.008	0.147
H37Rv RvΔCcomp	1.000	0.583	0.258	0.288
H37Rv wild type	0.917	0.950	0.067	0.203
Erdman EΔsigC	1.000	0.950	0.025	0.245
Erdman EΔCcomp	1.000	0.900	0.050	0.241
Erdman wild type	0.833	0.967	0.100	0.241

[a]Seventy-two total spectra used, 12 for each MTB strain. Before calculation, spectra were pre-processed using Savitzky-Golay 1st derivatives, vector normalization, and mean-centering.

[b]Six latent variables, accounting for 85.54% of the captured variance, were used in this model.

[c]CV, cross-validation based on Venetian blinds method with 6 splits

[d]Class. Error, classification error after cross-validation

[e]RMSECV, root-mean square error after cross-validation

5.5 Conclusions

Mycolic acids are the most characteristic high molecular weight bioactive lipid component of the mycobacterial cell envelope; their structures vary greatly according to mycobacterial species. In this study we were able to show that multivariate statistical analysis of SERS spectra can be used to accurately differentiate several NTM and MTB strains by analyzing mycolic acid cleavage products obtained from lipophilic extracts of bacterial cell walls.

Prior to Raman analysis, we characterized the components of the lipophilic extracts using chromatography and NMR spectroscopy. First, we determined the identity of the mycolic acid fatty acid methyl esters extracted from the mycobacterial strains using GC-MS (Table5.1). The results were found to be identical with previous studies, confirming the presence of the expected lipids in our samples. ^{1}H-NMR spectra were then used to identify the structure of the mycolic acids in the samples, with special attention given to the types of cyclopropane rings present. Only trace amounts of cyclopropane groups were found in the NTM species (figure.5.1). However, the MTB Erdman and H37Rv strains show that ^{1}H resonances attributable to *cis*-cyclopropyl groups predominate (figures.5.2 and 5.3). This evidence further supports previous studies that specific cyclopropane structures generate a homologous family of mycolic acid subclasses indicative of *M. tuberculosis* pathogenesis.

SERS spectra of the NTM (figure.5.4) and MTB (figures.5.5 and 5.6) species showed vibrational bands characteristic of the mycolic acid lipid species. These Raman spectra were used to develop a classification method based on chemometrics for quick identification of the mycobacterial species. PCA, HCA, and PLS-DA (figure.5.7) plots of the SERS spectra enabled differentiation of NTM mycobacteria from one another with

100% accuracy. These same methods were applied to clinically isolated *M. tuberculosis* strains Erdman and H37Rv, the derivatives of each strain defective in production of sigma factor C, and the *sigC* mutants complemented with a wild-type copy of *sigC* to restore production. For both the Erdman and H37Rv strains, PCA and HCA fully separated the mutant strains from the wild type, but did not fully separate the wild type and complement strains. A PLS-DA model (figure.5.8) was able to fully classify Erdman wild type, EΔsigC knockout, and EΔCcomp complement strains from each other with 100% sensitivity and specificity. However, in the case of H37Rv, the PLS-DA model (figure.5.9) resulted in 91.7% specificity for the RvΔComp strain, and 83.3% sensitivity and 79.2% specificity for the RvΔsigC mutant strain. A simultaneous analysis of all six MTB strains (figure.5.10) using a combined PLS-DA model was able to discriminate with 83 – 100% sensitivity and 58 – 98% specificity. The low value for the specificity of the RvΔComp strain (58%) is likely due to its high biochemical similarity to the parent H37Rv wild type strain from which it was derived.

The current work is the first report on discrimination of mycobacteria strains based on the SERS spectra of mycolic acids contained in the lipophilic extracts of mycobacteria. These results demonstrate that SERS, in combination with multivariate statistical methods, can be used as an accurate, quantitative, and sensitive method for species and strain discrimination in mycobacteria.

Acknowledgements

The work described here was supported through a University of Georgia Franklin College Faculty Research Seed Grant.

The following reagent was obtained through BEI Resources, NIAID, NIH: *Mycobacterium tuberculosis*, Strain H37Rv, Purified Mycolic Acid Methyl Esters, NR-14854.

Appendix1 Information. 1) Representative total ion chromatograms of four of the *Mycobacterium* species investigated by GC-MS (Figure A1.1). 2) Mycolic acid chemical structures (Figure A1.2). 3) Unprocessed SERS spectra of MTB, strain Erdman (Figure A1.3). 4) PCA scores plot corresponding to the NTM species (Figure A1.4). 5) A hierarchical cluster analysis dendrogram derived from the PC scores of the NTM species (Figure A1.5). 6) PCA scores plot corresponding to *M. tuberculosis* Erdman strains (Figure A1.6). 7) A hierarchical cluster analysis dendrogram derived from the PC scores of the *M. tuberculosis* Erdman strains (Figure A1.7). 8) PCA scores plot corresponding to *M. tuberculosis* H37Rv strains (Figure A1.8). 9) A hierarchical cluster analysis dendrogram derived from the PC scores of the *M. tuberculosis* H37Rv strains (Figure A1.9). 10) Observed NMR chemical shift resonances appearing in the ^{1}H-NMR spectra of the non-tuberculous mycobacteria *M. smegmatis*, *M. avium* and *M. bovis* BCG (Table A1.1). 11) Observed NMR chemical shift resonances appearing in the ^{1}H-NMR spectra of the *M. tuberculous* clinical Erdman strain (Table A1.2). 12) Observed NMR chemical shift resonances appearing in the ^{1}H-NMR spectra of the *M. tuberculous* laboratory-passaged H37Rv strains (Table A1.3). 13) Representative Raman bands appearing in the SERS spectra of the non-tuberculous mycobacteria *M. smegmatis*, *M. avium* and *M. bovis* BCG

(Table A1.4). 14) Representative Raman bands appearing in the SERS spectra of the *M. tuberculosis* clinical Erdman strains (Table A1.5). 15) Representative Raman bands appearing in the SERS spectra of the *M. tuberculosis* laboratory-passaged H37Rv strains (Table A1.6). 16) Quantitative statistics calculated from the PLS-DA model of the three NTM species (Table A1.7). 17) Quantitative statistics calculated from the PLS-DA model of the three *M. tuberculosis* Erdman strains (Table A1.8). 18) Quantitative statistics calculated from the PLS-DA model of the three *M. tuberculosis* H37Rv strains (Table A1.9).

CHAPTER 6

Layer-by-Layer Polyelectrolyte Encapsulation of *Mycoplasma pneumoniae* for Enhanced Raman Detection

Rivera-Betancourt, O.E., Sheppard, E.S., Krause, D.C., Dluhy, R.A.

2014**,** *Analyst*, 139, 4287-4295.

6.1 Abstract

Mycoplasma pneumoniae is a major cause of respiratory disease in humans and accounts for as much as 20% of all community-acquired pneumonia. Existing mycoplasma diagnosis is primarily limited by the poor success rate at culturing the bacteria from clinical samples. There is a critical need to develop a new platform for mycoplasma detection that has high sensitivity, specificity, and expediency. Here we report the layer-by-layer (LBL) encapsulation of *M. pneumoniae* cells with Ag nanoparticles in a matrix of the polyelectrolytes poly(allylamine hydrochloride) (PAH) and poly(styrene sulfonate) (PSS). We evaluated nanoparticle encapsulated mycoplasma cells as a platform for the differentiation of *M. pneumoniae* strains using surface enhanced Raman scattering (SERS) combined with multivariate statistical analysis. Three separate *M. pneumoniae* strains (M129, FH and II-3) were studied. Scanning electron microscopy and fluorescence imaging showed that the Ag nanoparticles were incorporated between the oppositely charged polyelectrolyte layers. SERS spectra showed that LBL encapsulation provides excellent spectral reproducibility. Multivariate statistical analysis of the Raman spectra differentiated the three *M. pneumoniae* strains with 97 – 100% specificity and sensitivity, and low (0.1 – 0.4) root mean square error. These results indicated that nanoparticle and polyelectrolyte encapsulation of *M. pneumoniae* is a potentially powerful platform for rapid and sensitive SERS-based bacterial identification.

6.2 Introduction

Mycoplasma pneumoniae is a significant human respiratory pathogen, causing bronchitis and atypical or "walking" pneumonia. *M. pneumoniae* accounts for 20% of all community-acquired pneumonia and is the leading cause of pneumonia in older children and young adults.[34] [238] Serologic testing is a common method for diagnosis due to significant challenges posed by direct culture, but suffers from severe limitations, including the need for paired sera obtained at separate physician visits, and thus is impractical for rapid testing.[34] Detection of *M. pneumoniae* by polymerase chain reaction (PCR) yields high specificity, but is prone to false-negatives.[239] The inability to provide rapid and definitive diagnosis delays initiation of appropriate treatment, prolongs morbidity, and increases the likelihood of continued transmission, secondary infections, and long-term sequelae, including chronic lung disease associated with COPD and asthma.[239] Lack of a simple, reliable, rapid diagnostic test is thus a critical barrier to the improved control of *M. pneumoniae*.

Our laboratories have used a combination of surface-enhanced Raman scattering (SERS)-based nanotechnology methods with pattern-recognition approaches to yield direct, rapid, and sensitive detection of infectious agents.[90] Nanofabrication by oblique angle vapor deposition produces Ag nanorods arrays exhibiting extremely high electromagnetic field enhancements.[90,95,240] Paired with chemometric analysis, this platform can rapidly detect and distinguish with great sensitivity and specificity the Raman spectra of viruses and bacteria, including mycoplasmas.[151,207,234,241]

This current study reports on a new SERS platform for mycoplasma detection that is based on modification of Ag nanoparticles (AgNP) to increase their affinity for the bacteria. The

direct placement of Ag nanoparticles onto living cells can affect the viability of cells either during the process of deposition of nanoparticles or shortly after. Therefore, we have adapted the use of layer-by-layer (LBL) encapsulation techniques, which are widely used for modification of substrates such as planar surfaces and nanoparticles.[242-244]

The LBL technique utilizes the consecutive deposition of oppositely charged polyelectrolytes onto surfaces,[245] allowing consecutive layers to be formed.[246-248] A general outline of the LBL assembly procedure begins with a polycation such as poly(allylamine hydrochloride) (PAH), followed with a polyanion such as poly(styrene sulfonate) (PSS). The LBL process is repeated until the planned shell architecture is realized. Biological cells are suitable templates for LBL coatings because they can be used as cores for the development of polyelectrolyte microcapsules while their biological activity is preserved. LBL deposition of polyelectrolyte assemblies affords nanoscale control over the construction of multilayers with charged nanoparticles.[249,250]

Many microbial and human cells are negatively charged.[251] Polyelectrolyte assemblies can thus facilitate adhesion of nanoparticles to cells and provide stability to the sandwich-like polyelectrolyte/nanoparticle coating. The use of electrostatic LBL encapsulation of bacterial cells with SERS-active nanoparticles has previously been explored by several research groups,[161,165,251-253] and has recently been reviewed.[251] Layer-by-layer polyelectrolyte assembly with colloidal Au and Ag nanoparticles has been demonstrated with fungi[252] and bacteria.[161] In this study we used LBL techniques to encapsulate three different strains of *M. pneumoniae* with Ag nanoparticles (AgNPs) for SERS analysis. Our hypothesis is that the charged polyelectrolyte layers should increase the number of contact points between the AgNPs nanoparticles and the bacterial cell for improved SERS spectral

quality, thereby increasing accuracy in identification and differentiation of different mycoplasma strains.

Our laboratories have previously used planar Ag nanorod array substrates to detect and differentiate *M. pneumoniae* strains with statistically significant sensitivity and specificity.[151] The current work uses LBL encapsulation as an alternative SERS preparation method to avoid issues with pleomorphism and lysis due to the absence of a cell wall in mycoplasmas. LBL-SERS methods have not previously been reported for detection and identification of mycoplasmas. We used *M. pneumoniae* wild-type strain M129 as a model organism to illustrate the LBL encapsulation procedure. The results presented in this study showed that the LBL method identified three *M. pneumoniae* strains with 97 – 100% specificity and sensitivity, and with extremely low root-mean-square errors.

6.3 Materials and Methods

Chemicals. Poly(allylamine hydrochloride) (PAH, M_w ~15,000), sodium (polystyrene sulfonate) (PSS, M_w ~70,000), fluorescein-isothiocyanate-PAH (FITC-PAH, M_w ~15 kDa), and 4′,6-diamidino-2-phenylindole dihydrochloride (DAPI) were purchased from Sigma-Aldrich (St. Louis, MO). PELCO® NanoXact™ citrate-capped Ag colloid nanoparticles (50 nm) were purchased from Ted Pella, Inc., (Redding, CA). Non-functionalized (SiOH) silica microspheres (600 nm) were purchased from Bangs Laboratories, Inc., (Fishers, IN).

Culture and Preparation of Bacterial Strains. Two major wild-type *M. pneumoniae* subtypes, M129 and FH,[254] as well as strain II-3, a spontaneously arising avirulent mutant derived from M129,[255,256] were used in this study. Mycoplasmas were grown to log phase with a 1 µl/ml inoculation. The wild-type and mutant strains were grown in 25 ml of SP4 medium[257,258] in cell culture flasks at 37˚C for 72-96 h and harvested when the phenol red pH indicator turned orange (pH approx. 6.5). The growth medium for the M129 and FH strains was poured off and the cells were scraped from the flask surface into 2.5 ml of fresh SP4 medium. For the II-3 strain, which fails to attach to plastic, cell suspensions were collected through centrifugation at 25,000×g for 25 min at 4˚C and then suspended in 2.5 ml of fresh SP4. Mycoplasma suspensions were syringe-passaged 10 times with a 25-gauge needle to disperse the cells, and aliquots of each were serially diluted for plating to measure colony-forming units (CFU). A 500 µl aliquot of each strain was transferred to a separate tube and fixed in SP4 by adding 500 µl of 8% formaldehyde (pH 7.0-7.5) for a final 4% formaldehyde concentration and stored at 4˚C until used for cell encapsulation.

Polyelectrolyte Encapsulation. A three-step wet chemical assembly process was used for encapsulation of the mycoplasma cells.

Step 1. Mycoplasma Phase. The first step involved encapsulating the bacterial cells in a layer-by-layer fashion by alternating depositions of PAH/PSS/PAH. Polyelectrolyte solutions were dissolved in 0.5M NaCl at the concentration of 1 mg/ml. The procedure began by coating with PAH; 500 µl of the cell suspension was combined with 250 µl of 1 mg/ml PAH and 250 µl of 1 mg/ml PSS and mixed for 15 min at room temperature. This mixture was centrifuged for 10 min at 17,000 rpm at 4˚C, excess polyelectrolyte solution was discarded, and the cells suspended and washed two additional times with cold ultrapure water. The cell suspension was centrifuged for 10 min at 17,000 rpm at 4˚C after each wash. To the same tube were added 250 µl of 1 mg/ml solution of PAH, and cold water to a final volume of 1 ml. The suspension was mixed for 15 min, and then centrifuged for 10 min at 17,000 rpm at 4˚C. Excess polyelectrolyte solution was discarded and the cells suspended and washed two times with cold ultrapure water. The supernatant was discarded at this point leaving the pelleted cells.

Step 2. AgNP Phase. The second step involved coating the AgNP colloidal suspension with PAH and PSS. 1 ml of the Ag colloid suspension was centrifuged at 7000 rpm for 10 min at 4˚C. The supernatant was discarded and an additional 900 µl of colloidal suspension was added. To this AgNP suspension was added 50 µl of 1 mg/ml PAH and the contents mixed for 15 min. Then 50 µl of 1 mg/ml PSS were added and the contents again mixed for 15 min. The suspension was then centrifuged for 10 min at 17,000 rpm at 4˚C. Excess

polyelectrolyte solution was discarded and the AgNP's suspended and washed two times with cold ultrapure water.

Step 3. Encapsulation of Cells. The encapsulated mycoplasma pellet from Step 1 was mixed with the polyelectrolyte-coated AgNPs from Step 2 for 15 min and then centrifuged for 10 min at 17,000 rpm at 4˚C. The supernatant was discarded and the pellet suspended and washed twice with cold ultrapure water, centrifuging for 10 min at 17,000 rpm after each wash. At the end of the process, the cells were suspended in water.

Characterization of the Encapsulated Cells. The uniformity of the LBL polyelectrolyte coating was investigated using FITC-PAH and DAPI nucleic acid stains. A Nikon A1R confocal microscope with a CFI Plan APO VC 60× oil immersion objective with NA=1.4 and 0.13 mm working distance was used to image the *M. pneumoniae* cells. Scanning electron microscopy (SEM) images of the uncoated bacteria were obtained using a Zeiss (Jena, Germany) 1450EP SEM. For the encapsulated cells, images were obtained using an FEI (Hillsboro, OR) Inspect F FEG-SEM. Samples for SEM were fixed as described elsewhere,[35] with modifications. Samples of cells were prepared by dispersing 100 µL of a cellular suspension on the surface of a glass coverslip pre-coated with poly-L-lysine and incubated overnight at 37˚C. The samples were fixed in 2% glutaraldehyde in Na cacodylate buffer for one hour and then washed twice in Na cacodylate buffer for 5 min each. The samples were post-fixed in 1% OsO_4 in Na cacodylate buffer for one hour, washed once afterwards with Na cacodylate buffer for 10 minutes, and then rinsed with water twice for 5 min. The SEM coverslips were treated with a sequential ethanol

dehydration series (5 min each step) with 25, 50, 75, 85, 95, and 3× 100% washes, critical point dried, and sputter coated with Au for examination.

SERS Measurements of the M. pneumoniae Strains. SERS spectra were acquired using a Renishaw (Hoffman Estates, IL) inVia confocal Raman microscope system using a 785 nm near-IR diode laser as the excitation source. Radiation from the diode laser was attenuated to <15 mW using a series of neutral density filters and focused onto the sample using a 20× microscope objective. Spectra were collected between 1800 – 400 cm^{-1} and integrated for 30s per scan with 1 scan per spectrum. The SERS spectra of the encapsulated cells with polyelectrolytes and AgNP were collected applying a 10 µl sample droplet to a copper foil substrate that was cleaned thoroughly with copious amounts of methanol and acetone. The drop was dried in an incubator at 75˚C and then rinsed thoroughly with ultrapure water and dried under a stream of N_2 prior to analysis. A minimum of ten spectra were collected for each bacterial strain from different locations on each individual substrate. Duplicate samples of the LBL-AgNP assemblies for each *M. pneumoniae* strain were prepared to test for reproducibility of the method.

Multivariate Statistical Analysis. Raman spectra were imported into GRAMS AI (Version 8.0 Thermo Electron Corp, Waltham, MA) for spectral averaging and baseline correction. Chemometric analysis was carried out with MATLAB version 7.2 (The Mathworks, Inc., Natick, MA), using PLS Toolbox version 7.0 (Eigenvector Research Inc., Wenatchee, WA). SERS spectra in the range 1650 – 700 cm^{-1} were used for classification. Prior to analysis, first derivatives of the SERS spectra were calculated using the Savitzky-Golay method with a 2^{nd} order polynomial and a fifteen-point window. Each data set was then

vector normalized and mean centered. Multivariate statistical analysis of the mycoplasma spectra was performed using principal components analysis (PCA), hierarchical cluster analysis (HCA), and partial least squares discriminant analysis (PLS-DA) using the PLS Toolbox software. The calculated principal components were used as inputs to the HCA algorithm, which used the Ward's method algorithm to evaluate minimum variances between clusters.

6.4 Results and Discussion

Characterization of LBL Encapsulated Mycoplasma. We used *M. pneumoniae* wild-type strain M129 as a model organism to illustrate the LBL encapsulation procedure. Synthetic polycation/polyanion pairs, *i.e*. PAH/PSS, were used to produce layered shells that covered the bacteria. PAH was deposited as the first layer to balance the negative surface charge of the mycoplasma cells, followed by the polyanion PSS, and then a final layer of PAH was added. Therefore, the final mycoplasma LBL structure was PAH/PPS/PAH.

PAH was also used as the first polyelectrolyte layer on the citrate-reduced AgNPs, followed by a layer of PSS. This final polyanion layer provides higher stability to the structure, as well as sensitivity to temperature and permeability.[243,248,251] Figure 6.1 shows a schematic representation of the deposition process. The bacteria were treated with PAH/PPS/PAH rather than just one layer of PAH since the three-layer system dramatically increased the quality of the resulting SERS spectra of the encapsulated cells. Additional layers of polyelectrolytes also increases the electrostatic interaction of the bacteria with the Ag nanoparticles, as well as the number of contact points between the nanoparticles and bacterial surface.

The cells were washed after every LBL deposition cycle to remove any excess polyelectrolytes. Deposition of the polyelectrolyte-coated AgNPs on the mycoplasma cells could be monitored visually. Originally, the cell suspensions were clear; however, the suspended cells acquired a brownish-yellow color during the deposition of the AgNPs. The color of the suspension was due to the presence of the bound AgNPs in the LBL matrix and not to free, unbound nanoparticles, as these were removed during the washing steps. The final washing steps resulted in a clear supernatant, indicating no AgNPs were being released from the cell suspensions.

Cell morphology of wild-type *M. pneumoniae* M129 wild type and mutant II-3 was directly characterized by SEM, as previously described.[35] Mycoplasmas have a marked tendency toward pleomorphism due to the absence of a cell wall.[259,260] Therefore, the use of an electron source for high-resolution imaging can be a challenge. We employed two different SEM's for this work. The first was an environmental SEM (ESEM). This technique is commonly employed for the detection of different living organisms on a "wet" and/or uncoated state.[261] ESEM micrographs of the uncoated, wild-type M129 cells are shown in figures 6.2A and 6.2B. The cells in these images appeared elongated with a well-defined, tapered tip structure and long, filamentous tail; their size was approximately 1-2 μm in length and 0.1-0.2 μm in width. While the predominant morphological forms present in the ESEM micrographs observed were elongated, minor amounts of ovoid and pleomorphic forms were also seen (data not shown).

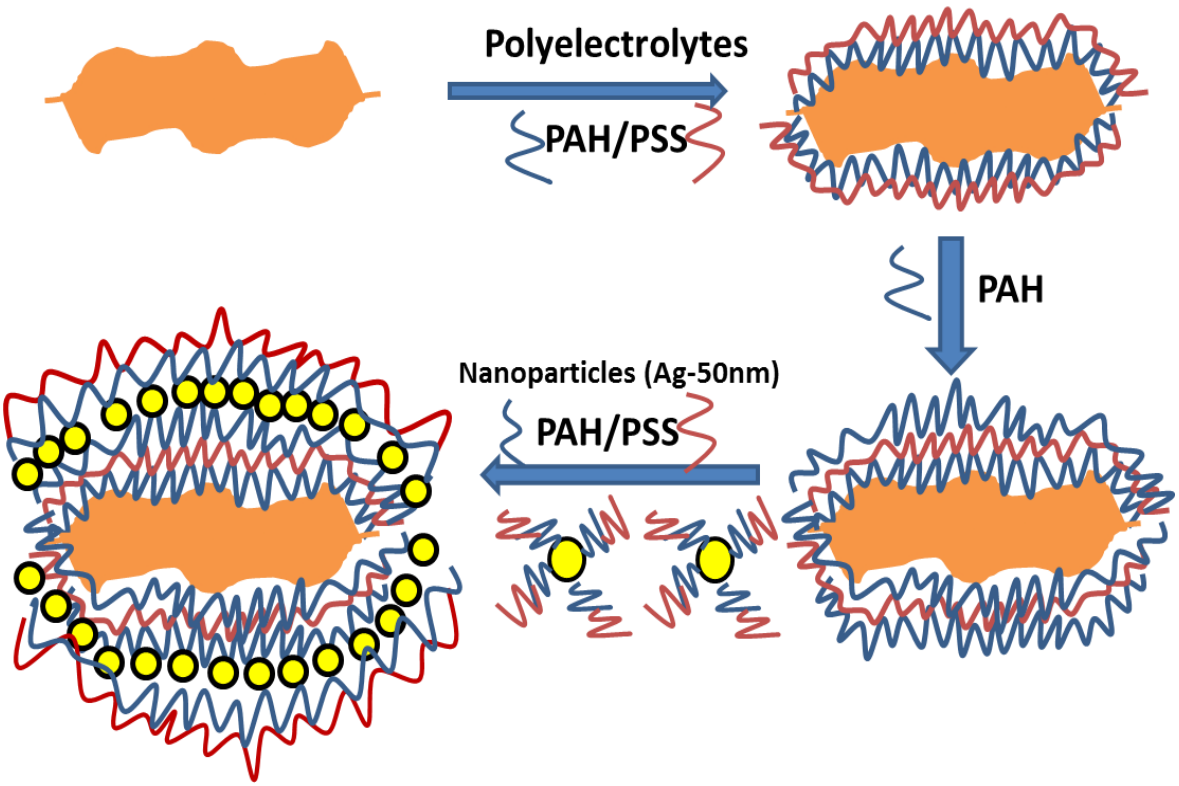

Figure 6.1 Schematic illustration of the encapsulation of the *M. pneumoniae* whole cells into the polyelectrolyte shells containing silver nanoparticles.

The polyelectrolyte-encapsulated cells were characterized by field emission SEM (FE-SEM), as seen in figures.6.2C and 6.2D. In comparison to the smooth surfaces of the uncoated cells in figure.6.2A and figure.6.2B, the encapsulated mycoplasmas in figures.6.2C and 6.2D showed large aggregates and roughened features. The Ag nanoparticles that we used in this work were ~50 nm in diameter; however, the extent of aggregation seen in figure.6.2 makes it difficult to determine the size and location of the Ag nanoparticle complexes within the LBL-mycoplasma structures. Aggregation has previously been reported in LBL polyelectrolyte encapsulation of other bacterial species, with the extent of aggregation depending on the fixation protocols as well as the bacterial cell surface biochemistry.[161]

The presence of the polyelectrolyte-encapsulated AgNPs enhances the aggregation of the mycoplasma cells into clusters. The chemical natures of the polyelectrolytes, as well as the solution ionic strength, have a strong influence on the polyion complexes that are formed. [248,262-264] In the current case, we employed a weak polycation (PAH) at relatively low ionic strength (0.5M NaCl) that led to a heterogenous surface topography previously noted as characteristic for PAH/PSS polyelectrolyte systems.[112,263]

We also conducted fluorescence labeling experiments to confirm the co-localization of the AgNPs on the polyelectrolyte-encapsulated mycoplasmas. These images are shown in figure. 6.3. Two different stains were used for this procedure. First, a commonly used nucleic acid dye, DAPI, was used to stain the nucleoid of mycoplasma cells.[265,266]

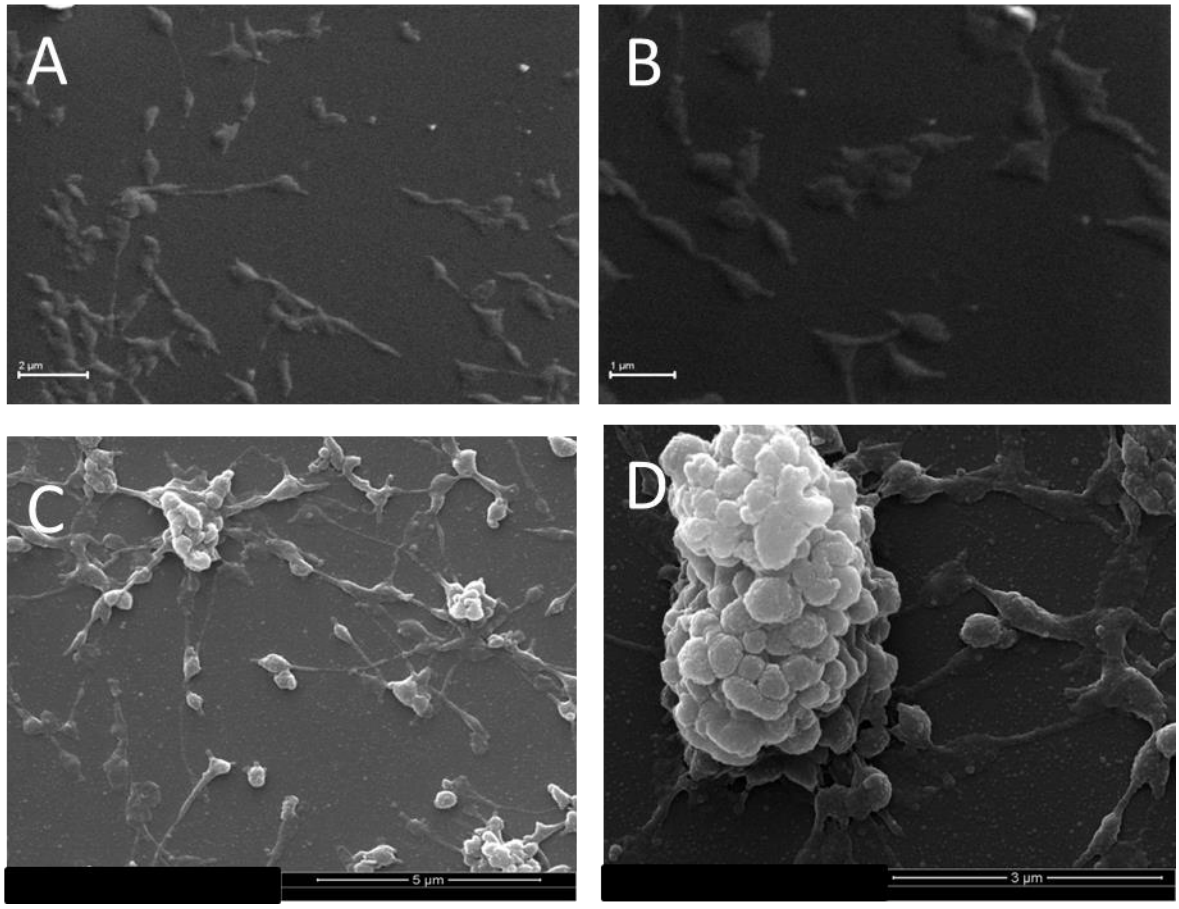

Figure 6.2 SEM images of: *M. pneumoniae* (A) M129 uncoated whole cells, scale bar equals 2 µm; (B) M129 uncoated whole cells at higher magnification, scale bar equals 1 µm; (C) M129/PAH/PSS/PAH/Ag/PAH/PSS cell aggregates, scale bar equals 5 µm; and (D) M129/PAH/PSS/PAH/Ag/PAH/PSS cell aggregates at higher magnification, scale bar equals 3 µm.

DAPI is a blue fluorescent nucleic acid stain that preferentially stains A-T complexes in double-stranded DNA (dsDNA).[265,266] This is illustrated in figure. 6.3A, in which the blue images mark the presence of likely individual un-encapsulated *M. pneumoniae* M129 cells stained with DAPI. Figure 6.3B shows the DAPI image of the mycoplasma after encapsulation with the PAH/PSS polyelectrolytes and AgNPs, and clearly shows the encapsulated bacteria more aggregated than the uncoated bacteria in figure. 6.3A.

In conjunction with the DAPI stain, FITC labeled PAH was used to determine whether the PAH/PSS/PAH polyelectrolyte layers and the PAH/PSS coated AgNPs were bound to the encapsulated mycoplasmas.[267] Figure 6.3C shows the green fluorescence emission of FITC-PAH incorporated into the PAH/PSS/PAH layers on *M. pneumoniae* M129 for the same field as figure. 6.3B. A comparison of figure. 6.3B (DAPI-stained mycoplasma emission) with figure.6.3C (FITC-PAH emission) showed significant fluorescence overlap between the two images, consistent with co-localization of the polyelectrolytes with the bacteria. Figure 6.3D is a merged image of DAPI-labeled *M. pneumoniae* (blue) after deposition of the FITC-PAH layers (green). This image was taken at a different location than those of figure.6.3B and figure.6.3C. Figure 6.3D shows that not all of the FITC signal is associated with the DAPI signal. This suggests that some of the polyelectrolyte may be interacting with sample components other than whole mycoplasma bacterial cells. The additional bacterial components in figure 6.3 may be due to cellular debris, as mycoplasma does not contain a cell wall, and cell lysis potentially occurs.[260]

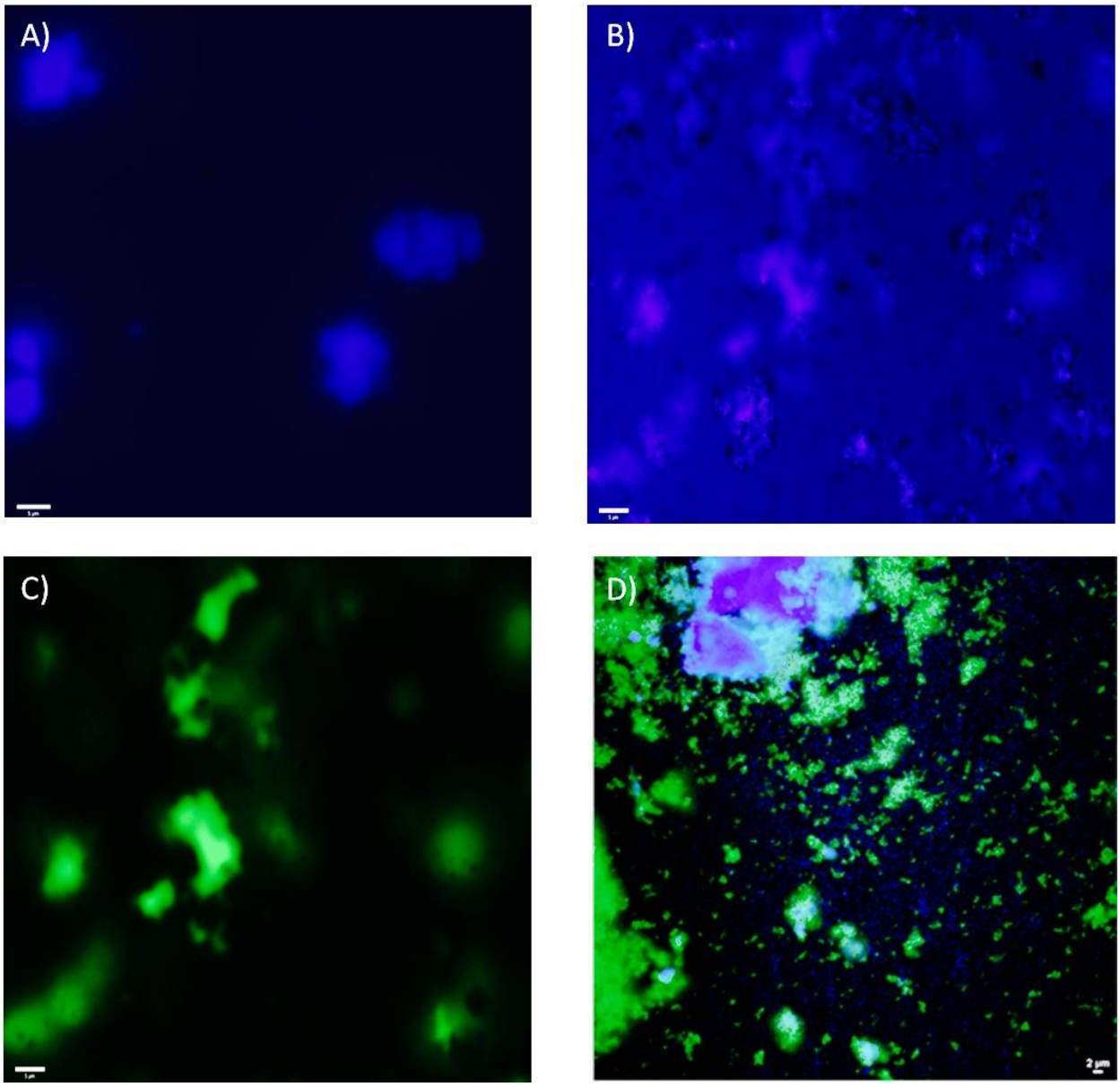

Figure 6.3 Fluorescence microscopy image of: (A) *M. pneumoniae* M129 cells, scale bar equals 1 μm; (B) LBL array of DAPI-M129, scale bar equals 1 μm; (C) LBL array of FITC-PAH on M129 cells, scale bar equals 1 μm; and (D) overlaid image of FITC-PAH coated and DAPI-stained *M. pneumoniae* M129 cells, scale bar equals 6 μm.

SERS Spectra of the M. pneumoniae Strains. The Raman spectra of bacteria reflect predominantly phenotypic information arising from proteins, nucleic acids, lipids, carbohydrates and endogenous biomolecules.[268,269] In order to detect and identify pathogens of interest, it is necessary to ensure that the observed Raman bands are specific to the targeted organisms but not to the overall environment, i.e., the media or solvents. The mycoplasma SP4 growth medium is removed from the sample prior to spectral analysis; therefore, it was not included in the background analysis. Instead, the background control used in these samples was from the last part of the LBL assembly process that includes both polyelectrolytes as well as the AgNPs, *i.e.*, the silver nanoparticle-polyelectrolyte layer Ag/PAH/PSS.

We also incorporated a second negative control sample in these experiments. This negative control utilized non-functionalized silica microspheres (SiMS) of 600 nm diameter. The size of these microspheres closely resembles the actual size of a mycoplasma bacterium cell, ~1000 nm. These experiments used the silica microspheres as a sacrificial or electroactive core to simulate the bacterium as a negative control. We performed all the experimental protocols using the SiMS in place of the bacteria. Therefore, these Si microsphere were coated with PAH and PSS and functionalized with Ag nanospheres, resulting in a LBL structure of SiMS/PAH/PSS/PAH/Ag/PAH/PSS. For brevity, this will be referred to as the SiMS negative control in the remainder of the article.

Figure 6.4 illustrates the reproducibility of the SERS spectra obtained from mycoplasmas prepared using this LBL encapsulation process. Samples containing M129/PAH/PSS/PAH/Ag/PAH/PSS were prepared as described above, and then spot dried on a piece of copper foil. Spectra of the dried samples were obtained from 10 different

locations. The SERS experiments were repeated several times with a newly prepared sample each time. The same approach was used when acquiring the SERS spectra of the other two *M. pneumoniae* strains, as well as the Ag/PAH/PSS background and the SiMS negative controls. The spectra in figure.6.4 were the raw, unprocessed, and baseline un-corrected spectra of the M129 mycoplasma strain. The overlaid spectra demonstrate the reproducibility of the LBL encapsulation process.

SERS is highly dependent on, and sensitive to, the proximity of the AgNPs to the cell surface. Functional groups such as COO^- and NH_2^+ may define the modes of interaction between the bacterial surface and the AgNPs.[159,253,270,271] Figure 6.5 illustrates the average SERS spectra in the 1800 – 400 cm^{-1} range of the three *M. pneumoniae* strains, the Ag/PAH/PSS background and the SiMS negative control . For each sample, including the controls, ten spectra were taken at each of two different spots, baseline-corrected, and normalized for visualization. The Raman vibrations at 1504, 1459, 1277, 1130 and 1080 cm^{-1} were found in all spectra, although their intensities varied.[272] The spectra of the different strains did exhibit some differences; for example, the bands at 1580 and 1299 cm^{-1} were characteristic for *M. pneumoniae* M129, while the bands at 1383, 1359 and 668 cm^{-1} were characteristic for *M. pneumoniae* FH and II-3. The most significant spectral differences occurred in the 1000 – 400 cm^{-1} region. To improve the resolution of overlapping bands and eliminate potential artifacts induced by baseline correction, we also compared first derivative spectra of these samples, as seen in figure. 6.6.

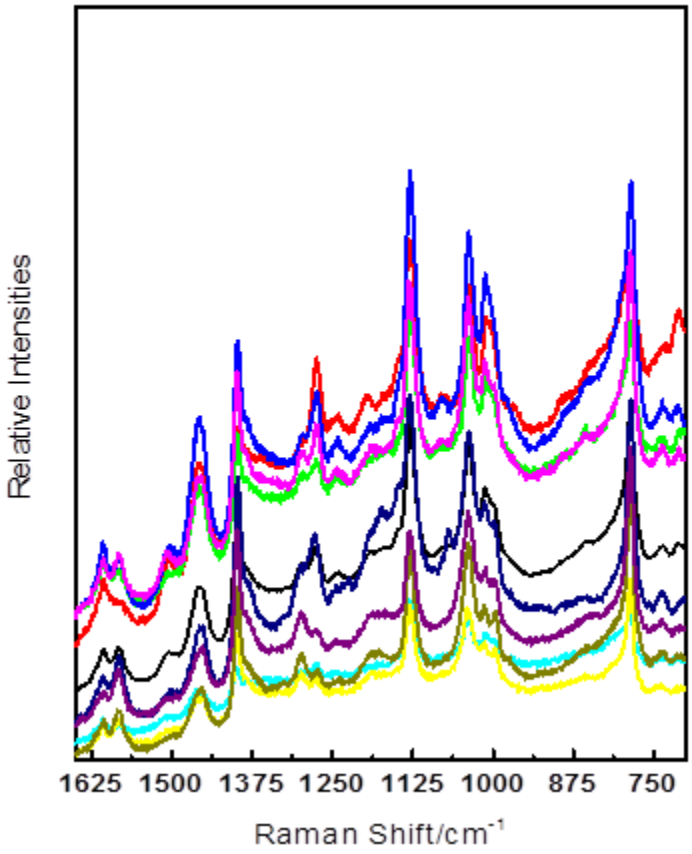

Figure 6.4 Representative SERS spectra of *M. pneumoniae* M129 showing the reproducibility of the spectra collected in ten random locations. Spectra presented here are the original spectra as collected, without further processing.

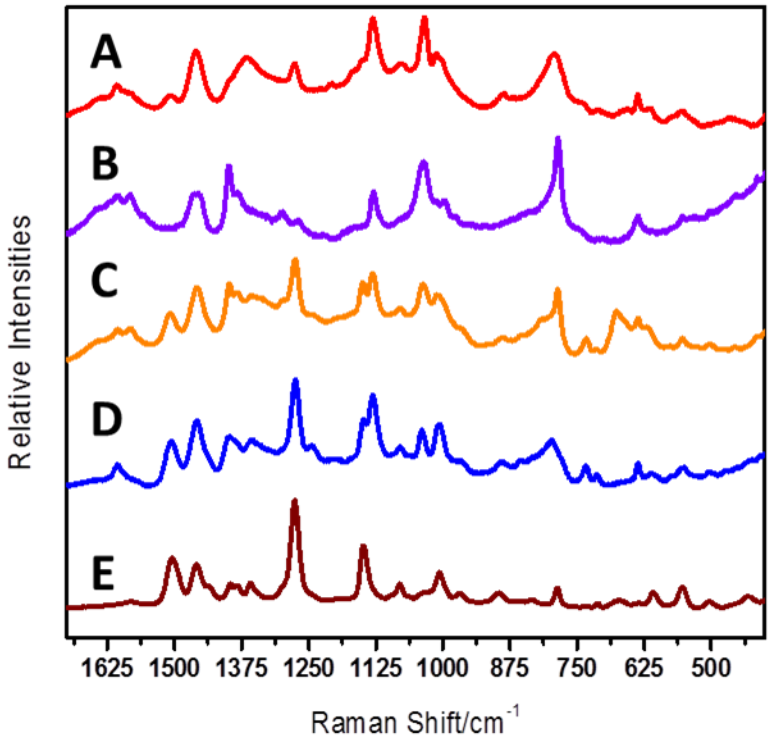

Figure 6.5 Representative SERS spectra for (A) the background sample Ag/PAH/PSS, (B) the negative control sample SiMS/PAH/PSS/PAH/Ag/PAH/PSS and the *M. pneumoniae* strains: (C) M129, (D) FH and, (E) II-3. Each is an average of 10 spectra obtained per sample. The spectra have been baseline corrected and normalized for visualization.

A list of the observed Raman vibrations attributed to the mycoplasma strains in figures.6.5 and 6.6 as well as their tentative assignments are found in Table 6.1. Detailed spectral band assignments have been published elsewhere.[89,159,161,253,270,272,273]

Classification of the M. pneumoniae Strains. The excellent spot-to-spot and sample-to-sample reproducibility of the SERS Raman spectra offered by the LBL methods allows for multivariate analysis as a method for classification and identification. The statistical basis for the application of chemometric techniques to vibrational spectroscopy is well established.[274] Spectral interpretation can be accompanied by unsupervised pattern recognition methods such as principal component analysis (PCA)[271] and hierarchical cluster analysis (HCA), as well as supervised methods such as partial least squares discrimination analysis (PLS-DA).[90,269,275]

We have previously used chemometric methods to analyze nanorod-array SERS spectra of human and avian mycoplasma species extracted into a water-formalin mixture for inactivation of the bacteria, a process that potentially lyses the cells.[150,151] The current work is the first to use the LBL method to ensure preparation of whole, intact encapsulated mycoplasma cells for use in SERS classification studies. The high quality and reproducibility of the LBL SERS spectra, as seen in figures. 6.4 – 6.6, demonstrate that these spectra are suitable for use in subsequent multivariate statistical processing steps.

The statistical methods PCA, HCA, and PLS-DA were used to determine whether it is possible to discriminate between these three different mycoplasma strains and controls

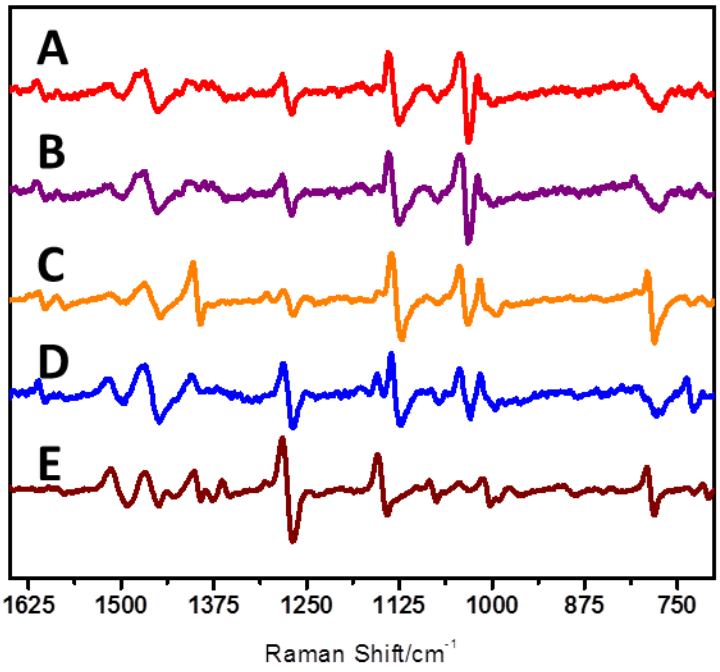

Figure 6.6 First derivative spectra for (A) the background sample Ag/PAH/PSS, B) the negative control sample SiMS/PAH/PSS/PAH/Ag/PAH/PSS and the *M. pneumoniae* strains: (C) M129, (D) FH and, (E) II-3.

Table 6.1. Raman bands appearing in the LBL SERS spectra of *M. pneumoniae* strains.

Raman Shift, cm^{-1}			Vibrational Band Assignment
M129	FH	II-3	
1606	1606		Phenylalanine
1580			Guanine, Adenine (ring stretch)
1504	1504	1504	C-O-H bend; $(CH_2)_n$ in-phase twist
1459	1459	1459	$\delta(C\text{-}H_2)$ sci.; CH_3 antisym. bend
		1435	$\delta(C\text{-}H_2)$ sci
1396	1396	1396	C-O-H bend; (CH_2 in-phase twist
	1383	1383	C-H def
	1359	1359	C-H def
1299			Amide III
1277	1277	1277	C-O-H bend, Amide III
1247	1247		v_{as}(COC); δ(CH),Amide III
1204	1200		$C\text{-}C_6H_5$ str., Phe,Trp
1193		1193	δ (C-H), Tyr
		1149	NH_3^+def Pro
1130	1130		C-N and C-C stretch
1080	1080	1080	C-O stretch,
1039	1039		C-H in plane, Phe, C-N Gly
1008	1008	1008	phenylalanine
		969	C-C str
		898	COC str
858	858	859	"buried" tyrosine
		831	"exposed" Tyrosine
	792		Cytosine, uracil (str, ring); CH_2 in-phase rock
785		785	Cytosine, Uracil
738	738	738	Adenosine
712	712	712	Adenine, COO^- def
	668	668	Guanine
	618		Phenylalanine (skeletal)

based solely on their LBL SERS spectra. PCA reduces dataset dimensionality by calculating orthogonal eigenvector projections, and facilitates establishing patterns and grouping of similar spectra.[240,270,276] Principal components were calculated from the SERS spectra of the two *M. pneumoniae* M129 and FH wild-type strain, the mutant II-3 strain, a control sample consisting of Ag/PAH/PSS background and the SiMS negative control sample. The PC model consisted of 50 raw spectra, 10 spectra for each *M. pneumoniae* strain and the controls, and was calculated using LBL SERS spectra in the 1650 – 700 cm^{-1} range. As shown in figure.6.7, comparison of the processed spectra using principal components 1 and 2 couldn't clearly differentiate between the three mycoplasma strains and the control Ag/PAH/PSS sample into individual groups. But this PC plot easily separates the bacterial samples and background from the SiMS negative control.

In addition, a hierarchical cluster analysis of the LBL SERS spectra was calculated using their principal components. The resulting dendrogram, calculated using the Ward's linkage method (figure.6.8), shows that the three mycoplasma strains, the Ag nanoparticle background, and the SiMS negative control form clearly differentiated clusters.

In addition to the unsupervised PCA and HCA methods, we used PLS-DA to quantitatively determine statistically significant differences among the strains. PLS-DA, unlike HCA and PCA, is a full spectrum, multivariate, supervised method whereby prior knowledge of the classes is used to yield more robust differentiation, minimizing class variation while emphasizing latent variables between or among classes.[186,277]

Figure 6.9 presents the results of a PLS-DA analysis of the mycoplasma strains. The model was generated using 50 spectra (10 each for the mycoplasma strains, Ag/PAH/PSS background, and SiMS negative control). The horizontal red line in each panel in figure.6.9

is a calculation of a threshold value of prediction for each modeled class. Spectra with predicted values above the threshold level are determined to belong to a particular class, while spectra with the predicted values below are excluded. Each panel corresponds to the predictions made for the Ag/PAH/PSS background (figure.6.9A), the SiMS negative control (figure.6.9B), M129 wild-type (figure.6.9C), FH wild-type (figure.6.9D), and II-3 mutant (figure.6.9E). This analysis used 4 latent variables that captured 85.09% of the total variance, followed by cross-validation using Venetian blinds with 7 splits. It is clear from all the panels that PLS-DA was able to classify each spectrum in its class with 100% accuracy. Table 2 provides the statistics calculated from the PLS-DA model, with a root-mean square error after cross-validation (RMSECV) of 0.05 – 0.1 for all classes. This is remarkable classification sensitivity, considering that both M129 and FH are pathogenic wild type strains.

The ability of LBL SERS to accurately differentiate *M. pneumoniae* strains M129 and FH into separate classes may be due to the spectral differences in expressed surface proteins.[35,151] The high discriminatory ability of LBL SERS seen in this study is similar to that seen when using Ag nanorod-based SERS methods to analyze the spectra of human and avian mycoplasmas, albeit with lower root mean square errors.[150,151]

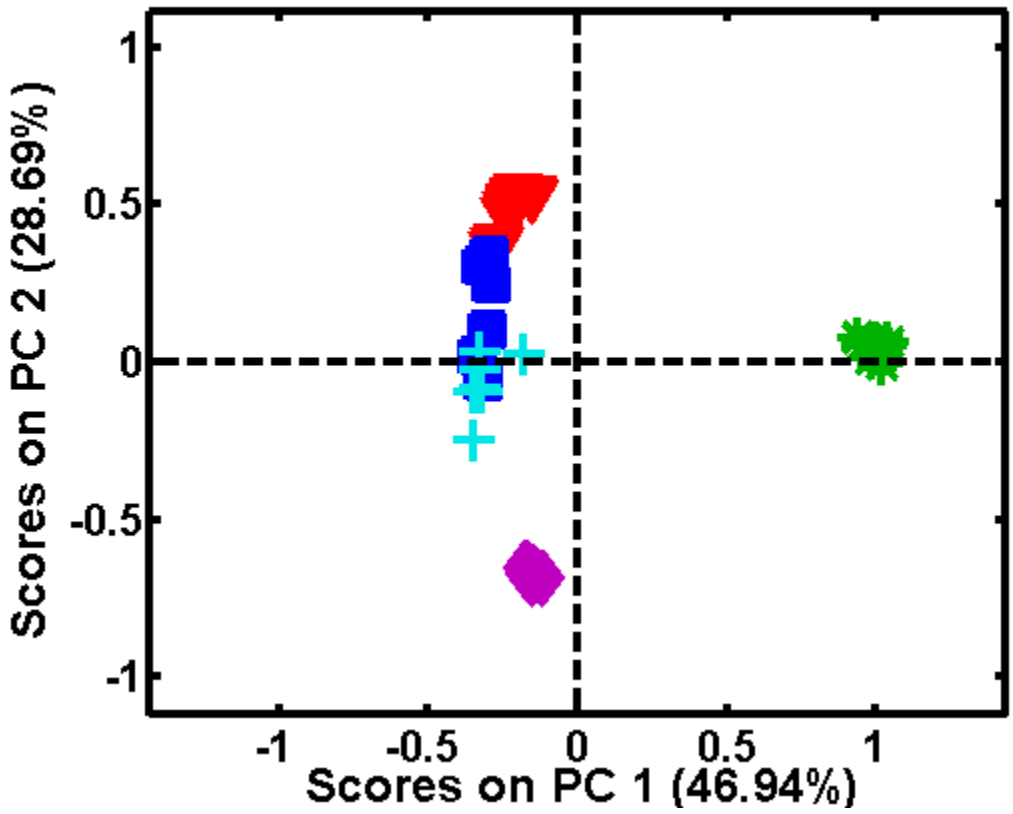

Figure 6.7 PCA scores plot corresponding to: the background sample Ag/PAH/PSS (⓿), the negative control SiMS/PAH/PSS/PAH/Ag/PAH/PSS(*), and *M. pneumoniae* strains FH (⑤), M129 (+), and II-3 (◇) The PC model was constructed from the SERS spectra of the corresponding species using the spectral range 1650-700 cm^{-1}.

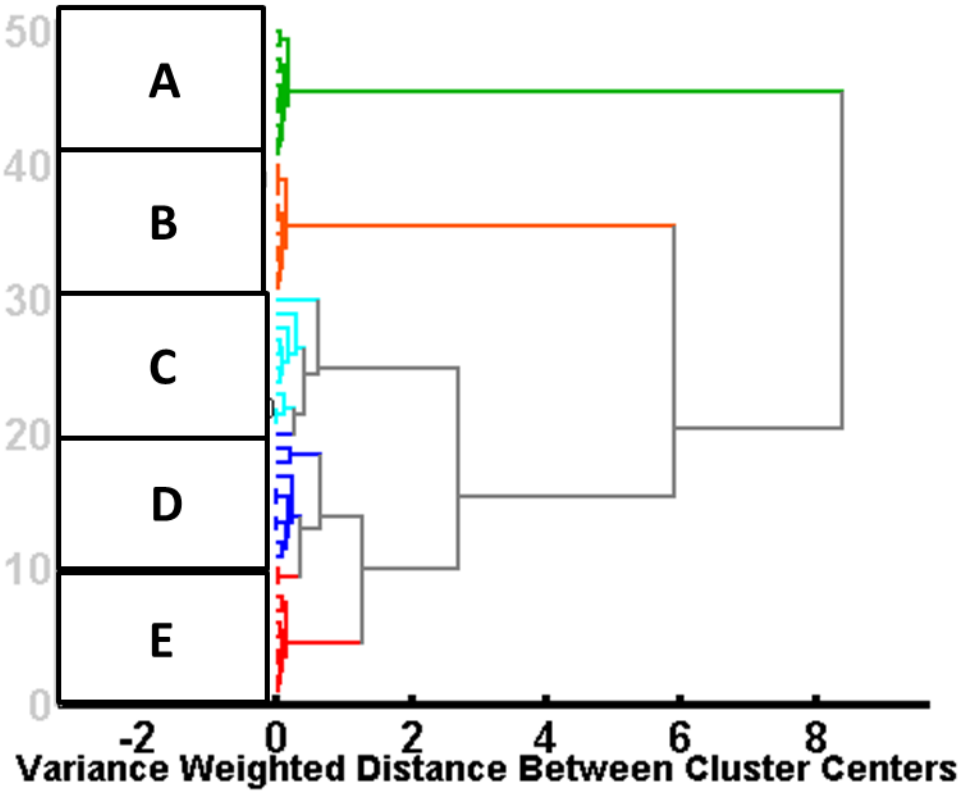

Figure 6.8 A hierarchical cluster analysis dendrogram derived from the PC scores of the *M. pneumonie* species and controls. The nodes group into five recognized clusters and are labeled according to the samples: (A) negative control SiMS/PAH/PSS/PAH/Ag/PAH/PSS; *M pneumoniae* strains: (B) II-3; (C) M129; (D) FH; and (E) background sample Ag/PAH/PSS.

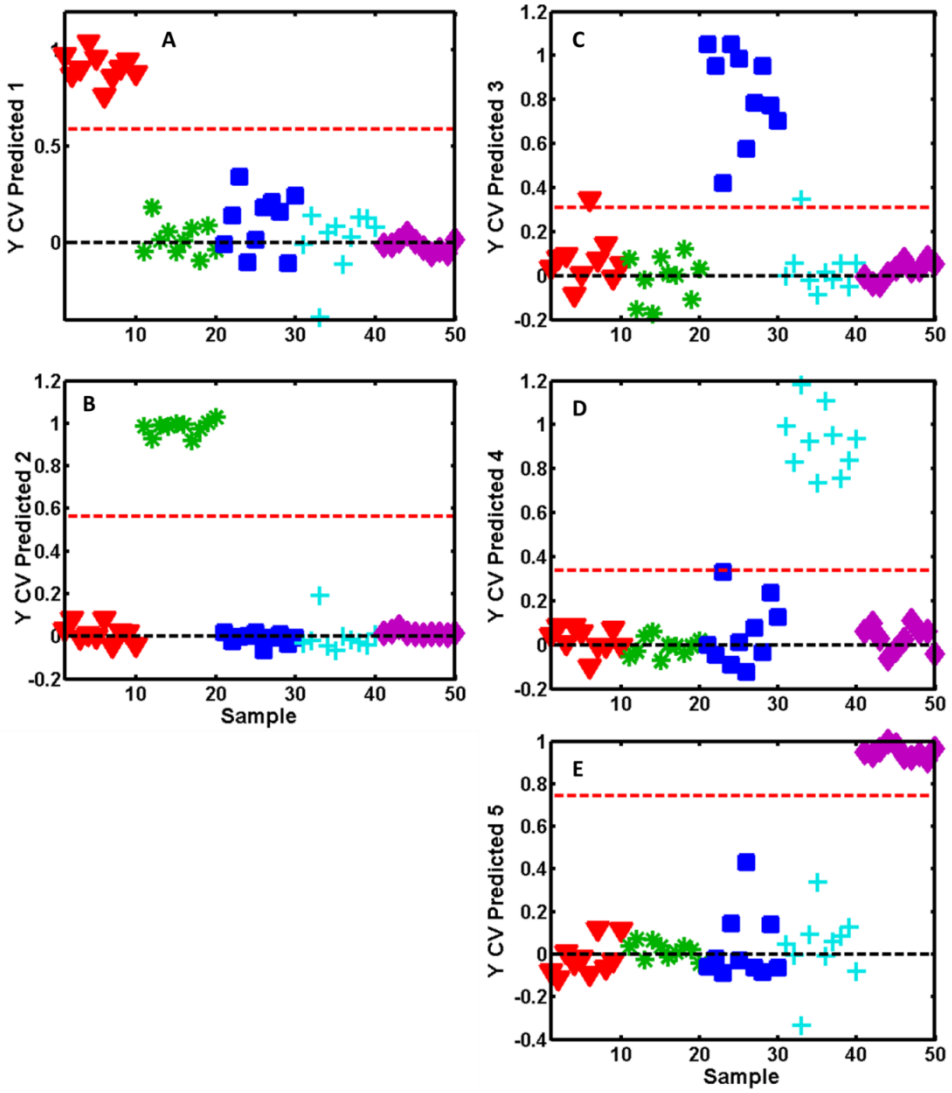

Figure 6.9 PLS-DA cross-validated prediction plot based on the LBL SERS spectra of the *M. pneumoniae* strains. Horizontal red line denotes calculated class prediction threshold level. Predictions for: (A) background sample Ag/PAH/PSS(⓿);(B)negative control SiMS/PAH/PSS/PAH/Ag/PAH/PSS (☒); and *M. pneumoniae* strains: (C) FH;(⑤); (D) M129 (+); and; (E) II-3 (◇). Fifty spectra, corresponding to 10 spectra in each sample category, are represented in this plot.

Table 6.2 Discrimination results from PLS-DA analysis of the LBL SERS spectra of three *M. pneumoniae* strains, including a control sample (Ag/PAH/PSS).

Modeled Class[a,b]	Sensitivity[d] (CV)	Specificity (CV)	Class Error[e] (CV)	RMSECV[f]
Control[c]	1.000	1.000	0.000	0.123
Negative Ctrl[g]	1.000	1.000	0.000	0.428
M129	1.000	0.975	0.012	0.151
FH	1.000	0.975	0.012	0.104
II-3	1.000	1.000	0.000	0.113

[a] Fifty total spectra were used, 10 for each modeled class. Before calculation, spectra were pre-processed by calculating 1st derivatives, followed by vector normalization and mean centering.
[b] Four latent variables, accounting for 85.09% of the captured variance, were used in this model.
[c] Background sample consisted of Ag nanoparticles derivatized with PAH and PSS layers.
[d] CV, cross-validation based on Venetian blinds method with 7 splits.
[e] Class Error, classification error after cross-validation.
[f] RMSECV, root-mean square error after cross-validation.
[g] Negative control sample consisted of silica microspheres coated with PAH and PSS, and modified with polymer-derivatized Ag nanoparticles..

6.5 Conclusions

M. pneumoniae is a respiratory pathogen that accounts for widespread bronchitis and pneumonia. Unfortunately, the complexity of laboratory culture complicates diagnostic strategies. Currently, the lack of a simple, rapid, clinical diagnostic test delays initiation of appropriate treatment, and increases the risk of continued transmission and long-term sequelae. The purpose of this research was to determine whether charged polyelectrolyte layers could increase the number of contact points between the AgNPs nanoparticles and the bacterial cell for improved SERS spectral quality, thereby increasing accuracy in identification and differentiation of different mycoplasma strains.

In this study, *M. pneumoniae* whole cells were encapsulated layer-by-layer with polyelectrolyte thin films incorporating Ag nanoparticles (figure.6.1). Three strains of *M. pneumoniae* were used as model organisms to illustrate the effectiveness of the LBL encapsulation procedure. The encapsulated bacteria were investigated using both SEM and fluorescence microscopy techniques. SEM images (figures.6.2A and 6.2B) of the uncoated bacteria showed the expected elongated morphology for mycoplasma, while LBL encapsulation resulted in significant aggregation of the mycoplasma cells into multicellular clusters (figures. 6.2C and 6.2D). Fluorescence microscopy using both a mycoplasma-specific dye (figures.6.3A and 6.3B) as well as a dye-labeled polyelectrolyte (figures.6.3C and 6.3D) showed that the polyelectrolytes co-localized with the bacteria, although the areas of polyelectrolyte coverage were not uniform. This behavior is not unexpected for this bacterium, which demonstrates a tendency to clump and aggregate in both clinical isolates and laboratory cultures.[278]

SERS spectra of the LBL encapsulated *M. pneumoniae* strains showed a high degree of reproducibility with good signal-to-noise, making detailed spectral band assignments possible (figures. 6.4-6.6 and Table 6.1). Unsupervised methods of multivariate statistical analysis, including PCA (figures.6.7) and HCA (figures.6.8), showed a high degree of qualitative class discrimination based on principle components calculated from the SERS spectra. The model-dependent method PLS-DA (figure.6.9 and Table 6.2) provided quantitative statistical measurements of the sensitivity and specificity of the LBL encapsulation method for discrimination between the three *M. pneumoniae* strains, background, and negative control. In this case, both sensitivity and specificity were between 97 – 100% for all classes modeled, with a low (0.1 – 0.4) root mean square error. This study demonstrated that LBL polyelectrolyte encapsulation combined with Ag nanoparticle SERS provides a promising platform for accurate identification and differentiation of *M. pneumoniae* strains. The advantage of the LBL method is that charged polyelectrolyte layers should increase the number of contact points between the AgNPs nanoparticles and the bacterial cell for improved SERS spectral quality. In the case of mycoplasma, the use of LBL encapsulation also solves the problem of cell lysis that may complicate spectral analysis.[151] The technique also shows promise for adaptation to sample preparation of *M. pneumoniae* infections in clinical specimens.

Acknowledgements

The research described here was supported by the U.S. Public Health Service through NIH Grants AI096364 and GM102546.

CHAPTER 7

Polyelectrolyte Matrix Encapsulation of Mycoplasma Cells Enhanced Raman Study of the Development of a Potential Bacterial Diagnostic Tool

Rivera-Betancourt, O.E, Krause, D.C., Dluhy, R.A., To be submitted to

7.1 Abstract

Mycoplasma pneumoniae is a cell wall-less bacterial pathogen which is the major cause of respiratory disease in humans that accounts up to 20% of all community-acquired pneumonia. Modern diagnostic and detection methods of mycoplasma infections are limited primarily by several factors but mainly by poor success rate at culturing the bacteria from clinical samples. We have developed and previously described a new platform for mycoplasma detection by layer-by-layer (LBL) encapsulation of *M. pneumoniae* cells. Here we report the coating of *M. pneumoniae* cells and commensal mycoplasma strains with Ag nanoparticles in a matrix of the polyelectrolytes poly(diallyldimethylammonium chloride) (PDADMAC) and poly(styrene sulfonate) (PSS) to obtain high quality SERS spectra to increase the accuracy in identifying and differentiating different mycoplasma strains. It was found that the silver nanoparticles as well as the polyelectrolytes were incorporated on the mycoplasma cells depending on the ionic strength of the salt solution. The thickness as well as the topography of the encapsulated cells were found to vary due to the incorporation of different molecular weights of PDADMAC used to develop the capsule as seen on the SEM and AFM images. SERS spectra of the LBL encapsulated whole cell mycoplasma strains showed a high degree of reproducibility with good signal-to-noise, making detailed spectral band assignments possible. Unsupervised methods of multivariate statistical analysis, including PCA and HCA, although it didn't show a high degree of qualitative class discrimination, it did classify separately the *M. pneumoniae* strains from the clinical commensal strains for both encapsulation methods. The model-dependent method PLS-DA provided quantitative statistical measurements of the sensitivity and specificity for both LBL encapsulation methods for discrimination between

the *M. pneumoniae*, commensal strains and control. The technique shows promise for adaptation to sample preparation of *M. pneumoniae* infections in clinical specimens and represents a valuable alternative to current bacterial diagnostic techniques.

7.2 Introduction

One of the most important causes of respiratory tract infections, including pneumonia, is *Mycoplasma pneumoniae*. This cell wall-less bacterium is human pathogen that causes around 20-40% of all community acquired pneumonia (CAP) and approximately 18% of cases requiring hospitalization for children.[34,52,238,279] *M. pneumoniae* may also play a role in a wide range of extrapulmonary infections, autoimmune disorders and chronic diseases, such as asthma and arthritis.[280] Infections are acquired through respiratory secretions and manifest in a nonspecific upper respiratory tract symptoms which may progress to tracheobronchitis and atypical bronchopneumonia.

It is classified as a member of the class of *Mollicutes* and represents one of the smallest self-replicating species with respect to genome size as well as cellular dimensions.[281] *M. pneumoniae* strains are divided into two main groups (subtypes 1 and 2) and some minor variants.[282] The first studies of molecular discrimination in these two subtypes[283] showed that the gene encoding the major adhesion *M. pneumoniae*, the 170 kDa P1 protein, can exist in two regions within the P1. The P1 protein is major antigenic factor of *M. pneumoniae* , typing of the P1 strains does have epidemiological value. But thus far has not been informative regarding differences among *M .pneumoniae* isolates relevant to clinical features of the bacterium.[284] Serologic testing such as enzyme-linked immunoassay is the most widely used commercially available test to diagnose *M. pneumoniae* infections.

Results may be comparably sensitive as those obtained by PCR.[34] PCR can exhibit high sensitivity and yield positive detection sooner than serological testing but is limited by issues of reliability, standardization and cost.[280] Lack of a simple, reliable, rapid diagnostic test is thus a critical barrier to the improved control of *M. pneumoniae* disease.

Application of nanotechnology to biosensor development is yielding direct, rapid, and sensitive pattern-recognition approaches for detection of infectious agents. It has been shown that nanofabrication by glancing angle vapor deposition produces Ag nanorods arrays (NA) exhibiting extremely high electromagnetic field enhancements for surface-enhanced Raman spectroscopy (SERS).[95 90 240] Paired with chemometric analysis this platform can rapidly detect and distinguish with great sensitivity and specificity the Raman spectra of viruses,[241 207 234] mycoplasma,[151] and mycobacteria,[92] and shows great promise in its potential to improve of *M. pnumoniae* infections.

In this work the platform that will be used is based on a layer-by-layer assembly, this technique has been used in the preparation of planar multilayer films[243 248,285] and the encapsulation of living bacteria cells.[286,287 112,288] Our strategy will be based on the placement of silver nanoparticles after the deposition of the polyelectrolytes on the surface of *M. pneumoniae* whole cells. Many methods for cell encapsulation, coating and entrapment within polymers have been investigated in the past.[247,286,289,290] Many of the work in the past consisted of thin film buildup via layer-by-layer (LBL) assembly has been particularly attractive for cell coating applications due to its nanoscale precision and the ability to modify surface characteristics. The basis of this method is the assembly of composite layers by alternating absorption of oppositely charged species onto a charged template. Cells are usually suitable templates for coatings because they can be used for

sacrificial cores for the development of microcapsules while their metabolic activity is preserved.[246,291] It is usually known that living microbial or human cells are negatively charged.[251] Through this process, the negatively-charged cell membrane can be encased within a cationic polymeric shell; oppositely charged polymers can be subsequently adsorbed to achieve a desired thickness and surface composition.

Polyelectrolytes facilitate adhesion of nanoparticles to biological cells, thus provide stability of sandwhich like polyelectrolyte/nanoparticle coat and suppress the particles to reach the internal part of the cell wall. Layer by layer assembly with colloidal metal nanoparticles such as gold and silver has been demonstrated with fungi[252] and bacteria.[161] The positively charged polyelectrolytes and charged AgNPs should utilize the LBL technique to increase the number of contact points between nanoparticles and the bacterial cell for the improved SERS spectral quality. Incorporation of nanoparticles into the cell wall architecture will possibly allow the modification of cells as templates and elements of biosensors. Herein we developed a SERS biosensing platform with the LBL encapsulated mycoplasma whole cells along with chemometrics, for the detection and differentiation of *M. pneumoniae* strains with statistically significant sensitivity and specificity.

7.3 Materials and Methods

Chemicals. Poly(diallyldimethylammonium chloride) solution (PDADMAC, M_w <100,000), Poly(diallyldimethylammonium chloride) solution (PDADMAC, M_w ~100,000-200,000) and Sodium (polystyrene sulfonate) (PSS, M_w~70,000) were purchased from Sigma-Aldrich, USA; 70 nm PELCO® NanoXact™ Silver Colloids/Nanoparticles- Citrate Capped were purchased from Ted Pella, Inc, USA.

Culture and Preparation of Bacterial Strains. Two major wild-type *M. pneumoniae* subtypes, M129 and FH, [292] as well as strain II-3, a spontaneously arising avirulent mutant derived from M129,[293 256] were used in this study. Mycoplasmas were grown to log phase with a 1 µl/ml inoculation. The wild-type and mutant strains were grown in 25 ml of SP4 medium[257 258] in cell culture flasks at 37˚C for 72-96 h and harvested when the phenol red pH indicator turned orange (pH approx. 6.5). The growth medium for the M129 and FH strains was poured off and the cells were scraped from the flask surface into 2.5 ml of fresh SP4 medium. For the II-3 strain, which fails to attach to plastic, cell suspensions were collected through centrifugation at 25,000×g for 25 min at 4˚C and then suspended in 2.5 ml of fresh SP4. Mycoplasma suspensions were syringe-passaged 10 times with a 25-gauge needle to disperse the cells, and aliquots of each were serially diluted for plating to measure colony-forming units (CFU). A 500 µl aliquot of each strain was transferred to a separate tube and fixed in SP4 by adding 500 µl of 8% formaldehyde (pH 7.0-7.5) for a final 4% formaldehyde concentration and stored at 4˚C until used for cell encapsulation. *M. orale* and *M. salivarum* strains were obtained from the American Type Culture Collection and

kindly shipped, cultured, and harvested by the University of Alabama-Birmingham (UAB) diagnostic mycoplasma laboratory.

The encapsulation of the cells was done in three different steps. The formation of the layers was prepared by a wet chemical assembly strategy through electrostatic interaction. The first sets of steps involved encapsulating the bacterial strains in a layer-by-layer fashion alternating deposition of PDADMAC/PSS. Polyelectrolyte solutions were dissolved in 3M NaCl at the concentration of 1mg/ml for PSS and 0.1% w/v for the PDADMAC solutions of different molecular weights. The second set of steps involved coating the silver nanoparticles (AgNP) with PDADMAC, this is done to develop the last layer of the LBL architecture. The encapsulation of the intact mycoplasma whole cells was done as two separate procedures. In our work we used two different PDADMAC solutions. They just both differed in molecular weight as described in the materials section. The encapsulated cells prepared with the poly(diallyldimethylammonium chloride) solution (PDADMAC, M_w <100,000), we will refer to them as LMW or low molecular weight. The encapsulated cells prepared with the (PDADMAC, M_w ~100,000-200,000) we will refer to them as HMW or high molecular weight.

STEP 1- Bacteria/PDADMAC/PSS-Phase. The surface charge of the bacteria is assumed to be negative, so the procedure was started by coating with a positively charged polymer such as PDADMAC. A 300 µl of the cell suspension was mixed with 350µl of 0.1% w/v solution of PDADMAC and 350µl of 1mg/ml solution of PSS and mix for 15 min at room temperature. Centrifuge the mix for 20min at 17,000 rpm at 4˚C. Discard excess polyelectrolyte solution and resuspend the cells. Wash with cold ultrapure water.

Centrifuge the cell suspension for 10 min at 17,000 rpm at 4˚C after each wash. Discard the excess water at this point don't resuspend the cells.

STEP 2- AgNP/PDADMAC-Phase- In a centrifuge tube add 1ml of Ag colloid solution, centrifuge at 7000rpm for 10min at 4˚C. Discard excess colloidal solution and add an additional 900 µl of colloidal solution. To this content add 100ul of 0.1%w/v PDADMAC. Mix the contents for 15 mins.

STEP 3- Encapsulation of Cells. Mix the contents of STEP 2 into the contents STEP 1, Mix the contents for 15 mins. Centrifuge for 10min at 17,000 rpm at 4˚C. Discard excess solution, resuspend, wash two times with cold ultrapure water. Centrifuge for 10min at 17,000 rpm after each wash. Resuspend the cells in water at the end.

The LMW *M. pneumoniae* encapsulated cells, to fully resuspend the capsule in water an additional 30 minutes of sonication in a heated water bath were required. For the HMW encapsulated cells an additional 60 minutes of sonication were required.

SEM Characterization of the Encapsulated Cells

We imaged the mycoplasmas using two procedures, one involved growing the mycoplasmas directly on glass coverslips (procedure not shown) and a more simple approach in which it is the method that will be discussed within the contents.

Scanning Electron Microscopy (SEM) images of the uncoated bacteria and the encapsulated cells were obtained using a Zeiss 1450EP SEM. The samples were fixated as described elsewhere[35] but with a few modifications. Samples of cells were prepared by

dispersing 100 µL of a cellular suspension on the surface of a glass coverslip precoated with poly-L-Lysine and incubated overnight at 36˚ C. Then samples were primarily washed twice in sodium cacodylate buffer, 5 min each wash. Then samples were then postfixed in 1% OsO4 in sodium cacodylate buffer for one hour. The samples were then washed once afterwards with sodium cacodylate buffer for ten minutes and then rinsed with water twice for five minutes. The SEM coverslips were then treated with an ethanol dehydration series sequentially (5 mins each step) with 25, 50, 75, 85, 95, and three 100% washes, critical point dried , and sputter coated with 20 nm diameter gold prior to examination.

AFM Characterization of the Encapsulated Cells

Atomic force microscopy (AFM) images were obtained using a Bruker Innova Microscope using tapping mode with a large area scanner. The samples were prepared by dispersing 100 µL of the encapsulated cellular suspension on the surface of a glass coverslip precoated with poly-L-Lysine and incubated overnight at 36˚ C. Then the samples were primarily washed twice in sodium cacodylate buffer, 5 min each wash. The samples were then washed twice with water for five minutes. The glass coverslips were then treated with an ethanol dehydration series sequentially (5mins each step) with 25, 50, 75, 85, 95, and three 100% washes and finally critical point dried.

SERS Measurements of the *M. pneumoniae* Strains

SERS spectra were acquired using a Renishaw (Hoffman Estates, IL) inVia confocal Raman microscope system using a 785 nm near-IR diode laser as the excitation source. Radiation from the diode laser was attenuated to <15 mW using a series of neutral density

filters and focused onto the sample using a 20× microscope objective. Spectra were collected between 1800 – 400 cm^{-1} and integrated for 30 s per scan with 1 scan per spectrum. The SERS spectrum of each of the encapsulated cells with polyelectrolytes and AgNP were collected applying a 30 µl sample droplet to a copperfoil substrate that was cleaned thoroughly with copious amounts of methanol and acetone.[294] The droplet was dried in an incubator and then rinsed thoroughly with ultrapure water and stream of nitrogen prior to analysis. A minimum of ten spectra were collected for each bacterial strain from different locations on each individual substrate. The encapsulated cells were prepared in duplicates for each *M. pneumoniae* strain to test for reproducibility of the LBL method.

Classification using chemometric analysis

Raman spectra were imported into GRAMS AI (Version 8.0 Thermo Electron Corp, Waltham, MA) for spectral averaging and baseline correction. Chemometric analysis was carried out with MATLAB version 7.2 (The Mathworks, Inc., Natick, MA), using the PLS Toolbox version 7.0 (Eigenvector Research Inc., Wenatchee, WA). SERS spectra in the range 1650 – 700 cm^{-1} were used for classification. Prior to analysis, first derivatives of the SERS spectra were calculated using the Savitzky-Golay method with a 2^{nd} order polynomial and a fifteen-point window. Each data set was then vector normalized and mean centered. Multivariate statistical analysis of the extracted mycobacterial lipids was performed using principal components analysis (PCA), hierarchical cluster analysis (HCA), and partial least squares discriminant analysis (PLS-DA) using the PLS Toolbox software. The calculated principal components were used as inputs to the HCA algorithm, which used the Ward's method algorithm to evaluate minimum variances within clusters.

7.4 Results and Discussion

Characterization of LBL Encapsulated Mycoplasma. We used *M. pneumoniae* wild-type strain FH as a model organism to illustrate the LBL encapsulation procedure. Synthetic polycation/polyanion pairs, i.e. PDADMAC/PSS, were used to produce layered shells that covered the bacteria. Two PDADMAC samples of different molecular weights were used to encapsulate the mycoplasma as separate procedures in order to explore the effects it can potentially have in the overall appearance and hardiness of the capsule. The procedures were repeated numerous times. Since this is a wet chemical procedure, the mycoplasma cells were suspended with the polyelectrolyte solutions, centrifuged, washed a number of times in between the consecutive alternation of PDADMAC and PSS. The polyelectrolytes were used to produce layered shells that covered the cell wall-less bacteria. PDADMAC was deposited as the first layer to balance the negative surface charge of the mycoplasma cells, followed by the polyanion PSS, and then a final layer of PDADMAC was added before the addition of the final layer of AgNP colloids. This final layer provides higher stability to the structure, as well as sensitivity to temperature and permeability.[243 285 251] Figure 7.1 shows a schematic representation of the deposition process. The purpose of our study was not to analyze in depth the effect of the ionic strength of the salt solution (e.g. NaCl in water) to prepare the polyelectrolyte solutions, we didn't end up using 0.5M NaCl due to the poor aggregation of the polyelectrolyte pair on the surface of the mycoplasma whole cells(data not shown). As we increased the ionic strength of the salt solution to 3.0M NaCl, this resulted in an increase of the adsorption of the PDADMAC/PSS pair towards the mycoplasma cells due to decreasing electrosteric repulsion interactions[295 296 297 298]

which could lead to a thicker and more coiled structure as will be discussed later in this section.

The cells were washed after every LBL deposition cycle to remove any excess polyelectrolytes. Deposition of the polyelectrolyte-coated mycoplasma cells with AgNPs could initially be monitored visually. Originally, the cell suspensions were clear; however, the suspended cells acquired a brownish-yellow color during the deposition of the AgNPs. The color of the suspension was due to the presence of the bound AgNPs in the LBL matrix and not to free, unbound nanoparticles, as these were removed during the washing steps. The presence of the polyelectrolyte-encapsulated AgNPs enhances the aggregation of the mycoplasma cells into clusters. Aggregation has previously been reported in LBL polyelectrolyte encapsulation of other bacterial species, with the extent of aggregation depending on the fixation protocols as well as the bacterial cell surface biochemistry.[161] The chemical natures of the polyelectrolytes, as well as the solution ionic strength, have a strong influence on the polyion complexes that are formed as well.[285 263 262 264]

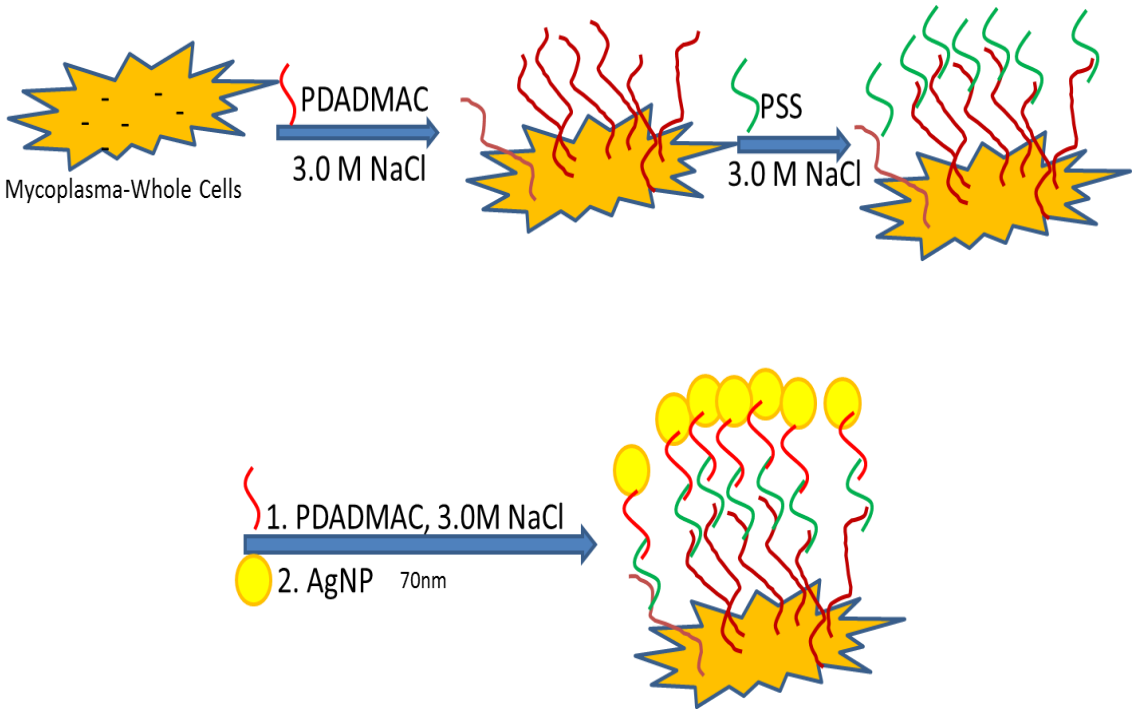

Figure 7.1 Schematic illustration of the encapsulation of the *M. pneumoniae* whole cells into the polyelectrolyte shells containing silver nanoparticles.

Cell morphology of wild-type *M. pneumoniae* M129 wild type and mutant II-3 was directly characterized by SEM, as previously described.[35] Mycoplasmas have a marked tendency toward pleomorphism due to the absence of a cell wall.[259 260 299] We used an environmental SEM (ESEM) to monitor the deposition process for our work. This technique is commonly employed for the detection of different living organisms on a "wet" and/or uncoated state.[261] ESEM micrograph of the uncoated, wild-type FH cells is shown in figure.7.2A. The cell in this image appeared as a round ovoidal shape with a smooth surface. The cells tended to clump together. The size variation was approximately 1.5-2 μm in length and 0.5-1 μm in width.

The polyelectrolyte-encapsulated cells were also characterized by ESEM, as seen in figures.7.2B and 7.2C. Comparing the smooth surfaces seen on the uncoated bacteria cells in figure.7.2A, the encapsulated FH cells when using HMW PDADMAC incorporating AgNP's of figure.7.2.B, show rough and more peaky features. The encapsulated FH cells when using LMW PDADMAC incorporating AgNP's of figure.7.2.C shows a smooth and in some points of the surface, grainy features. In both cases we see more of an organized encapsulated ovoidal structures in comparison to the uncoated cell. Although the density of the coverage with nanoparticles is not visibly high, a number of single and aggregated nanoparticles can be seen on cell surfaces.

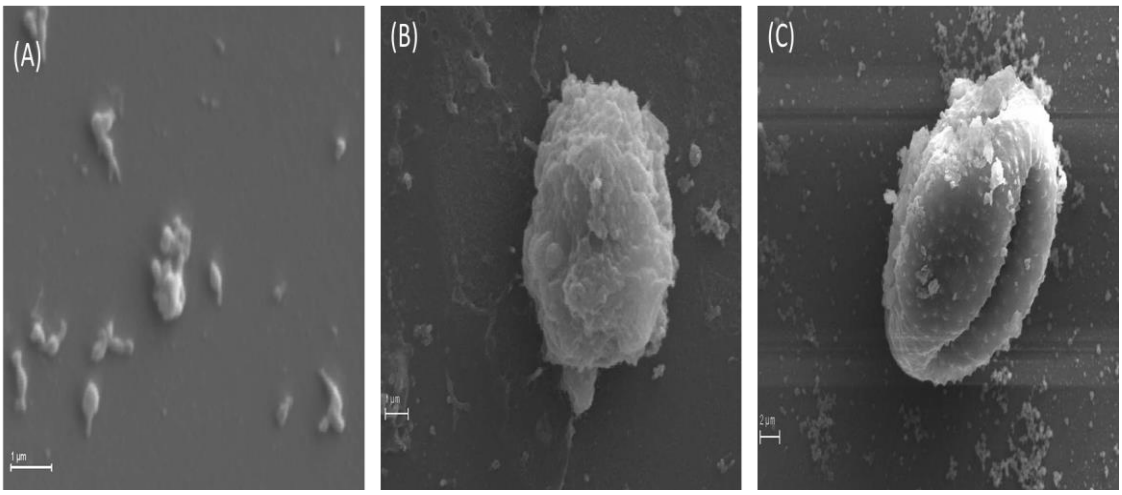

Figure 7.2 SEM images of (A) *M. pneumoniae* FH uncoated whole cells, (B) *M. pneumoniae* FH/PDADMAC/PSS/PDADMAC/AgNP encapsulated cell when using PDADMAC-HMW, (C) *M. pneumoniae* FH/PDADMAC/PSS/PDADMAC/AgNP encapsulated cell when using PDADMAC-LMW.

Atomic force microscopy was applied to visualize the surface structure and measure the thickness of the polyelectrolyte encapsulated FH cells. We carried out all experiments in tapping mode in order to investigate the morphology, thickness and robustness of the cells.[288] Images taken of the encapsulated FH cells when using LMW PDADMAC incorporating AgNP's are shown in figures 7.3A and 7.3B. The topographical images reveal a smooth and porous surface mostly throughout the encapsulated cell. When rastering the surface throughout the coverslip we noticed that there were minor capsule height differences found that ranged from 250-408 nm. In addition, the phase images (figure 7.3C) reveal consistency of polymer as well as nanoparticle coverage around a given cell. Images taken of the encapsulated FH cells when using HMW PDADMAC incorporating AgNP's are shown in figures 7.3D and 7.3E. As seen from the SEM micrographs the AFM images also reveal show rough and more peaky features in comparison to the encapsulated cells done with LMW PDADMAC. We also noticed smaller as well as minor capsule height differences found that ranged from 134-300 nm, the phase images (figure 7.3F) reveal consistency of polymer as well as nanoparticle coverage around a given cell. As can be seen in the AFM images depicting the cells encapsulated in shells doped with silver nanoparticles form thin shells. This could be due to the tendency of an exponential growth mechanism of PDADMAC/PSS [263] on the bacteria due strong polymer interactions that led to a decrease in film thickness.[286 300] We noticed experimentally that the capsules prepared at a higher molecular weight of the PDADMAC polyelectrolyte and according to the core-shell morphology of the LBL encapsulated mycoplasma cell structure, the thickness of this capsule should be higher.[301]. In both types of capsules showed within since we used 3.0 M NaCl solutions to prepare the

polyelectrolyte solutions, at high salt concentrations the polyelectrolyte complex formed encapsulating the bacteria should have a coiled conformation explained by the screening of the opposing polyelectrolyte charges of the chains, which diminishes the internal repulsion of the charged monomers.[302] [132]

SERS Spectra of the M. pneumoniae Strains. The Raman spectra of bacteria reflect predominantly phenotypic information arising from proteins, nucleic acids, lipids, carbohydrates and endogenous biomolecules.[303] [304] In order to detect and identify pathogens of interest, it is necessary to ensure that the observed Raman bands are specific to the targeted organisms but not to the overall environment, e.g. the media or solvents. The mycoplasma SP4 growth medium is removed from the sample prior to spectral analysis; therefore, it was not included in the background analysis. Instead, the background and controls used in these samples were from the LBL assembly process that includes both polyelectrolytes as well as the AgNPs, i.e., the silver nanoparticle-polyelectrolyte layer Ag/PDADMAC(**HMW**)/PSS and Ag/PDADMAC(**LMW**)/PSS on separate experiments.

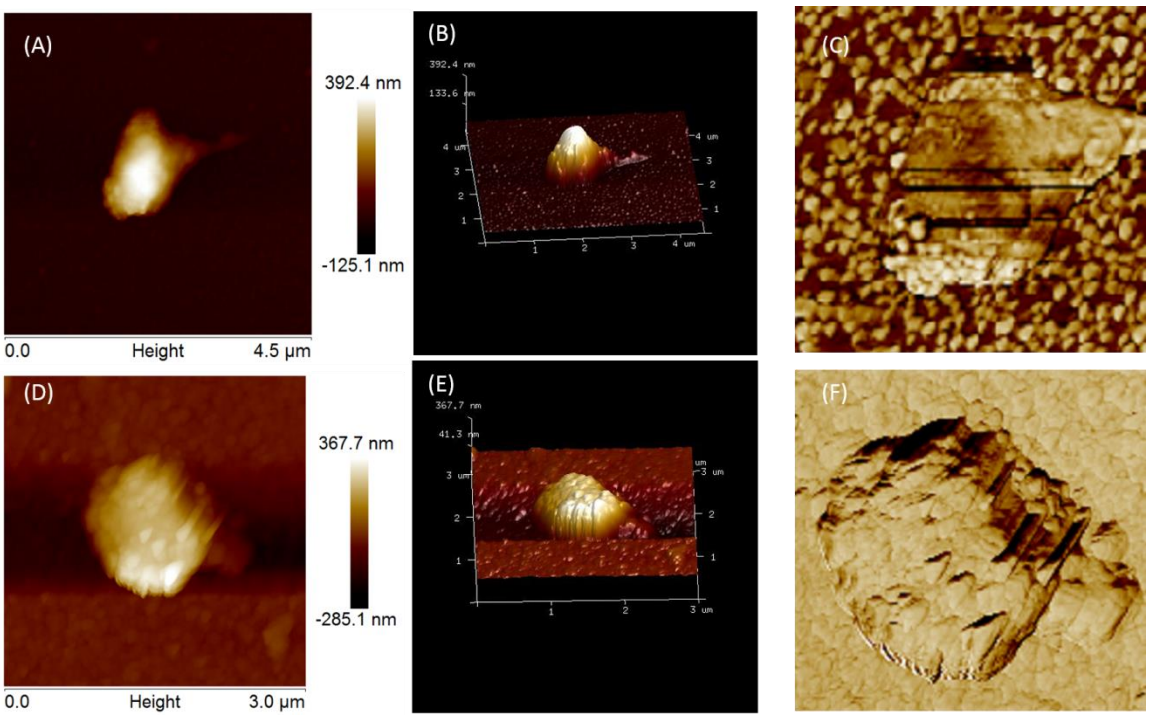

Figure 7.3 Tapping mode AFM images of (A) topographic image 2D view of FH coated with the polyelectrolytes and AgNP's when LMW PDADMAC is used, (B) topographic image 3D view of FH coated with the polyelectrolytes and AgNP's when LMW PDADMAC is used, (C) phase image view of FH coated with the polyelectrolytes and AgNP's when LMW PDADMAC is used,(D) topographic image 2D view of FH coated with the polyelectrolytes and AgNP's when HMW PDADMAC is used, (E) topographic image 3D view of FH coated with the polyelectrolytes and AgNP's when HMW PDADMAC is used, and (F) phase image view of FH coated with the polyelectrolytes and AgNP's when HMW PDADMAC is used.

The SERS spectra obtained experimentally were highly reproducible for both LBL mycoplasma encapsulation methods. Figure 7.4 and figure 7.5 show the SERS spectra taken from samples with the encapsulation procedures Mycoplasma/PDADMAC(**HMW**)/PSS/PDADMAC(**HMW**)/AgNP and Mycoplasma/PDADMAC(**LMW**)/PSS/PDADMAC(**LMW**)/AgNP respectively. The samples were prepared as described above, and then spot dried on a piece of copper foil. Spectra of the dried samples were obtained from 10 different locations. The SERS experiments were repeated several times with a newly prepared sample each time. This approach was used when acquiring the SERS spectra of all the *M. pneumoniae* strains, as well as the Ag/PDADMAC/PSS control in separate experiments.

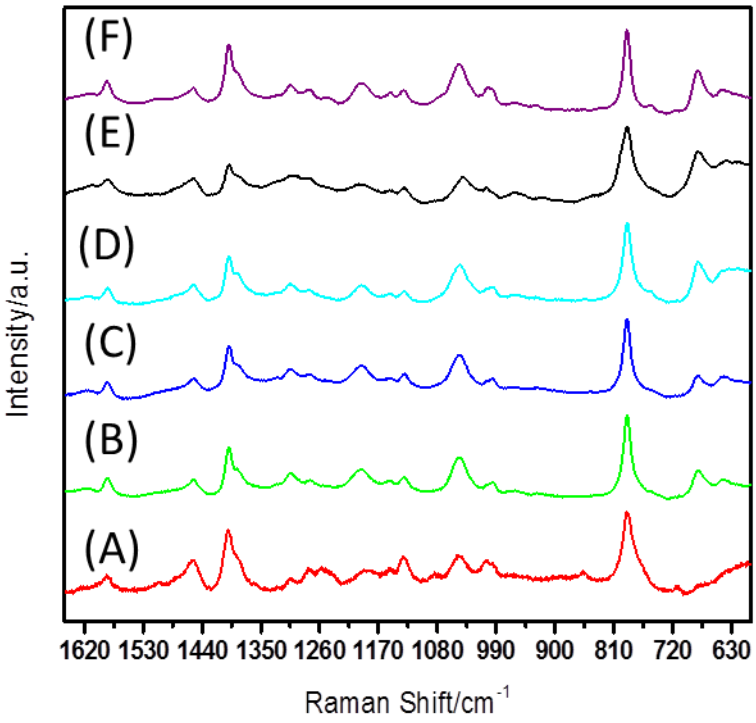

Figure 7.4 Representative SERS spectra when the bacterial encapsulation procedure involved HMW PDADMAC for (A) the control sample AgNP/PDADMAC/PSS. The *M. pneumoniae* strains, (B) M129, (C) FH and, (D) II-3. The mycoplasma commensal strains, (E) *M. orale* and (F) *M. salivarum*. Each spectrum is an average of 10 spectra obtained for each sample. The spectra have been baseline corrected and normalized for visualization.

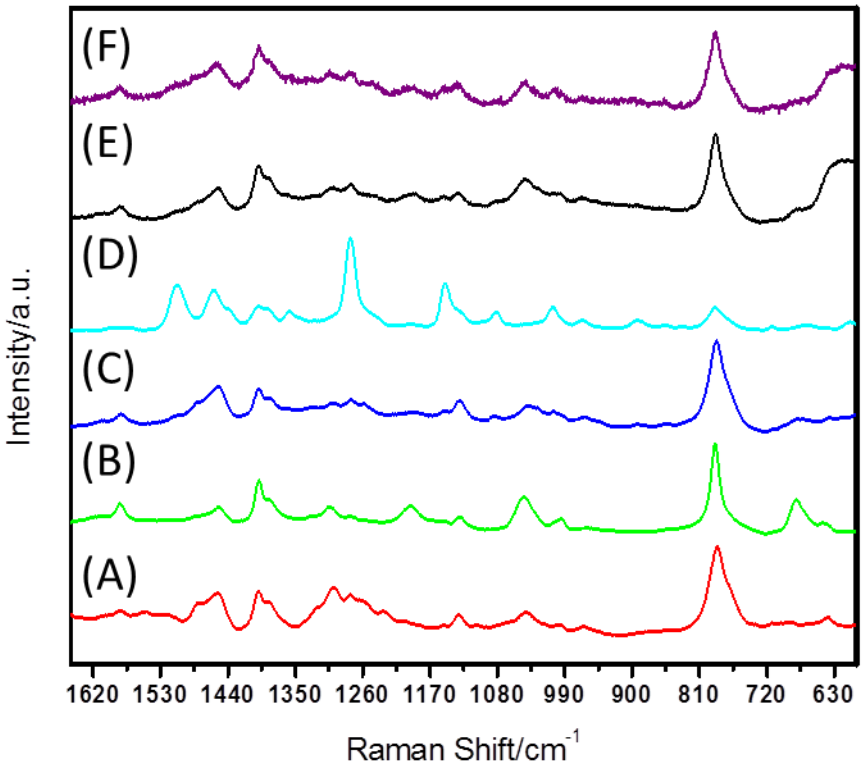

Figure 7.5 Representative SERS spectra when the bacterial encapsulation procedure involved LMW PDADMAC for (A) the control sample AgNP/PDADMAC/PSS. The *M. pneumoniae* strains, (B) M129, (C) FH and, (D) II-3. The mycoplasma commensal strains, (E) *M. orale* and (F) *M. salivarum*. Each spectrum is an average of 10 spectra obtained for each sample. The spectra have been baseline corrected and normalized for visualization.

SERS is highly dependent on, and sensitive to, the proximity of the AgNPs to the cell surface. Functional groups such as COO^- and NH_2^+ may define the modes of interaction between the bacterial surface of the intact mycoplasma whole cells and the AgNPs.[60 253 159 142] Figures 7.4 and 7.5 illustrate the average SERS spectra in the 1800 – 400 cm^{-1} range of the three *M. pneumoniae* strains, two commensal mycoplasma strains *M. orale* and *M. salivarum* and the Ag/PDADMAC(**HMW**)/PSS and Ag/PDADMAC(**LMW**)/PSS control samples used in separate experiments. For each sample, including the control, ten spectra were taken at each of two different spots, baseline-corrected, and normalized for visualization. Inspection of the SERS spectra from figures 7.4 and 7.5 we saw the Raman vibrations, with the exception of a few differences found in the spectra of the some of the strains. From figure 7.4 the Raman vibrations at 1614, 1587, 1452, 1401, 1382 and 1304cm^{-1} among other bands were found in all spectra, although their intensities varied.[272] The spectra of the different strains did exhibit some differences; for example, the band at 1611cm^{-1} was present just in M129 and FH while a vibration at 1356 cm^{-1} was just present in the mutant strain II-3. The commensal strains *M. orale* and *M. salivarum* did not show any characteristic bands that visually could distinguish these strains separately from the two wild type strains as well as the mutant strain II-3. The spectra obtained from the other encapsulation method is shown in figure 7.5, SERS spectra showed similar vibrations as those found in figure 7.4. The Raman vibrations at 1611, 1585, 1452, 1399, 1274 and 964 cm^{-1} among other bands were found in all spectra, although their intensities varied. As stated previously some different vibrational bands are found among all strains and the commensal strains did not show any characteristic bands that could visually distinguish these strains separately from the others. The most significant spectral differences were seen

to occur in the 1000 – 400 cm^{-1} region in both encapsulation methods. A list of the observed Raman vibrations in figures.7.4 and 7.5 as well as their tentative assignments are found in Table 7.1 and Table 7.2 for both encapsulation methods discussed within the contents. Detailed spectral band assignments have been published elsewhere.[161] [60,89,159,253,272]

Classification of the M. pneumoniae Strains. The cell wall-less prokaryote Mycoplasma pneumoniae causes bronchitis and atypical pneumonia in humans.[155,305,306] It is mostly bound by only a cell membrane containing numerous surface-exposed membrane proteins and glycolipids. Raman bands that where present in the encapsulated intact whole cells obtained experimentally were more frequently associated with vibrations present in amino acids and lipids.

Table 7.1 Raman bands appearing in the LBL SERS spectra of *M. pneumoniae* strains when using the Mycoplasma/PDADMAC(**HMW**)/PSS/PDADMAC(**HMW**)/AgNP encapsulation

Raman Shift cm^{-1}					
M129	FH	II-3	*M.orale*	*M. salivarum*	Vibrational Mode Assignment
1613	1614	1615	1616	1617	ν_s(C=O) carboxylic acid
1587	1587	1587	1587	1587	Guanine, Adenine (ring stretch)
1452	1452	1452	1452	1452	δ(C-H_2) sci.; CH3 antisym. bend
1401	1401	1401	1401	1401	C-O-H bend; (CH2)n in-phase twist
1382	1382	1382	1382	1382	C-H def
		1356			Amide III
1304	1304	1304	1304	1304	Amide III
1279	1279	1279	1279	1279	C-O-H bend,Amide III
		1247			ν_{as}(COC); δ(CH),Amide III
1195	1195	1195	1195	1195	δ (C-H), Tyr
1152	1152	1152	1152	1152	C-N and C-C stretch
1131	1131	1131	1131	1131	C-N and C-C stretch
1048	1048	1048	1039	1048	C-C skel. str in alkane; ν_{as}(COC) ; C-N and C-C stretch
996	994	994	1001	1003	phenylalanine
962	962	962	962	962	C-C str
930	930	930	930	930	Thr, Trp, C-COO stretch Tyr
792	792	792	792	792	Cytosin, uracil (str, ring); CH2 in-phase rock
		752		752	Trp; Glucose, Galactose
684	684	684	684	684	adenine
641	641	641	641	641	Phenylalanine (skeletal)

Table 7.2 Raman bands appearing in the LBL SERS spectra of *M. pneumoniae* strains when using the Mycoplasma/PDADMAC(**LMW**)/PSS/PDADMAC(**LMW**)/AgNP encapsulation

Raman Shift cm^{-1}					
M129	FH	II-3	*M. orale*	*M. salivarum*	Vibrational Mode Assignment
1611	1611				v_s(C=O) carboxylic acid, Tyr
1585	1585	1585	1585	1585	Guanine, Adenine (ring stretch)
		1508			C-O-H bend; $(CH_2)_n$ in-phase twist
1452	1452	1452	1452	1452	δ(C-H_2) sci.; CH3 antisym. bend
1399	1399	1399	1399	1399	C-O-H bend; (CH2)n in-phase twist
1388	1388	1388	1388	1388	C-H def
1299	1299		1299	1299	Amide III
1276	1274	1274	1274	1274	C-O-H bend,Amide III
	1257				v_{as}(COC); δ(CH),Amide III
1193	1193	1193	1193	1193	δ (C-H), Tyr
1150	1150	1150			C-N and C-C stretch
1126	1126		1129	1129	C-N and C-C stretch
1043	1043		1043	1043	C-C skel. str in alkane; v_{as}(COC) ; C-N and C-C stretch
1005	1005	1005	1005	1005	phenylalanine
964	964	964	964	964	C-C stretch, PO_4
891	891	891			COO- str carboxylic acid
792	792	792	792	792	Cytosin, uracil (str, ring); CH_2 in-phase rock
684	684	684	684	684	adenine
639	639	639	639	639	Phenylalanine (skeletal)

The excellent spot-to-spot and sample-to-sample reproducibility of the SERS Raman spectra offered by the LBL methods allows for multivariate analysis as a method for classification and identification.[274] The statistical basis for the application of chemometric techniques to vibrational spectroscopy is well established. Spectral interpretation can be accompanied by unsupervised pattern recognition methods such as principal component analysis (PCA)[142] and hierarchical cluster analysis (HCA), as well as supervised methods such as partial least squares discrimination analysis (PLS-DA).[90,94,269] We have previously used chemometric methods to analyze nanorod-array SERS spectra of human and avian mycoplasma species extracted into a water-formalin mixture for inactivation of the bacteria, a process that potentially lyses the cells,[150,151] we are also as far as we know to be the first to use the LBL method to ensure preparation of whole, intact encapsulated mycoplasma cells for use in SERS classification studies.[37] The high quality and reproducibility of the LBL SERS spectra, as seen in figures.7.4 – 7.5, demonstrate that these spectra are suitable for use in subsequent multivariate statistical processing steps. The statistical methods PCA, HCA, and PLS-DA were used to determine whether it is possible to discriminate between these five different mycoplasma strains and controls based solely on their LBL SERS spectra. PCA is the most commonly employed technique, it reduces dataset dimensionality by calculating orthogonal eigenvector projections, and facilitates establishing patterns and grouping of similar spectra,[60,240,276] reduces spectral noise and maximizes total spectral variance among spectral fingerprints for all the mycoplasma strains. Principal components were calculated from the SERS spectra of the two *M. pneumoniae* M129 and FH wild-type strain, the mutant II-3 strain, and two mycoplasma commensals *M. orale* and *M. salivarum*. Two control samples consisting of

Ag/PDADMAC(HMW)/PSS and Ag/PDADMAC(LMW)/PSS used in separate experiments depending on the encapsulation method being analyzed. The PC model consisted of 60 raw spectra, 10 spectra for each *M. pneumoniae* strain, the two commensal strains, and the control, and was calculated using LBL SERS spectra in the 1650 – 700 cm^{-1} range.

As shown in figures 7.6 and 7.7 are the PCA plots for the encapsulation methods Mycoplasma/PDADMAC(HMW)/PSS/PDADMAC(HMW)/AgNP and Mycoplasma/PDADMAC(LMW)/PSS/PDADMAC(LMW)/AgNP respectively. As seen in figure 7.6 and 7.7 these plots resulted in five different clusters in which all of them had clusters that significantly overlapped. In figure 7.6 the encapsulation method provided the best comparison of the processed spectra using principal components 1 and 2 clearly differentiated two of the mycoplasma strains, which belong to the commensal strains respectively. The control Ag/PDADMAC(HMW)/PSS sample as well as the other *M. pneumoniae* strains clustered in a region with a significant degree of overlap. On figure 7.7 we saw better results, the control Ag/PDADMAC(LMW)/PSS clustered almost entirely from that of the *M. pneumoniae* strains as well as the commensal strains. We can also observe that the model clustered separately the *M. pneumoniae* strains M129, FH, and II-3 from the mycoplasma commensal strains *M. orale* and *M. salivarum*, even though in both clustered regions there is significant overlap between the strains that the model grouped there respectively. In addition, a hierarchical cluster analysis

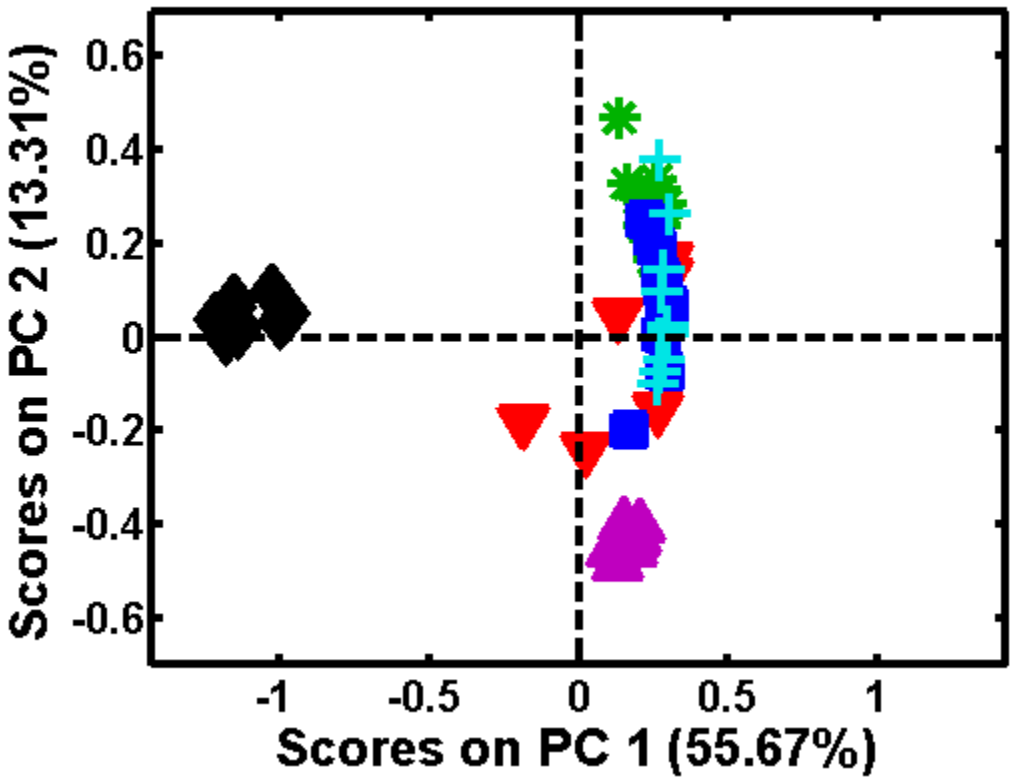

Figure 7.6 PCA scores plot corresponding to the raw SERS spectra when the bacterial encapsulation procedure involved HMW PDADMAC, control Ag/PDADMAC/PSS (⓿), and *M. pneumoniae* strains FH (⑤), M129 (☒) and II-3 (*) and the commensal strains, *M. orale* (◆), and *M. salivarum* (▲). The PC model was constructed from the SERS spectra of the corresponding species using the spectral range 1650-700 cm^{-1}.

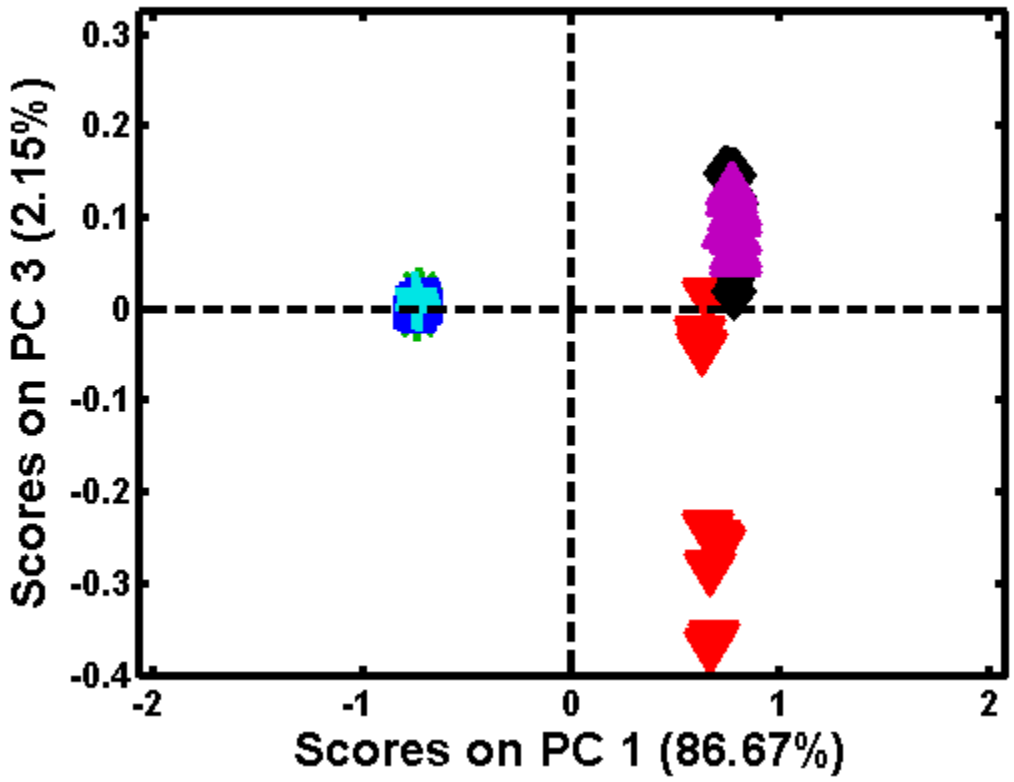

Figure 7.7 PCA scores plot corresponding to the raw SERS spectra when the bacterial encapsulation procedure involved LMW PDADMAC, control Ag/PDADMAC/PSS (0), and *M. pneumoniae* strains FH (5), M129 (☒) and II-3 (*) and the commensal strains, *M. orale* (⬢), and *M. salivarum* (▲). The PC model was constructed from the SERS spectra of the corresponding species using the spectral range 1650-700 cm^{-1}.

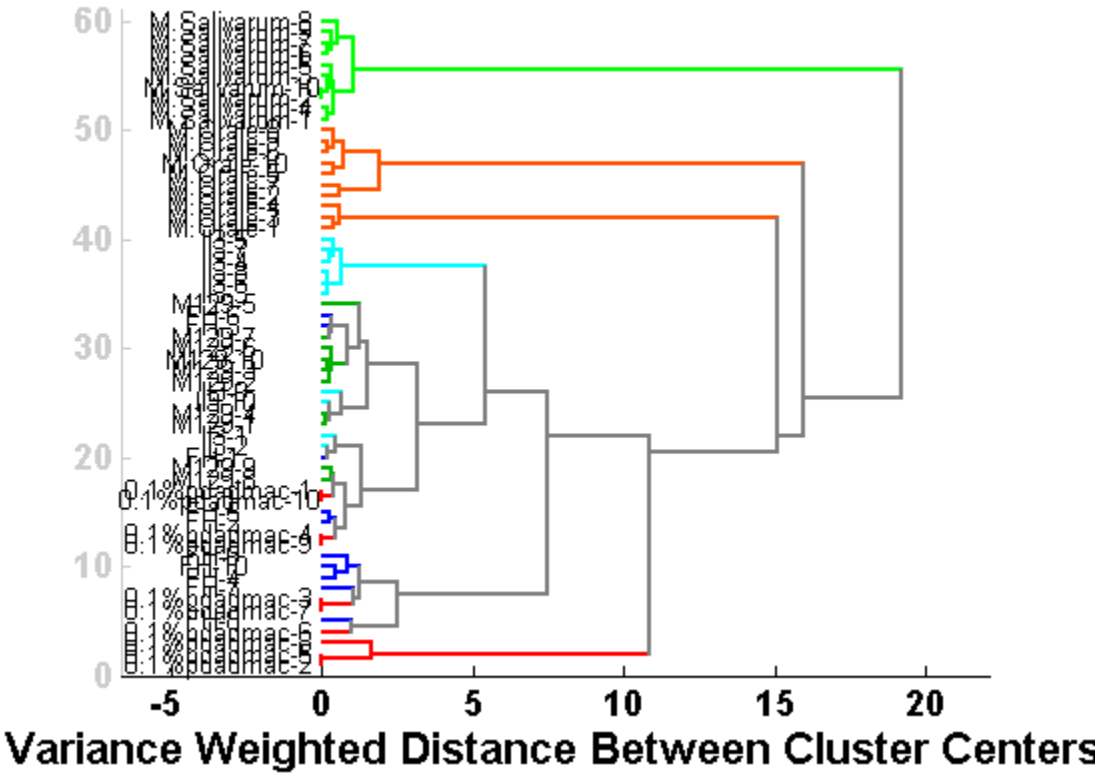

Figure 7.8 A hierarchical cluster analysis dendrogram derived from the PC scores of the *M. pneumoniae* and commensal species when the bacterial encapsulation procedure involved HMW PDADMAC. The nodes group into four recognized clusters and are labeled according to the samples. (A) II-3 (sky blue) ; (B) M129 (green) ; (C) FH (blue); (D) *M. orale* (orange); (E) *M. salivarum* (light green) and control sample (D) Ag/PAH/PSS(red).

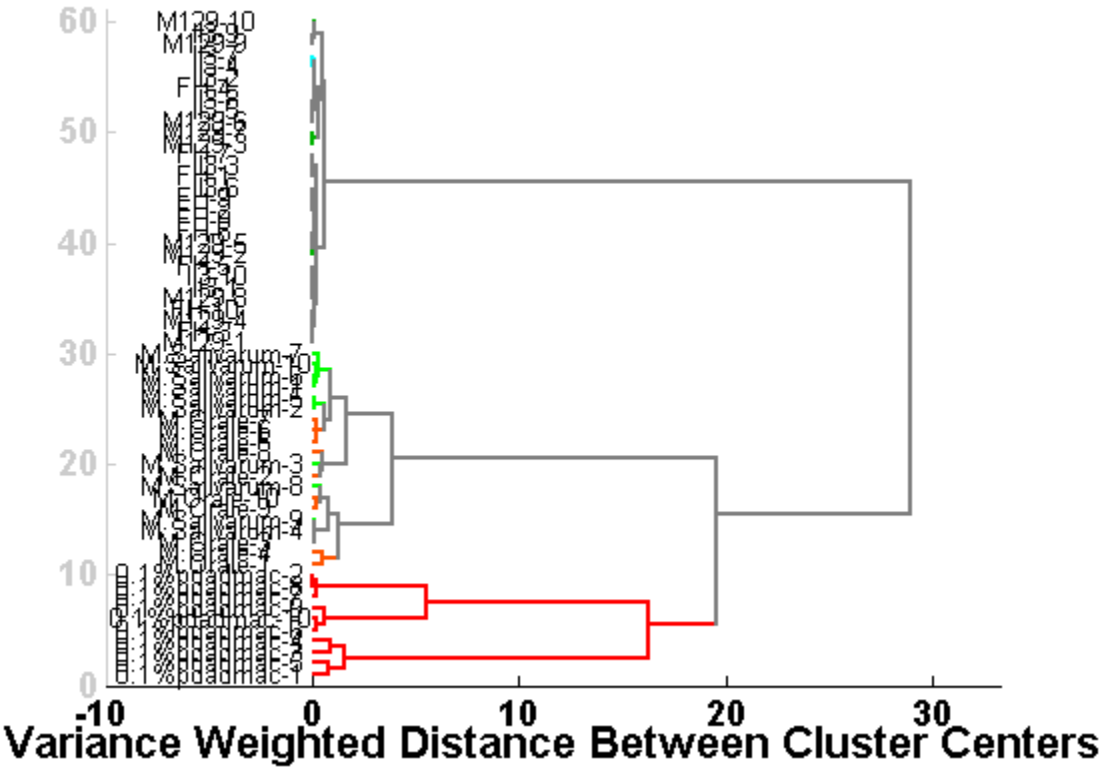

Figure 7.9 A hierarchical cluster analysis dendrogram derived from the PC scores of the *M. pneumoniae* and commensal species when the bacterial encapsulation procedure involved LMW PDADMAC. The nodes group into four recognized clusters and are labeled according to the samples. (A) II-3 (sky blue) ; (B) M129 (green) ; (C) FH (blue); (D) *M. orale* (orange); (E) *M. salivarum*(light green) and control sample (D) Ag/PAH/PSS(red).

(HCA) of the LBL SERS spectra was calculated using their principal components. The resulting dendrograms, calculated using the Ward's linkage for both encapsulation methods (figures. 7.8 and 7.9), shows that there was significant overlap with the clustered mycoplasma strains. But still the algorithm used still could cluster separately the *M. pneumoniae* strains from the mycoplasma commensal strains.

In addition to the unsupervised PCA and HCA methods, we used PLS-DA to quantitatively determine statistically significant differences among the strains. PLS-DA, unlike HCA and PCA, is a full spectrum, multivariate, supervised method whereby prior knowledge of the classes is used to yield more robust differentiation, minimizing class variation while emphasizing latent variables between or among classes.[186,307] When using PLSDA, it is important to include an appropriate control to avoid over- or under-fitting the statistical models. For this purpose we again included the control samples Ag/PDADMAC(HMW)/PSS and Ag/PDADMAC(LMW)/PSS respectively on separate models because of the two different encapsulation methods as seen in figures 7.10 and 7.11.

Figures 7.10 and 7.11 present the results of the PLS-DA analyses of the mycoplasma strains with both encapsulation methods. The models were built separately, each were generated using 60 spectra (10 per mycoplasma strain, and 10 for the control samples). The horizontal red line in each panel of figures. 7.9 and 7.10 is a calculation of a threshold value of prediction for each modeled class. Spectra with predicted values above the threshold level are determined to belong to a particular class, while spectra with the predicted values below are excluded. Each panel correspond to the predictions made for the control samples Ag/PDADMAC/PSS (figures. 7.10A and 7.11A), M129 wild-type (figures.7.10B and

7.11B), FH wild-type (figures. 7.10C and 7.11C), II-3 mutant (figures. 7.10D and 7.11D), *M. orale* (figures 7.10E and 7.11E) and *M. salivarum* (figures 7.10F and 7.11F). From figure 7.10 the analysis used 5 latent variables that captured 78.65% of the total variance, followed by cross-validation using Venetian blinds with 7 splits. It is clear from all the panels that PLS-DA was able to classify each spectrum in its class with 90-100% accuracy. The analysis from figure 11 used 6 latent variables that captured 94.87% of the total variance, followed by cross-validation using Venetian blinds with 7 splits and classifying every spectrum with a slight variability with the accuracy percentage 80-100%. Tables 3 and 4 provide the statistics from the PLS-DA models, with root-mean square error after cross-validation (RMSECV) of 0.05-0.1 for all classes.

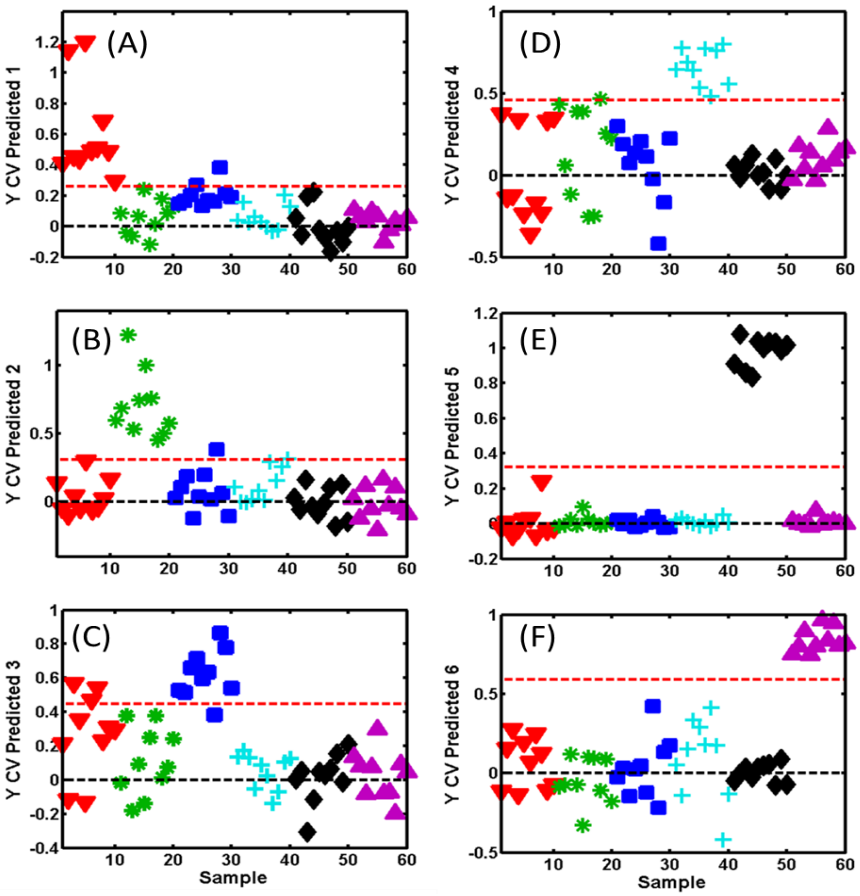

Figure 7.10 PLS-DA cross validated prediction plot based on the SERS spectra of the *M. pneumoniae* and mycoplasma commensal species when the bacterial encapsulation procedure involved HMW PDADMAC. Horizontal red line denotes calculated class prediction threshold level. (a) Predictions for control Ag/PDADMAC/PSS (⓿), (b) Predictions for Predictions for *M. pneumoniae* M129 (☒), (c) *M. pneumoniae* FH (⑤), (d) Predictions for *M. pneumoniae* II-3 (*), (e) Predictions for *M. orale* (⬢), and (f) Predictions for *M. salivarum* (▲) . Sixty spectra, corresponding to 10 spectra in each sample category, are represented in this plot.

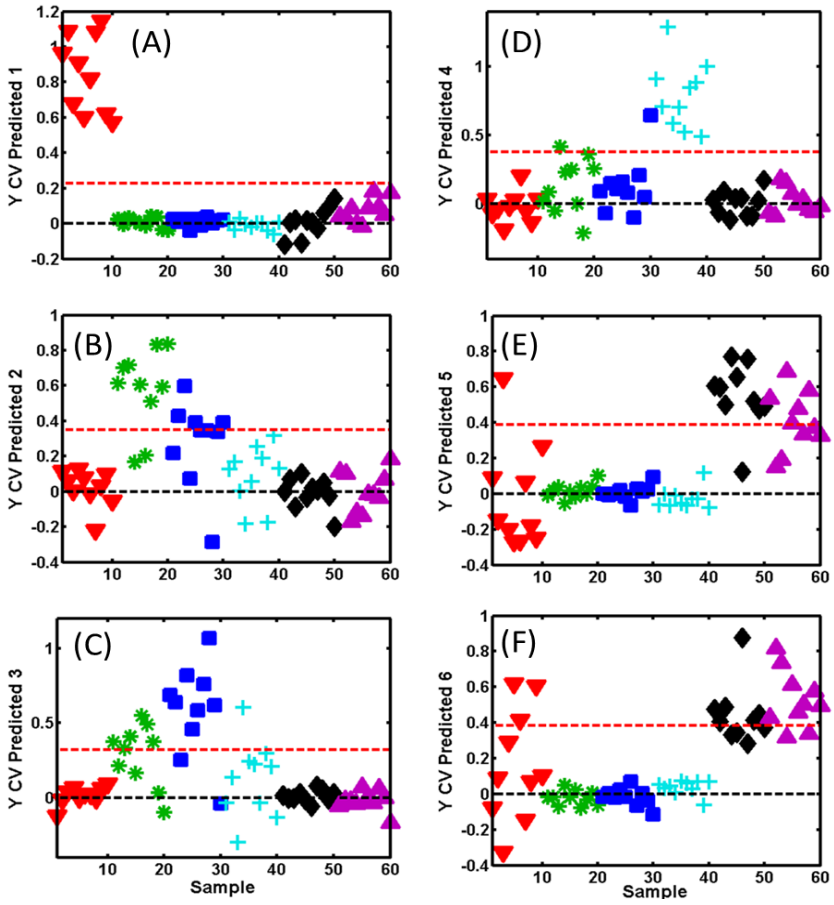

Figure 7.11 PLS-DA cross validated prediction plot based on the SERS spectra of the *M. pneumoniae* and mycoplasma commensal species when the bacterial encapsulation procedure involved LMW PDADMAC. Horizontal red line denotes calculated class prediction threshold level. (a) Predictions for control Ag/PDADMAC/PSS (⓿), (b) Predictions for Predictions for *M. pneumoniae* M129 (☒), (c) *M. pneumoniae* FH (⑤), (d) Predictions for *M. pneumoniae* II-3 (*), (e) Predictions for *M. orale* (⬢), and (f) Predictions for *M. salivarum* (▲) . Sixty spectra, corresponding to 10 spectra in each sample category, are represented in this plot

It was remarkable to see form both models the high classification sensitivity, considering that both M129 and FH are both pathogenic wild types. Key differences in several major surface proteins have been described,[308,309] such as the protein P65[310] which differs in both strains. II-3 is a mutant strain derived from the wild type strain M129, in both models it was classified separately. One of the main reasons could be attributed to the lack of production of the P30 integral membrane protein located at the distal end of the terminal organelle.[311] The lack of production of this protein makes this strain unable to attach to host cells. In both models the *M. pneumoniae* strains clearly classified separately from the commensal strains. The commensals strains are usually respiratory and urogenital tract inhabitants, but they can become pathogenic.[312] *M. orale* and *M. salivarium*, usually commensals of the oro-pharynx, may be found in the lower respiratory tract of patients with chronic bronchitis, although it is not clear that they have an effect on the severity of bronchitis.[313] Several potential challenges exist to widespread clinical application of LBL for mycoplasma detection. For example, the increased biochemical complexity and variability in clinical specimens could confound interpretation of the spectral patterns.

7.5 Conclusions

M. pneumoniae is a significant human respiratory tract pathogen with respect to both incidence and impact, but diagnostic strategies are complicated by un-describable symptoms, complex disease presentation, and the numerous challenges posed by direct culture. Serologic testing in the past has been the method of choice for diagnosis but suffers which make it impractical and unreliable. Currently, the lack of a simple, rapid, clinical diagnostic test delays initiation of appropriate treatment, and increases the risk of continued

transmission and long term sequelae. The research stated within was to develop an experimental platform in which *M. pneumoniae* and clinical mycoplasma commensal whole cells were encapsulated layer-by-layer (LBL) with polyelectrolyte thin films incorporating Ag nanoparticles to obtain high quality SERS spectra to increase the accuracy in identifying and differentiating these different strains. A diagnostic method that can identify pathogens rapidly and distinctively with minimum sample preparation has major benefits in the prevention of epidemic outbreak and infectious agents.

We previously established an LBL platform consisting of polyelectrolytes such as PAH and PSS incorporating silver nanoparticles exhibit robust and reproducible spectrum enhancement for biosensing applications for *M. pneumoniae* whole cells. Here we extend those findings we used two different polyelectrolyte layer arrays consisting of PDADMAC at different molecular weights over the mycoplasma cells to illustrate the effectiveness of the LBL encapsulation procedure. We used three strains of *M. pneumoniae* as model organisms and incorporated two mycoplasma commensal strains. The encapsulated bacteria were investigated using SEM and AFM microscopy techniques. The SEM micrographs showed the uncoated bacteria cells in figure.7.2A have very smooth surfaces in comparison the encapsulated FH cells when using HMW PDADMAC incorporating AgNP's of figure 7.2.B, show rough and more peaky features. The encapsulated FH cells when using LMW PDADMAC incorporating AgNP's of figure 7.2.C show a smooth and in some points of the surface grainy features. In both cases we saw more of an organized ovoidal encapsulated structures in comparison to previous studies done in our facilities.[37]

The AFM taken of the encapsulated FH cells when using LMW PDADMAC incorporating AgNP's shown in figures 7.3A and 7.3B revealed a smooth and porous

surface mostly throughout the encapsulated cell. When rastering the surface throughout the coverslip we noticed that there were minor capsule height differences found that ranged from 250-408 nm. AFM images taken of the encapsulated FH cells when using HMW PDADMAC figures 7.3C and 7.3D in comparison with the SEM reveal rough and more peaky features in comparison to the encapsulated cells done with LMW PDADMAC. We also noticed smaller as well as minor capsule height differences found that ranged from 134-300 nm. Based on experimental observation, the capsules prepared at a higher molecular weight of the PDADMAC polyelectrolyte and according to the core-shell morphology of the LBL encapsulated mycoplasma cell structure, the thickness should be higher.[301]

Table 7.3 Discrimination results from PLS-DA analysis of the LBL SERS spectra of three *M. pneumoniae* strains and two mycoplasma commensal strains, including a control sample (Ag/PDADMAC/PSS). Samples done with PDADMAC-high molecular weight (HMW).

Modeled class[a,b]	Sensitivity[d] (CV)	Specificity (CV)	Class Error[e] (CV)	RMSEC[f]
Control[c]	1.000	0.960	0.010	0.207
M129	1.000	0.960	0.020	0.166
FH	0.900	0.940	0.080	0.210
II-3	1.000	1.000	0.000	0.058
M.orale	1.000	1.000	0.000	0.040
M. salivarum	1.000	1.000	0.000	0.155

[a] Seventy total spectra were used, 10 for each modeled class. Before calculation, spectra were pre-processed by calculating 1st derivatives, followed by vector normalization and mean centering.
[b] Five latent variables, accounting for 78.65% of the captured variance, were used in this model.
[c] Control sample consisted of Ag nanoparticles derivatized with PDADMAC and PSS layers.
[d] CV, cross-validation based on Venetian blinds method with 7 splits.
[e] Class Error, classification error after cross-validation.

Table 7.4 Discrimination results from PLS-DA analysis of the LBL SERS spectra of three *M. pneumoniae* strains and two mycoplasma commensal strains, including a control sample (Ag/PDADMAC/PSS). Samples done with PDADMAC-low molecular weight (LMW).

Modeled class	Sensitivity (CV)	Specificity (CV)	Class Error (CV)	RMSEC
Control	1.000	1.000	0.000	0.082
M129	0.800	0.880	0.160	0.220
FH	0.800	0.860	0.170	0.248
II-3	1.000	0.960	0.020	0.159
M. orale	1.000	1.000	0.110	0.210
M. salivarum	1.000	1.000	0.190	0.230

[a] Seventy total spectra were used, 10 for each modeled class. Before calculation, spectra were pre-processed by calculating 1st derivatives, followed by vector normalization and mean centering.
[b] Six latent variables, accounting for 94.87% of the captured variance, were used in this model.
[c] Control sample consisted of Ag nanoparticles derivatized with PDADMAC and PSS layers.
[d] CV, cross-validation based on Venetian blinds method with 7 splits.
[e] Class Error, classification error after cross-validation.

SERS spectra of the LBL encapsulated whole cell mycoplasma strains showed a high degree of reproducibility with good signal-to-noise, making detailed spectral band assignments possible (figures. 7.4-7.5 and Tables 1 and 2).Unsupervised methods of multivariate statistical analysis, including PCA and HCA, although it didn't show a high degree of qualitative class discrimination, it did classify separately the *M. pneumoniae* strains from the clinical commensal strains for both encapsulation methods. The model-dependent method PLS-DA (figure.7.9 and Table 7.2) provided quantitative statistical measurements of the sensitivity and specificity for both LBL encapsulation methods for discrimination between the three *M. pneumoniae*, two commensal strains and control. It is clear from all the panels that PLS-DA was able to classify each spectrum in its class with 90-100% accuracy when using the model that incorporated PDADMAC at a high molecular weight within the LBL structure. PLS-DA models that incorporated PDADMAC at a low molecular weight within the LBL structure was able to classify each spectrum in its class with 85-100% accuracy.

This study demonstrated that LBL polyelectrolyte encapsulation combined with Ag nanoparticle SERS provides a promising platform for accurate identification and differentiation of *M. pneumoniae* strains. The speed, specificity, and ease of implementation of the LBL technique along with SERS represents a valuable alternative to current bacterial diagnostic techniques. The advantage of the LBL method is that charged polyelectrolyte layers should increase the number of contact points between the AgNPs nanoparticles and the bacterial cell for improved SERS spectral quality. The technique also shows promise for adaptation to sample preparation of *M. pneumoniae* infections in clinical specimens. A more comprehensive analysis of *M. pneumoniae* stains

with mycoplasma commensal clinical isolates is underway to assess further the discriminatory capacity of this LBL technique as a SERS platform.

CHAPTER 8

Multivariate Statistical Classification for Differentiation of Human Mycoplasma Strains by using Layer-by-Layer (LbL) Encapsulation Method

`

Choi, J.Y.‡, Rivera-Betancourt, O.E.‡, Henderson, K.C., Krause, D.C., Dluhy, R.A.

‡ Represents co-authorship

To be submitted to

8.1 Abstract

Development of a biosensing platform for a rapid, sensitive, accurate, and convenient detection of bacteria and infectious diseases has been challenging. Here, we report a powerful self-assembled layer-by-layer (LbL) based SERS technique for whole bacterial cells detection. By consecutively alternating polyelectrolytes of opposite charges, multilayer films are formed and different mycoplasma cells are encapsulated with polyelectrolytes and Ag nanoparticles and used for a rapid and sensitive identification of biologically similar bacterial strains. Pathogenic *Mycoplasma pneumoniae* M129 subtype strain and 11 human commensal strains (*Mycoplasma genitalium*, *Mycoplasma orale*, *Mycoplasma salivarium*, *Mycoplasma amphoriforme*, *Mycoplasma fermentans*, *Mycoplasma hominis*, *Ureaplasma parvum, Mycoplasma penetrans*, *Mycoplasma pirum*, *Mycoplasma spermatophilum*, *Ureaplasma urealyticum*) were analyzed for detection. Here, we determine the spectral features by ANOVA (Analysis of Variance) along with post-hoc Tukey Honestly Significance Difference (HSD) test to identify the wavebands that best differentiate between different bacteria strains.

Multivariate analysis such as Partial Least Squares Discriminant Analysis (PLS-DA) and Support Vector Machine Discriminant Analysis (SVM-DA) discriminate between different mycoplasma bacterial cells with high classification accuracy by having ~ 95 % classification accuracy for a 5 class model and ~85 % for a 13 class model, respectively. This study demonstrated the feasibility of SERS method for direct identification of whole bacterial cells without amplification or labeling.

8.2 Introduction

One of the most important causes of respiratory tract infections, including pneumonia, is *Mycoplasma pneumoniae*. This cell wall-less bacterium is a human pathogen that causes up to 20 % of all cases of community-acquired pneumonia (CAP).[34] In children, the contribution of *M. pneumoniae* to CAP may even be higher, reaching levels of more than 50 % among children aged 5 years or older, making *M. pneumoniae* the most common aetiological agent of CAP in this age group.[314] *M. pneumoniae* infection is acquired through respiratory secretions and spreads efficiently within close living quarters, with incubation periods as long as three weeks.[43,151] Symptoms tend to be nondescript and the disease often has complex and variable presentations, making definitive diagnosis challenging.[49,50,315] As a result, diagnosis of *M. pneumoniae* is often presumptive and relies heavily on the combination of physical findings and elimination of other possible causes.

Gold standard techniques in infectious disease diagnostics include microscopy, tissue culture, lateral flow immunoassays (also known as dipsticks or immunochromatographic tests — ICTs), and enzyme-linked immunosorbent assays (ELISA). These techniques are expensive, time-consuming, have limited ability to differentiate between multiple pathogens, and have a poor detection threshold. Recently, the polymerase chain reaction (PCR) technique has been adapted and utilized for pathogen detection. PCR has a higher detection limit than previous techniques (e.g., 1 million times more sensitive than lateral flow immunoassays) and can selectively differentiate pathogen strains.[316] However, both PCR and real-time PCR are limited by the same problems as the most commonly used diagnostic techniques (e.g., high equipment cost, reagent expenses and requirement of

skilled technicians) and have other hurdles such as the influence of contamination on the measurement result and analysis times of several hours.[317]

The development of an inexpensive, ultrasensitive, rapid and selective biosensing method for the detection of bacteria and infectious diseases has been challenging. In the past few years, research on metal nanoparticles has attracted great interest due to their catalytic,[318] biological,[319] and sensing[320] properties, which are different from bulk materials. Moreover, the unique optical properties that result from surface plasmon resonance in the visible range of the electromagnetic spectrum make them particularly attractive for optical applications. The enhancement of the Raman signal on certain metal nanoparticles (surface-enhanced Raman scattering, SERS)[68,69] has proven to produce signals with scattering cross sections which are able to achieve single molecule detection[72,73,321] making SERS a powerful analytical tool besides the investigation and structural characterization of interfacial and thin film systems. Numerous approaches, including physical vapor deposition, oblique angle vapor deposition,[88] and electrodepostion[322] are applied to produce SERS substrates consequently. Alternatively, SERS active substrates have been made by the simple and powerful self-assembled layer-by-layer (LbL) technique,[160,161,252] which is a distinct approach to prepare well-defined and controlled nanostructures. The characterization of bacterial cells using SERS is attractive because whole bacterial cells are used in the experiments by simple mixing with colloidal silver or gold nanoparticles most of the time. The use of colloidal nanoparticles rather than surfaces can help to bring the nanoparticles close to bacterial cell wall from as many points as possible.

The LbL technique makes multilayer film by alternating deposition of opposite charges polyelectrolytes. In our work, polyelectrolytes and silver colloidal nanoparticles

were used with different mycoplasma strains in order to improve spectral quality and reproducibility. We have previously reported a SERS platform for mycoplasma detection that is based on modification of Ag nanoparticles (AgNP) to increase their affinity towards the bacteria. The direct placement of Ag nanoparticles onto living cells can affect the viability of cells either during the process of the deposition of the nanoparticles or shortly after. Therefore, we used the of layer-by-layer (LbL) encapsulation technique, which commonly are widely used for the modification of substrates such as planar surfaces and nanoparticles.[242,254]

We recently evaluated nanoparticle encapsulated mycoplasma cells as a platform for the differentiation of *M. pneumoniae* strains using surface enhanced Raman scattering (SERS) combined with multivariate statistical analysis. Three separate *M. pneumoniae* strains (M129, FH and II-3) were studied. SERS spectra showed that the LBL encapsulation provides excellent spectral reproducibility. Multivariate statistical analysis of the Raman spectra differentiated the three *M. pneumoniae* strains with 97 – 100% specificity and sensitivity, and low (0.1 – 0.4) root mean square error. These results indicated that nanoparticle and polyelectrolyte encapsulation of *M. pneumoniae* is a potentially powerful platform for rapid and sensitive SERS-based bacterial identification.[37] In this paper, we extended our work to a panel of 11 other human commensal and pathogenic mycoplasma strain M129 to demonstrate that this biosensing platform could distinguish *M. pneumoniae* M129 from its clinically relevant closest phylogenetic relatives.

The complexity of the spectra obtained from a biological sample makes extracting relevant information and interpreting the data challenging. The most primitive data analysis procedures used for Raman studies such as peak-by-peak analysis and peak

deconvolution do not allow for extensive data extraction and are a major limitation, often involving tedious, error-prone manual analysis procedures. A wide variety of data analysis methods have recently begun to be increasingly employed for evaluating Raman spectra, including multivariate statistical methods such as principal component analysis (PCA), and various machine learning-based data mining and optimizing methods.[195] Recently, considerable effort has been directed towards developing and evaluating different procedures that objectively identify variables that contribute useful information while eliminating variables that contain mostly noise or unnecessary information.[193] Feature selection in multivariate analysis is a critical step because during this procedure, non-informative variables are removed while retaining important variables. As a consequence of this selection, better prediction results with simpler models can be obtained for quantitative determinations and/or discrimination between biologically similar samples, allowing interactive improvement of the quality of the data during the calibration procedure.[193] In this work, we report supervised classification algorithms such as partial least square discriminant analysis (PLS-DA)[37,92,151,207] and support vector machine discriminant analysis (SVM-DA) [323-325] to generate statistical models for classification to discriminate between *M. pneumoniae* M129 subtype strain versus a series of 11 human commensal strains that are biologically phylogenetically similar. These algorithms in conjunction with a feature selection method were used to select the critical spectral bands that best contribute to discriminate and classify cell types based on variations in biochemical compositions. The model accuracy results obtained from these two classification models were established based on spectral data using the whole experimental spectral range and a selected amount of wavebands were compared.

8.3 Experimental Methods

Chemicals. Poly(allylamine hydrochloride) (PAH, M_w ~15,000), sodium (polystyrene sulfonate) (PSS, M_w ~70,000) were purchased from Sigma-Aldrich (St. Louis, MO). PELCO® NanoXact™ citrate-capped Ag colloid nanoparticles (70 nm) were purchased from Ted Pella, Inc., (Redding, CA).

Culture and Preparation of Bacterial Strains. One of the major type of *M. pneumoniae* subtype M129 was used in this study. Mycoplasmas were grown to log phase with a 1 µl/ml inoculation. The wild-type strains was grown in 25 ml of SP4 medium[326 258] in cell culture flasks at 37˚C for 72 - 96 hr and harvested when the phenol red pH indicator turned orange (pH approx. 6.5). The growth medium for the M129 strain was poured off and the cells were scraped from the flask surface into 2.5 ml of fresh SP4 medium. Mycoplasma suspensions were syringe-passaged 10 times with a 25-gauge needle to disperse the cells, and aliquots of each were serially diluted for plating to measure colony-forming units (CFU). A 500 µl aliquot of each strain was transferred to a separate tube and fixed in SP4 by adding 500 µl of 8% formaldehyde (pH 7.0-7.5) for a final 4 % formaldehyde concentration and stored at 4˚C until used for cell encapsulation.

Preparation of non-M. pneumoniae human commensal and pathogenic species for layer-by-layer SERS analysis. 11 human commensal and pathogenic Mollicutes species closely related[327] to *M. pneumoniae* were grown and harvested at the University of Alabama at Birmingham (UAB). These included: *Mycoplasma amphoriforme* (ATCC A39, M6123), *Mycoplasma fermentans* (ATCC 19989), *Mycoplasma genitalium* (ATCC 49897), *Mycoplasma hominis* (ATCC Mh132), *Mycoplasma orale* (ATCC 23714), *Mycoplasma*

penetrans (UAB reference strain collection, year 1995), *Mycoplasma pirum* (ATCC 25960), *Mycoplasma salivarium* (ATCC 23064), *Mycoplasma spermatophilum* (ATCC 49695), *Ureaplasma parvum* (ATCC Up1), and *Ureaplasma urealyticum* (ATCC Uu11). For each culture, 500 µl to 1 ml of stock culture was inoculated into approximately 30 ml of SP4, Hayflick's, or 10B medium and incubated until the pH indicator turned a peach color for *M. genitalium*, *M. penetrans*, *M. pirum*, and *M. fermentans*; a rose color for *M. hominis*, *M. orale*, and *M. salivarium*; a pink color for *U. parvum*, and *U. urealyticum*; and for 72 hr for *M. spermatophilum*. At the time of harvest the cells and spent media were poured into 50 ml polycarbonate tubes and centrifuged at 8,000 rpm for 15 min, except for Ureaplasma species, which were centrifuged for 1 hr. The supernatants were decanted and the pellets suspended in 30 ml sterile PBS. The cells were washed by centrifugation at 8,000 rpm for 15 min as above, or 10,000 rpm for 1 hr for Ureaplasma species. The supernatants were then decanted and the pellets suspended in 1 ml sterile PBS, transferred to a 1.5 ml vial, and centrifuged at 14,000 rpm for 20 min. The supernatants were again decanted and the pellets suspended in 1 ml sterile PBS and syringe-passaged using a 26-gauge needle to disperse clumps. Aliquots were made for spotting onto a blood agar plate to test for contamination, and plating for CFU and color-changing unit (CCU) determination. Two 400 µl aliquots for each were centrifuged at 14,000 rpm for 20 min, the supernatant was removed, and the pellets were frozen for shipment to UGA, where they were stored at -80°C.

For SERS and quality control analysis, cell pellets were suspended in 1 ml sterile SP4 (pH 7.0) and syringe-passaged 10× with a 25 gauge needle to disperse clumps. Aliquots were then made for DNA extraction, genome equivalent determination, protein assay, and SERS

analysis. SERS samples were prepared by fixing 500 µl of suspended cells with 500 µl of 8 % formaldehyde in sterile SP4 (pH 7.0), and stored at 4°C until time of cell encapsulation.

Polyelectrolyte Encapsulation. A three-step wet chemical assembly process was used for encapsulation of the mycoplasma cells as previously described[37].

Step 1. Mycoplasma Phase. The first step involved encapsulating the bacterial cells in a layer-by-layer fashion by alternating depositions of PAH/PSS/PAH. Polyelectrolyte solutions were dissolved in 0.5 M NaCl at the concentration of 1 mg/ml. The procedure began by coating with PAH; 500 µl of the cell suspension was combined with 250 µl of 1 mg/ml PAH and 250 µl of 1 mg/ml PSS and mixed for 15 min at room temperature. This mixture was centrifuged for 10 min at 17,000 rpm at 4˚C, excess polyelectrolyte solution was discarded, and the cells were suspended and washed two additional times with cold ultrapure water. The cell suspension was centrifuged for 10 min at 17,000 rpm at 4˚C after each wash. To the same tube 250 µl of 1 mg/ml solution of PAH were added and cold water to a final volume of 1 ml. The suspension was mixed for 15 min, and then centrifuged for 10 min at 17,000 rpm at 4˚C. Excess polyelectrolyte solution was discarded and the cells suspended and washed two times with cold ultrapure water. The supernatant was discarded at this point leaving the pelleted cells.

Step 2. AgNP Phase. The second step involved coating the AgNP colloidal suspension with PAH and PSS. 1 ml of the Ag colloid suspension was centrifuged at 7000 rpm for 10 min at 4˚C. The supernatant was discarded and an additional 900 µl of colloidal suspension was added. To this AgNP suspension was added 50 µl of 1 mg/ml PAH and the contents mixed for 15 min. Then 50 µl of 1 mg/ml PSS were added and the contents again mixed

for 15 min. The suspension was then centrifuged for 10 min at 17,000 rpm at 4˚C. Excess polyelectrolyte solution was discarded and the AgNP's suspended and washed two times with cold ultrapure water.

Step 3. Encapsulation of Cells. The encapsulated mycoplasma pellet from Step 1 was mixed with the polyelectrolyte-coated AgNPs from Step 2 for 15 min and then centrifuged for 10 min at 17,000 rpm at 4˚C. The supernatant was discarded and the pellet was suspended and washed twice with cold ultrapure water, centrifuging for 10 min at 17,000 rpm after each wash. At the end of the process, the cells were suspended in water.

SERS Measurements of the Mycoplasma Strains. SERS spectra were acquired using a Renishaw (Hoffman Estates, IL) inVia confocal Raman microscope system using a 785 nm near-IR diode laser as the excitation source. Radiation from the diode laser was attenuated to <15 mW using a series of neutral density filters and focused onto the sample using a 20× microscope objective. Spectra were collected between 1800 – 400 cm^{-1} and integrated for 30s per scan with 1 scan per spectrum. The SERS spectra of the encapsulated cells with polyelectrolytes and AgNP were collected applying a 10 µl sample droplet to a copper foil substrate that was cleaned thoroughly with copious amounts of methanol and acetone. The droplet was dried in an incubator at 75˚C and then rinsed thoroughly with ultrapure water and dried under a stream of N_2 prior to analysis. A minimum of ten spectra were collected for each bacterial strain from different locations on each individual substrate. Duplicate samples of the LBL-AgNP assemblies for each *M. pneumoniae* strain were prepared to test for reproducibility of the method.

Multivariate Statistical Analysis. Raman spectra were imported into GRAMS AI (Version 8.0 Thermo Electron Corp, Waltham, MA) for spectral averaging and baseline correction. Chemometric analysis was carried out with MATLAB version 7.2 (The Mathworks, Inc., Natick, MA), using PLS Toolbox version 7.0 (Eigenvector Research Inc., Wenatchee, WA). SERS spectra in the range 1800 – 400 cm^{-1} were used for classification. Prior to analysis, first derivatives of the SERS spectra were calculated using the Savitzky-Golay method with a 2nd order polynomial and a fifteen-point window. Each data set was then vector normalized and mean centered. Throughout the preprocessing steps, non-analyte related spectral variances such as baseline variations or any other heterogeneities in the substrate were eliminated. The resulting spectra were used for all data analysis. The spectral quality was assessed by performing principal component analysis (PCA) to find an outlier. An outlier was determined based on the corresponding PCA scores and Hotelling T^2 and Q Residuals. ANOVA along with post-hoc Tukey honestly significance difference (HSD) test was used for a spectral feature selection method. Multivariate analysis such as PLS-DA and SVM-DA were performed for classification. The LibSVM software toolbox designed by Lin's lab was freely downloaded from http://www.csie.ntu.edu.tw/~cjlin/libsvm and used to perform SVM-DA.

8.4 Results and Discussion

SERS Spectra and Samples. LbL encapsulation method was utilized for mycoplasma bacterial cells detection and 10 SERS spectra for each strain were used in this study for further data analysis. SERS spectra were taken from different spots of the same sample prepared through the different time periods. They were spotted on separate copper foil cut ups. Figure 8.1 illustrates the averaged, baseline-corrected, and normalized SERS spectra

of 11 human commensal strains, M129 pathogenic strain and AgNP/PAH/PSS background (control) over the spectral range of 1800 - 400 cm^{-1} for visualization. SERS spectra of different strains show a high similarity of features with Raman vibrations at 792 (cytosine, uracil (str, rang), CH_2 in-phase rock), 1000 (phenylalanine), 1052 (C-C skel. Str in alkane, C-N and C-C stretch), 1126 (C-N and C-C stretch), 1392 (C-O-H bend, $(CH_2)_n$ in-phase twist), 1577 cm^{-1} (guanine, adenine (ring stretch)), which can be found in all spectra.

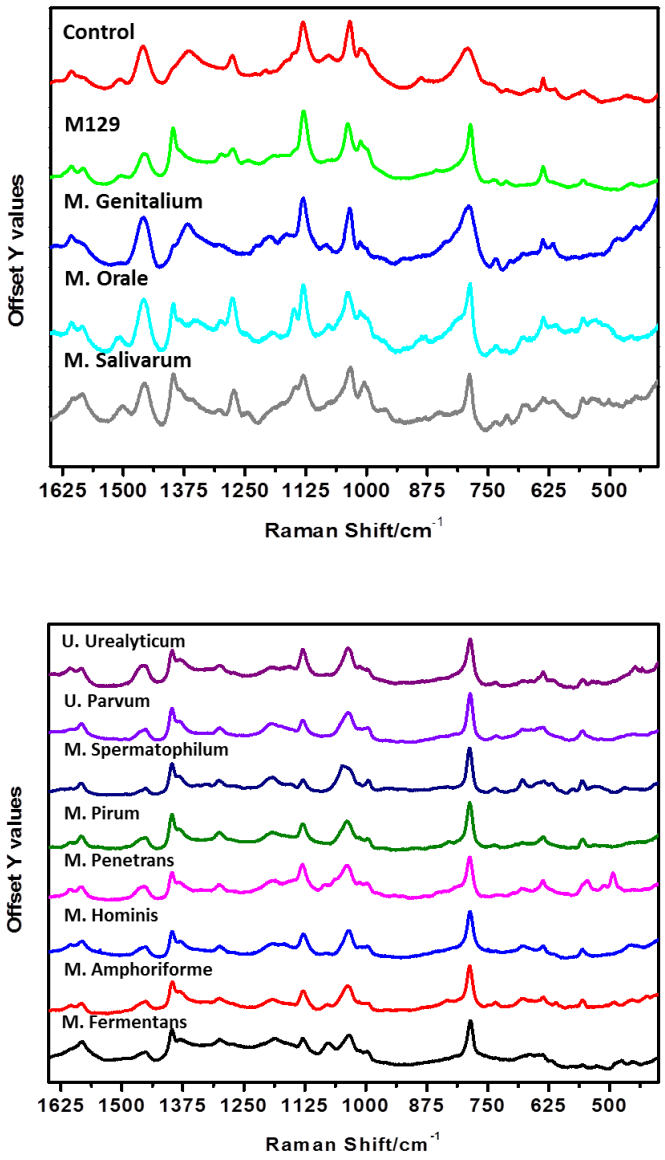

Figure 8.1 Un-preprocessed SERS spectra of 11 representative commensal strains, M129 pathogenic strains with the LBL architecture and AgNP/PAH/PSS alone (control).

Feature selection. In this study, two different models were established; a 5 class model including control, M129, *M. genitalium, M. orale, M. salivarum* and a 13 class model containing control, M129, *M.* a*mphiroforme, M. fermentans, M. genitalium, M. hominis, M. orale, M. penentrans, M. pirum, M. salivarum, M. spermatophilum, U. pavum, U. urealyticum*. One-way analysis of variance (ANOVA) followed by the post-hoc Tukey honestly significance difference (HSD) test were used for feature selection to determine the wavebands which contributes most for the best class discrimination. When we adopted the feature selection method and build the calibration model, the over-fitting and generalization issue needs to be considered. In this sense, ANOVA is a powerful method with a fast computational time to compare the difference between group means. This feature selection method can be generalized to more than two classes while avoiding increasing in Type 1 error. The spectral responses of different classes were statistically compared at every spectral location between 400 and 1800 cm^{-1}. Rejection of the null hypothesis ($H_o = \mu_1 = \mu_2 = \cdots = \mu_n$, where μ_n represents the mean spectra of the n^{th} class) was made by means of ANOVA at every spectral location with 99 % confidence limits ($\alpha = 0.01$). A total of 1780 and 2170 wavebands were selected for the 5 and 13 class models respectively, with p-values less than 0.01 at a 99 % confidence level. These selected wavebands have a discriminative capability for distinguishing spectral variance between classes from its variance within class. The post-hoc Tukey HSD test was carried out followed by ANOVA test for a statistical multiple pairwise comparison across each waveband to identify the spectral locations in which the all classes most differ from each other. A total of 312 and 940 statistically different spectral locations for the 5 and 13 class models were chosen by the post-hoc tukey test for further data analysis. A list of Raman

vibration bands selected by the feature selection methods are found in Table 8.1 and more detailed Raman band assignments can be found elsewhere.[89,160,161,270,272,273,328] These Raman vibrational bands are mostly associated with protein, nucleic acids, carbohydrates, and lipids found on the surface of the bacteria.

Classification. Multivariate analysis based on the selected wavebands by the feature selection method was exploited to overcome the spectral similarities and distinguish between biologically similar clinical human commensal mycoplasma strains for apparent classifications. Partial least-squares discriminant analysis (PLS-DA) was used to develop a classification model for discrimination. A 5 and 13 class model were under consideration for the PLS-DA classification method. Cross-validation (venetian blinds, 7 splits) was performed for the internal validation to establish the calibration model and the optimized classification model was used for the class predictions in the validation set. In the response vector Y, samples being considered as part of the group were assigned as 1 and outside of the group were assigned as 0 for each class model.

In the 5 class model, pathogenic strains *M. pneumoniae* M129, *M. genitalium*, and non-pathogenic human commensal strains of *M. salivarum* and *M. orale* were included as well as control. Table 8.2 summarizes the sensitivity, specificity, class error, and root mean square error of cross-validation (RMSECV) of the PLS-DA model for the 5 class model established based on the whole spectral range and feature selected wavebands. After the feature selection, overall sensitivity and specificity show ~95 % with RMSECV value of 0.000, 0.013, 0.038, 0.025, and 0.025, respectively for each classification model. More noticeably, when comparing PLS-DA models before and after feature selection, pathogenic strains *M. pneumoniae* M129 and *M. genitalium* were both differentiated with around

100% sensitivity and around 93 % - 98 % specificity. Structural and biological properties of *M. genitalium* that are similar to those of *M. pneumoniae,* including adherence to and invasion of epithelial cells, have suggested that *M. genitalium* might be pathogenic.[313] Because of nucleic acid sequence and surface protein similarities, other groups haven't been able to differentiate these two strains apart.[279] The classification results that we obtained through PLS-DA was highly accurate, which other group haven't been able to show. When comparing *M. salivarum* and *M. orale* non-pathogenic human commensal strains, the PLS-DA results improved after the feature selection. Both sensitivity and specificity were ranged from 95 % to 100 %. *M. salivarum* and *M. orale* are commensals of the oro-pharynx, that may be found in the lower respiratory tract of patients with chronic bronchitis, although it is not clear that they have an effect on the severity of bronchitis.[312] In some diagnosis, some groups haven't been able to tell them apart, and that leads them to screening those non-pathogenic strains together. [312] However, by using this PLS-DA method, we were able to separate them apart with high accuracy.

Table 8.1 Feature selected Raman bands in the LbL SERS spectra of the mycoplasma commensal strains.

Raman Shift, cm^{-1}	Vibrational Mode Assignment
1660	Amide I ; v_s(C=O) carboxylic acid
1616	v_s(C=O) carboxylic acid
1577	Guanine, Adenine (ring stretch)
1501	C-O-H bend; $(CH_2)_n$ in-phase twist
1445	δ(C-H_2) sci.; CH_3 antisym. bend
1392	C-O-H bend; (CH_2)n in-phase twist
1299	Amide III
1276	C-O-H bend,Amide III
1147	C-N and C-C stretch
1126	C-N and C-C stretch
1052	C-C skel. str in alkane; v_{as}(COC) ; C-N and C-C stretch
1000	phenylalanine
792	Cytosine, uracil (str, ring); CH_2 in-phase rock

A 13 class model was established based on *M. pneumoniae* M129 strain and the 11 clinical commensal strains as well as a control. Tables 8.3 and 8.4 summarizes the sensitivity, specificity, class error, and root mean square error of cross-validation (RMSECV) of the PLS-DA model for the 13 class model established based on before and after the feature selection method and the results were compared. The classification accuracy was greatly improved after the feature selection and more noticeably, specificity of *M. genitalium*, *M. orale*, *M. amphoriforme*, and *M. fermentans* were improved by a factor of ~1.5 after the feature selection. Figure 8.2A and B illustrates two representative PLS-DA scores plots of *M. orale* before and after the feature selection, respectively. Each different colored symbol represents the different strains. The optimum threshold calculated based on Bayes' Theorem is represented as the red dashed line for sample classification. In each figure, SERS spectra with predicted Y values greater than the threshold are classified as the group that was considered in each classification model while for all other spectra below the threshold were considered to be the outside of the group. Figure 8.2A represents the prediction results for *M. orale* in the 13 class model before the feature selection. Before the feature selection, *M. orale* (inverted cyan) great overlap with M129 (green square) which degrades the classification accuracy. However, after the feature selection, *M. orale* could be well separated from all other strains including M129 as shown in figure 8.2B. As indicated in Tables 8.3 and 8.4, specificity was greatly improved from 79 % to 100 %. One possibility to have a high degree of interferences between M129 and *M. orale* is because both of them are found in the respiratory tract of patients with bronchitis. However, the fact that the difference between M129 and *M. orale* were noticeable after the feature selection is because this method selects the features that are responsible for the class

differentiation. One of the possible explanations is that M129 is the actual pathogen that causes the disease. Meanwhile *M. orale* is believed to coexist with M129 but it is a non-pathogenic strain that resides in the lower respiratory tract. Similar to the *M. orale* classification model, the prediction results for *M. fermentans*(purple diamond) before the feature selection shows a highly degree of interference with M129 (green square), *M. orale* (inverted cyan), and *M. salivarum* (red star). However, after the feature selection shown in figure 8.4C, *M. fermentans* was significantly distinguishable from all other strains. The consistent results can be also found in Tables 8.3 and 8.4. The specificity of *M. fermentans* was significantly improved from 65 % to 96 %. This improvement may be related to the fact that *M. fermentans* is mostly found in a significant part of synovial fluids or biopsies from patients with different chronic inflammatory arthritides including rheumatoid arthritis (RA), but not from patients with gout, osteoarthritis or chondrocalcinosis.[329,330] These facts are significantly different from all other strains, since it has not been found in the respiratory tracts of patients. These results demonstrated that PLS-DA multivariate method in conjunction with feature selection provides a powerful tool for classification to differentiate biologically similar pathogenic and non-pathogenic strains with low values of RMSECV for both the 5 and 13 class classification models.

Table 8.2 PLS-DA results for the 5 class model using default method and feature selection method.

	Default Method				Feature Selection Method			
	Sensitivity (CV)	Specificity (CV)	Class Error (CV)	RMSECV	Sensitivity (CV)	Specificity (CV)	Class Error (CV)	RMSECV
Control	1.000	0.975	0.013	0.120	1.000	1.000	0.000	0.115
M129	1.000	1.000	0.000	0.146	1.000	0.975	0.013	0.156
genitalium	1.000	0.925	0.038	0.278	1.000	0.925	0.038	0.283
orale	0.9000	1.000	0.050	0.269	1.000	0.950	0.025	0.223
salivarum	1.000	0.925	0.038	0.224	1.000	0.950	0.025	0.203

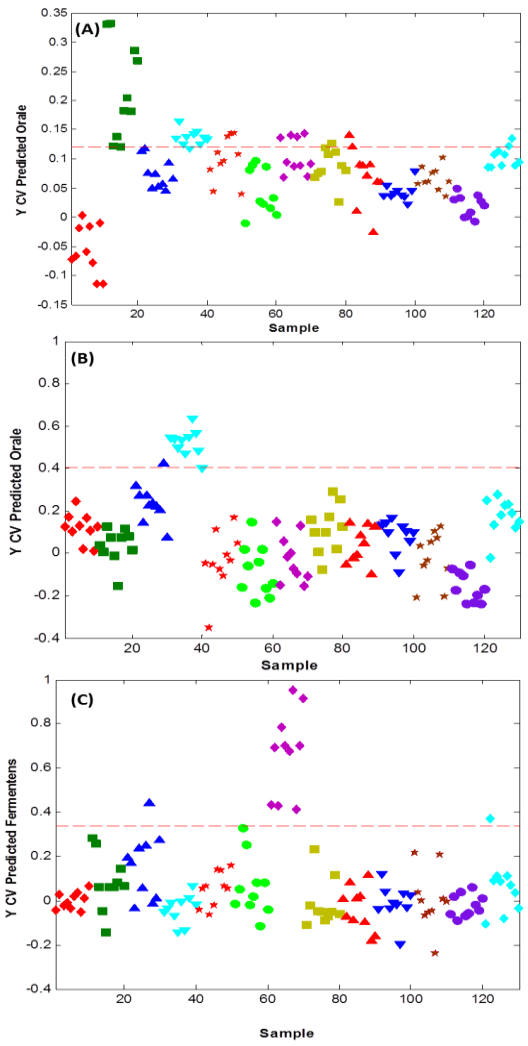

Figure 8.2 PLS-DA prediction plots for *M. orale* in the 13 class model established based on whole spectra range **(A)** and feature selected wavebands **(B)** respectively. Each colored symbol represents the PLS predicted value for an individual SERS. PLS-DA prediction plot for *M. fermentans* based on feature selected wavebands (**C**). Each colored symbol represents SERS spectra containing: M129 (◆), *M. genitalium* (■), *M. orale* (▲), *M. salivarum* (▼), *M. amphoriforme* (★), *M. fermentans* (●), *M. hominis* (◆), *Ureaplasma parvum* (■), *M. penetrans* (▲), *M. pirum* (▼), *M. spermatophilum* (★),*Ureaplasma urealyticum* (●), and AgNP/PAH/PSS was used as a control (◆).

Table 8.3 PLS-DA results for the 13 class model using the default method.

Modeled class	Sensitivity (CV)	Specificity (CV)	Class Error (CV)	RMSECV
Control	1.000	0.992	0.004	0.106
M129	1.000	0.983	0.008	0.135
M. amphoriforme	0.800	0.592	0.304	0.259
M. fermentans	0.600	0.650	0.375	0.265
M. genitalium	0.800	0.408	0.396	0.265
M. hominis	0.900	0.500	0.300	0.263
M. orale	0.900	0.792	0.154	0.259
M. penetrans	0.800	0.733	0.233	0.253
M. pirum	0.900	0.550	0.275	0.264
M. salivarum	0.900	0.908	0.096	0.211
M.spermatophilum	0.900	0.767	0.167	0.251
U. parvum	0.700	0.442	0.429	0.265
U. urealyticum	0.900	0.708	0.196	0.260

Table 8.4 PLS-DA Results for 13 class model using the feature selection method.

Modeled class	Sensitivity (CV)	Specificity (CV)	Class Error (CV)	RMSEC
Control	1.000	1.000	0.000	0.085
M129	1.000	1.000	0.000	0.103
M. amphoriforme	0.900	0.933	0.083	0.233
M. fermentans	0.900	0.958	0.071	0.146
M. genitalium	0.600	0.925	0.238	0.231
M. hominis	0.500	0.800	0.350	0.261
M. orale	0.900	1.000	0.050	0.172
M. penetrans	0.600	0.850	0.275	0.239
M. pirum	0.700	0.800	0.250	0.260
M. salivarum	1.000	0.992	0.004	0.140
M. spermatophilum	0.900	0.917	0.092	0.239
U. parvum	0.900	0.867	0.117	0.242
U. urealyticum	0.800	0.925	0.138	0.239

Support vector machine discriminant analysis (SVM-DA) is a supervised learning method for classification especially on complex and nonlinear hyperspectral datasets. This method would explain non-linearities, which are not taken into account with PLS-DA. The basic principle is to transform original data into a high-dimensional space using kernel functions. The radial basis function (RBF) kernel $K(\vec{X}_i, \vec{X}_j) = \exp(-\gamma\|\vec{X}_i - \vec{X}_j\|^2)$ was used in our work and SVM classification model was developed based on the optimal SVM parameters (C = penalty error, γ = radial width). A grid search was used for optimization of the penalty error (C) and the radial width (γ) by means of cross-validation (venetian blinds, 5 splits). The grid search space for C and γ are $[10^2, 10^{-3}]$ and $[10, 10^{-6}]$ in steps of $10^{0.5}$, respectively. The optimized C and γ through 5-fold cross-validation are 100 and 1 for both the 5 class and the 13 class model. A total number of 130 spectra were used in this analysis and 87 randomly selected spectra were assigned as calibration set and 43 randomly chosen spectra were assigned as validation set. 200 different calibration and validation sets were created and based on optimized SVM parameters, SVM was performed with 200 iterations and the results are shown in figures 8.3 and 8.4.

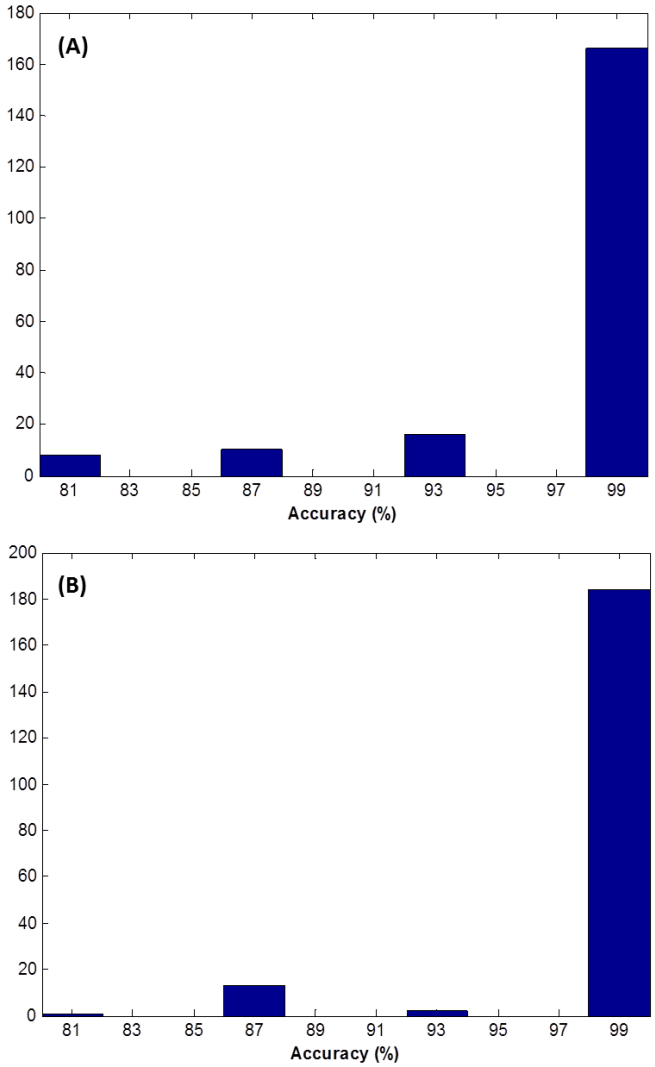

Figure 8.3 Histogram plot of the SVM-DA cross-validation (CV) classification accuracy results of 200 iterations in the 5 class model established based on the whole spectral range (A) and feature selected wavebands (B).

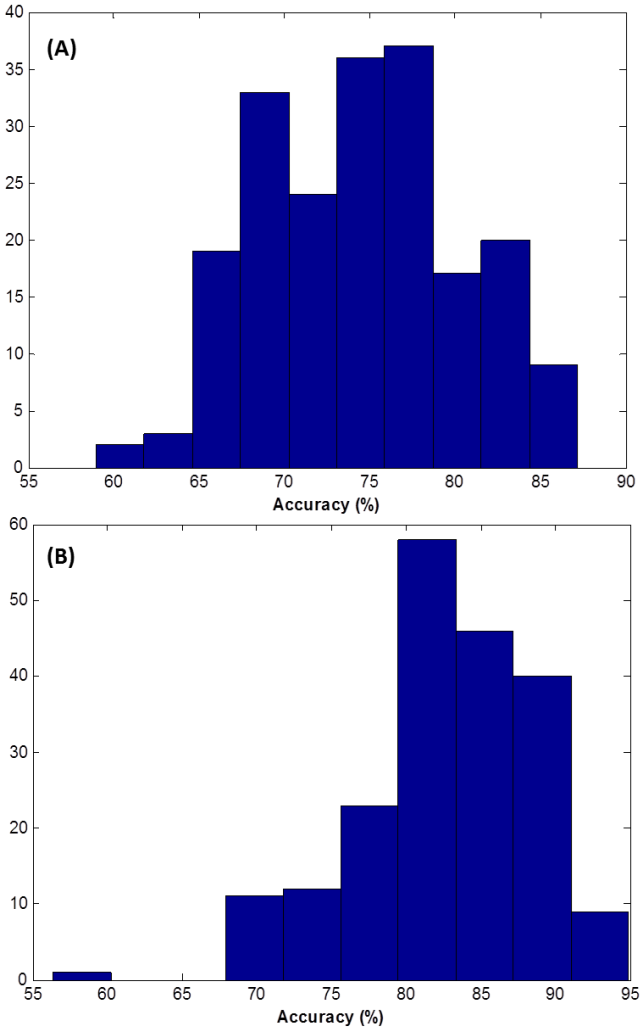

Figure 8.4 Histogram plot of the SVM-DA cross-validation (CV) classification accuracy results of 200 iterations in the 13 class model established based on the whole spectral range (A) and feature selected wavebands (B).

Figure 8.3A and B illustrate the SVM-DA classification results of 200 iterations for the 5 class model built based on the whole spectral range and selected wavebands features, respectively. The majority of the SVM-DA iterations results show classification accuracy of ~ 99 %, averaged classification accuracy of 97.93 % and 98.97 % for before and after feature selection, respectively. The same analysis was performed with the 13 class model and the results are illustrated in figure 8.4A and B. The averaged classification accuracy before and after the feature selection was 74.37 % and 82.19 %, respectively. The classification accuracy results show the complexity of the 13 class model in comparison to the 5 class model. It is interesting to note the range of classification accuracy before the feature selection method was from 60 % to 85 %, however, after feature selection, this range shifted upward and showed classification accuracies in the range of 70 % and 95 %. In spite of the complexity of the model, about 10 iterations provide above 90 % classification accuracy with feature selection which are not seen before feature selection.

8.5 Conclusions

We demonstrated an excellent spot-to-spot and sample-to-sample reproducibility of SERS spectra offered by the LbL method and recently evaluated nanoparticle encapsulated mycoplasma cells as a platform to differentiate *M. pneumoniae* strains using SERS. By means of SERS in conjunction with multivariate statistical analysis, pathogenic strain M129 and 11 other human commensal strains were able to be accurately classified and identified.

Feature selection was performed by using ANOVA along with post-hoc Tukey test to identify the spectral features that are significant to best discriminate between classes.

These selected spectral features match closely with protein, nucleic acid, lipids, and carbohydrates relevant peaks reported in the literature. PLS-DA and SVM-DA were performed based on the feature selected wavebands and the overall results show a high classification accuracy by having ~ 95 % classification accuracy for the 5 class model and ~85 % for 13 class model, respectively.

In spite of the complexity of the biologically similar mycoplasma strains, this study demonstrated the feasibility of LbL based SERS technique combined with multivariate statistical analysis tool for a rapid, accurate and direct identification of whole bacterial cells as a diagnosis tool without amplification or labeling.

Acknowledgements

This research project was supported by the U.S. Public Health Service through grant GM102546 from the National Institutes of Health.

References

1 McMahon, M. D., Rush, J. S. & Thomas, M. G. Analyses of MbtB, MbtE, and MbtF Suggest Revisions to the Mycobactin Biosynthesis Pathway in Mycobacterium tuberculosis. *J Bacteriol* **194**, 2809-2818, doi:Doi 10.1128/Jb.00088-12 (2012).

2 Stover, C. K. *et al.* A small-molecule nitroimidazopyran drug candidate for the treatment of tuberculosis. *Nature* **405**, 962-966, doi:Doi 10.1038/35016103 (2000).

3 Kalinowski, D. S. & Richardson, D. R. The evolution of iron chelators for the treatment of iron overload disease and cancer. *Pharmacol Rev* **57**, 547-583, doi:Doi 10.1124/Pr.57.4.2 (2005).

4 Ratledge, C. Iron, mycobacteria and tuberculosis. *Tuberculosis* **84**, 110-130, doi:DOI 10.1016/j.tube.2003.08.012 (2004).

5 De Voss, J. J., Rutter, K., Schroeder, B. G. & Barry, C. E. Iron acquisition and metabolism by mycobacteria. *J Bacteriol* **181**, 4443-4451 (1999).

6 Hider, R. C. & Kong, X. L. Chemistry and biology of siderophores. *Nat Prod Rep* **27**, 637-657, doi:Doi 10.1039/B906679a (2010).

7 Barclay, R., Furst, V. & Smith, I. A Simple and Rapid Method for the Detection and Identification of Mycobacteria Using Mycobactin. *J Med Microbiol* **37**, 286-290 (1992).

8 Barclay, R., Ewing, D. F. & Ratledge, C. Isolation, Identification, and Structural-Analysis of the Mycobactins of Mycobacterium-Avium, Mycobacterium-Intracellulare, Mycobacterium-Scrofulaceum, and Mycobacterium-Paratuberculosis. *J Bacteriol* **164**, 896-903 (1985).

9 Eckstein, T. M. *et al.* A major cell wall lipopeptide of Mycobacterium avium subspecies paratuberculosis. *J Biol Chem* **281**, 5209-5215, doi:DOI 10.1074/jbc.M512465200 (2006).

10 Bosne, S. & Levyfrebault, V. V. Mycobactin Analysis as an Aid for the Identification of Mycobacterium-Fortuitum and Mycobacterium-Chelonae Subspecies. *J Clin Microbiol* **30**, 1225-1231 (1992).

11 Greatban.D & Bedford, G. R. Identification of Mycobactins by Nuclear-Magnetic-Resonance Spectroscopy. *Biochem J* **115**, 1047-& (1969).

12 Madigan, C. A. *et al.* Lipidomic discovery of deoxysiderophores reveals a revised mycobactin biosynthesis pathway in Mycobacterium tuberculosis. *P Natl Acad Sci USA* **109**, 1257-1262, doi:DOI 10.1073/pnas.1109958109 (2012).

13 Boukhalfa, H. & Crumbliss, A. L. Chemical aspects of siderophore mediated iron transport. *Biometals* **15**, 325-339, doi:Doi 10.1023/A:1020218608266 (2002).

14 Luo, M. K., Fadeev, E. A. & Groves, J. T. Mycobactin-mediated iron acquisition within macrophages. *Nat Chem Biol* **1**, 149-153, doi:Doi 10.1038/Nchembio717 (2005).

15 Siegrist, M. S. *et al.* Mycobacterial Esx-3 is required for mycobactin-mediated iron acquisition. *P Natl Acad Sci USA* **106**, 18792-18797, doi:DOI 10.1073/pnas.0900589106 (2009).

16 Flo, T. H. *et al.* Lipocalin 2 mediates an innate immune response to bacterial infection by sequestrating iron. *Nature* **432**, 917-921, doi:Doi 10.1038/Nature03104 (2004).

17 Minnikin, D. E., Kremer, L., Dover, L. G. & Besra, G. S. The methyl-branched fortifications of Mycobacterium tuberculosis. *Chem Biol* **9**, 545-553, doi:Pii S1074-5521(02)00142-4
Doi 10.1016/S1074-5521(02)00142-4 (2002).

18 Laval, F., Laneelle, M. A., Deon, C., Monsarrat, B. & Daffe, M. Accurate molecular mass determination of mycolic acids by MALDI-TOF mass spectrometry. *Anal Chem* **73**, 4537-4544, doi:Doi 10.1021/Ac0105181 (2001).

19 Butler, W. R. & Guthertz, L. S. Mycolic acid analysis by high-performance liquid chromatography for identification of Mycobacterium species. *Clin Microbiol Rev* **14**, 704-726, doi:Doi 10.1128/Cmr.14.4.704-726.2001 (2001).

20 Fujita, Y., Naka, T., McNeil, M. R. & Yano, I. Intact molecular characterization of cord factor (trehalose 6,6 '-dimycolate) from nine species of mycobacteria by MALDI-TOF mass spectrometry. *Microbiol-Sgm* **151**, 3403-3416, doi:DOI 10.1099/mic.0.28158-0 (2005).

21 Liu, J. & Nikaido, H. A mutant of Mycobacterium smegmatis defective in the biosynthesis of mycolic acids accumulates meromycolates. *P Natl Acad Sci USA* **96**, 4011-4016, doi:DOI 10.1073/pnas.96.7.4011 (1999).

22 Chatterjee, D. The mycobacterial cell wall: structure, biosynthesis and sites of drug action. *Curr Opin Chem Biol* **1**, 579-588, doi:Doi 10.1016/S1367-5931(97)80055-5 (1997).

23 Kawamura, M. *et al.* Enzyme immunoassay to detect antituberculous glycolipid antigen (anti-TBGL antigen) antibodies in serum for diagnosis of tuberculosis. *J Clin Lab Anal* **11**, 140-145, doi:Doi 10.1002/(Sici)1098-2825(1997)11:3<140::Aid-Jcla4>3.0.Co;2-E (1997).

24 Hasegawa, T. & Leblanc, R. M. Aggregation properties of mycolic acid molecules in monolayer films: a comparative study of compounds from various acid-fast bacterial species. *Bba-Biomembranes* **1617**, 89-95, doi:DOI 10.1016/j.bbamem.2003.09.008 (2003).

25 Williams, C. Global Tuberculosis Control: WHO Report 2011. *Aust Nz J Publ Heal* **36**, 497-498, doi:DOI 10.1111/j.1753-6405.2012.00928.x (2012).

26 Lawn, S. D., Kerkhoff, A. D., Vogt, M. & Wood, R. Diagnostic accuracy of a low-cost, urine antigen, point-of-care screening assay for HIV-associated pulmonary tuberculosis before antiretroviral therapy: a descriptive study. *The Lancet. Infectious diseases* **12**, 201-209, doi:10.1016/S1473-3099(11)70251-1 (2012).

27 Steingart, K. R. *et al.* Sputum processing methods to improve the sensitivity of smear microscopy for tuberculosis: a systematic review. *The Lancet. Infectious diseases* **6**, 664-674, doi:10.1016/S1473-3099(06)70602-8 (2006).

28 Wallis, R. S. *et al.* Tuberculosis 4 Biomarkers and diagnostics for tuberculosis: progress, needs, and translation into practice. *Lancet* **375**, 1920-1937, doi:Doi 10.1016/S0140-6736(10)60359-5 (2010).

29 Kassa-Kelembho, E. *et al.* Poor performance of a novel serological test for diagnosis of pulmonary tuberculosis in Bangui, Central African Republic. *Clin Vaccine Immunol* **13**, 702-703, doi:Doi 10.1128/Cdli.00194-05 (2006).

30 Veenstra, H. *et al.* Changes in the kinetics of intracellular IFN-gamma production in TB patients during treatment. *Clin Immunol* **124**, 336-344, doi:DOI 10.1016/j.clim.2007.05.014 (2007).

31 Boehme, C. C. *et al.* Rapid Molecular Detection of Tuberculosis and Rifampin Resistance. *New Engl J Med* **363**, 1005-1015, doi:Doi 10.1056/Nejmoa0907847 (2010).
32 Boehme, C. C. *et al.* Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study. *Lancet* **377**, 1495-1505, doi:Doi 10.1016/S0140-6736(11)60438-8 (2011).
33 Evans, C. A. GeneXpert-A Game-Changer for Tuberculosis Control? *Plos Med* **8**, doi:ARTN e1001064
DOI 10.1371/journal.pmed.1001064 (2011).
34 Waites, K. B. & Talkington, D. F. Mycoplasma pneumoniae and its role as a human pathogen. *Clin Microbiol Rev* **17**, 697-+, doi:Doi 10.1128/Cmr.17.4.697-728.2004 (2004).
35 Romero-Arroyo, C. E. *et al.* Mycoplasma pneumoniae protein P30 is required for cytadherence and associated with proper cell development. *J Bacteriol* **181**, 1079-1087 (1999).
36 Dallo, S. F. & Baseman, J. B. Intracellular DNA replication and long-term survival of pathogenic mycoplasmas. *Microb Pathogenesis* **29**, 301-309, doi:DOI 10.1006/mpat.2000.0395 (2000).
37 Rivera-Betancourt, O. E., Sheppard, E. S., Krause, D. C. & Dluhy, R. A. Layer-by-layer polyelectrolyte encapsulation of Mycoplasma pneumoniae for enhanced Raman detection. *Analyst* **139**, 4287-4295, doi:Doi 10.1039/C4an00596a (2014).
38 Chanock, R. M., Hayflick, L. & Barile, M. F. Growth on Artificial Medium of an Agent Associated with Atypical Pneumonia and Its Identification as a Pplo. *P Natl Acad Sci USA* **48**, 41-&, doi:Doi 10.1073/Pnas.48.1.41 (1962).
39 Daxboeck, F., Krause, R. & Wenisch, C. Laboratory diagnosis of Mycoplasma pneumoniae infection. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* **9**, 263-273 (2003).
40 Kung, C. M. & Wang, H. L. Seroprevalence of mycoplasma pneumoniae in healthy adolescents in Taiwan. *Jpn J Infect Dis* **60**, 352-354 (2007).
41 Loens, K., Goossens, H. & Ieven, M. Acute respiratory infection due to Mycoplasma pneumoniae: current status of diagnostic methods. *Eur J Clin Microbiol* **29**, 1055-1069, doi:DOI 10.1007/s10096-010-0975-2 (2010).
42 Thibodeau, K. P. & Viera, A. J. Atypical pathogens and challenges in community-acquired pneumonia. *Am Fam Physician* **69**, 1699-1706 (2004).
43 Winchell, J. M., Thurman, K. A., Mitchell, S. L., Thacker, W. L. & Fields, B. S. Evaluation of three real-time PCR assays for detection of Mycoplasma pneumoniae in an outbreak investigation. *J Clin Microbiol* **46**, 3116-3118, doi:Doi 10.1128/Jcm.00440-08 (2008).
44 Dumke, R., Schurwanz, N. & Jacobs, E. Characterisation of subtype- and variant-specific antigen regions of the P1 adhesin of Mycoplasma pneumoniae. *Int J Med Microbiol* **298**, 483-491, doi:DOI 10.1016/j.ijmm.2007.06.002 (2008).
45 Jacobs, E. Mycoplasma pneumoniae: now in the focus of clinicians and epidemiologists. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* **17** (2012).

46 Schwartz, S. B., Mitchell, S. L., Thurman, K. A., Wolff, B. J. & Winchell, J. M. Identification of P1 Variants of Mycoplasma pneumoniae by Use of High-Resolution Melt Analysis. *J Clin Microbiol* **47**, 4117-4120, doi:Doi 10.1128/Jcm.01696-09 (2009).
47 Gerstenecker, B. & Jacobs, E. Topological Mapping of the P1-Adhesin of Mycoplasma-Pneumoniae with Adherence-Inhibiting Monoclonal-Antibodies. *J Gen Microbiol* **136**, 471-476 (1990).
48 Jacobs, E., Pilatschek, A., Gerstenecker, B., Oberle, K. & Bredt, W. Immunodominant Epitopes of the Adhesin of Mycoplasma-Pneumoniae. *J Clin Microbiol* **28**, 1194-1197 (1990).
49 Thurman, K. A. *et al.* Comparison of Laboratory Diagnostic Procedures for Detection of Mycoplasma pneumoniae in Community Outbreaks. *Clin Infect Dis* **48**, 1244-1249, doi:Doi 10.1086/597775 (2009).
50 Dumke, R. *et al.* Culture-independent molecular subtyping of Mycoplasma pneumoniae in clinical samples. *J Clin Microbiol* **44**, 2567-2570, doi:Doi 10.1128/Jcm.00495-06 (2006).
51 Schwartz, S. B., Thurman, K. A., Mitchell, S. L., Wolff, B. J. & Winchell, J. M. Genotyping of Mycoplasma pneumoniae isolates using real-time PCR and high-resolution melt analysis. *Clin Microbiol Infec* **15**, 756-762, doi:DOI 10.1111/j.1469-0691.2009.02814.x (2009).
52 Maquelin, K. *et al.* Raman spectroscopic typing reveals the presence of carotenoids in Mycoplasma pneumoniae. *Microbiology* **155**, 2068-2077, doi:10.1099/mic.0.026724-0 (2009).
53 Kelly, K. L., Coronado, E., Zhao, L. L. & Schatz, G. C. The optical properties of metal nanoparticles: The influence of size, shape, and dielectric environment. *J Phys Chem B* **107**, 668-677, doi:Doi 10.1021/Jp026731y (2003).
54 Raman, C. V. & Krishnan, K. S. A new type of secondary radiation (Reprinted from Nature, vol 121, pg 501-502, 1928). *Curr Sci India* **74**, 381-381 (1998).
55 Aroca, R. & RodriguezLlorente, S. Surface-enhanced vibrational spectroscopy. *J Mol Struct* **408**, 17-22, doi:Doi 10.1016/S0022-2860(96)09489-6 (1997).
56 Kudelski, A. Analytical applications of Raman spectroscopy. *Talanta* **76**, 1-8, doi:DOI 10.1016/j.talanta.2008.02.042 (2008).
57 Kerker, M. From the founding fathers of light scattering to surface enhanced Raman scattering. *Abstr Pap Am Chem S* **221**, U312-U313 (2001).
58 Chu, P. K. & Li, L. H. Characterization of amorphous and nanocrystalline carbon films. *Mater Chem Phys* **96**, 253-277, doi:DOI 10.1016/j.matchemphys.2005.07.048 (2006).
59 Maquelin, K. *et al.* Raman spectroscopic method for identification of clinically relevant microorganisms growing on solid culture medium. *Anal Chem* **72**, 12-19, doi:Doi 10.1021/Ac991011h (2000).
60 Premasiri, W. R. *et al.* Characterization of the Surface Enhanced Raman Scattering (SERS) of bacteria. *J Phys Chem B* **109**, 312-320, doi:Doi 10.1021/Jp040442n (2005).
61 Lombardi, J. R. & Birke, R. L. The theory of surface-enhanced Raman scattering. *The Journal of chemical physics* **136**, 144704, doi:10.1063/1.3698292 (2012).

62 Lombardi, J. R. & Birke, R. L. Time-dependent picture of the charge-transfer contributions to surface enhanced Raman spectroscopy. *The Journal of chemical physics* **126**, 244709, doi:10.1063/1.2748386 (2007).
63 Ren, B., Liu, G. K., Lian, X. B., Yang, Z. L. & Tian, Z. Q. Raman spectroscopy on transition metals. *Anal Bioanal Chem* **388**, 29-45, doi:DOI 10.1007/s00216-007-1141-2 (2007).
64 Huang, Q. J. *et al.* Extending surface Raman spectroscopic studies to transition metals for practical applications III. Effects of surface roughening procedure on surface-enhanced Raman spectroscopy from nickel and platinum electrodes. *Surf Sci* **427-28**, 162-166, doi:Doi 10.1016/S0039-6028(99)00258-7 (1999).
65 Yao, J. L. *et al.* Extending surface Raman spectroscopy to transition metals for practical applications IV. A study on corrosion inhibition of benzotriazole on bare Fe electrodes. *Electrochim Acta* **48**, 1263-1271, doi:Doi 10.1016/S0013-4686(02)00834-4 (2003).
66 Wiley, B. J. *et al.* Maneuvering the surface plasmon resonance of silver nanostructures through shape-controlled synthesis. *J Phys Chem B* **110**, 15666-15675, doi:Doi 10.1021/Jp0608628 (2006).
67 Campion, A. & Kambhampati, P. Surface-enhanced Raman scattering. *Chem Soc Rev* **27**, 241-250, doi:Doi 10.1039/A827241z (1998).
68 Albrecht, M. G. & Creighton, J. A. Anomalously Intense Raman-Spectra of Pyridine at a Silver Electrode. *J Am Chem Soc* **99**, 5215-5217, doi:Doi 10.1021/Ja00457a071 (1977).
69 Jeanmaire, D. L. & Vanduyne, R. P. Surface Raman Spectroelectrochemistry .1. Heterocyclic, Aromatic, and Aliphatic-Amines Adsorbed on Anodized Silver Electrode. *J Electroanal Chem* **84**, 1-20, doi:Doi 10.1016/S0022-0728(77)80224-6 (1977).
70 Stiles, P. L., Dieringer, J. A., Shah, N. C. & Van Duyne, R. R. Surface-Enhanced Raman Spectroscopy. *Annu Rev Anal Chem* **1**, 601-626, doi:DOI 10.1146/annurev.anchem.1.031207.112814 (2008).
71 Zhang, Y. X., Aslan, K., Previte, M. J. R. & Geddes, C. D. Metal-enhanced fluorescence: Surface plasmons can radiate a fluorophore's structured emission. *Appl Phys Lett* **90**, doi:Artn 053107
Doi 10.1063/1.2435661 (2007).
72 Kneipp, K. *et al.* Single molecule detection using surface-enhanced Raman scattering (SERS). *Phys Rev Lett* **78**, 1667-1670, doi:DOI 10.1103/PhysRevLett.78.1667 (1997).
73 Nie, S. M. & Emery, S. R. Probing single molecules and single nanoparticles by surface-enhanced Raman scattering. *Science* **275**, 1102-1106, doi:DOI 10.1126/science.275.5303.1102 (1997).
74 Willets, K. A. & Van Duyne, R. P. Localized surface plasmon resonance spectroscopy and sensing. *Annu Rev Phys Chem* **58**, 267-297, doi:DOI 10.1146/annurev.physchem.58.032806.104607 (2007).
75 Rodriguez-Gonzalez, B., Burrows, A., Watanabe, M., Kiely, C. J. & Liz-Marzan, L. M. Multishell bimetallic AuAg nanoparticles: synthesis, structure and optical properties. *J Mater Chem* **15**, 1755-1759, doi:Doi 10.1039/B500556f (2005).

76 Miller, M. M. & Lazarides, A. A. Sensitivity of metal nanoparticle surface plasmon resonance to the dielectric environment. *J Phys Chem B* **109**, 21556-21565, doi:Doi 10.1021/Jp054227y (2005).
77 Burda, C., Chen, X. B., Narayanan, R. & El-Sayed, M. A. Chemistry and properties of nanocrystals of different shapes. *Chem Rev* **105**, 1025-1102, doi:Doi 10.1021/Cr030063a (2005).
78 Li, H. X. & Rothberg, L. Colorimetric detection of DNA sequences based on electrostatic interactions with unmodified gold nanoparticles. *P Natl Acad Sci USA* **101**, 14036-14039, doi:DOI 10.1073/pnas.0406115101 (2004).
79 Rosi, N. L. & Mirkin, C. A. Nanostructures in biodiagnostics. *Chem Rev* **105**, 1547-1562, doi:Doi 10.1021/Cr030067f (2005).
80 Zhou, X. Z., Boey, F., Huo, F. W., Huang, L. & Zhang, H. Chemically Functionalized Surface Patterning. *Small* **7**, 2273-2289, doi:DOI 10.1002/smll.201002381 (2011).
81 Xie, Z. H. *et al.* Plasmonic Nanolithography: A Review. *Plasmonics* **6**, 565-580, doi:DOI 10.1007/s11468-011-9237-0 (2011).
82 Freeman, R. G. *et al.* Self-Assembled Metal Colloid Monolayers - an Approach to Sers Substrates. *Science* **267**, 1629-1632, doi:DOI 10.1126/science.267.5204.1629 (1995).
83 Liu, Y. J., Chu, H. Y. & Zhao, Y. P. Silver Nanorod Array Substrates Fabricated by Oblique Angle Deposition: Morphological, Optical, and SERS Characterizations. *J Phys Chem C* **114**, 8176-8183, doi:Doi 10.1021/Jp1001644 (2010).
84 Robbie, K., Sit, J. C. & Brett, M. J. Advanced techniques for glancing angle deposition. *J Vac Sci Technol B* **16**, 1115-1122, doi:Doi 10.1116/1.590019 (1998).
85 Abelmann, L. & Lodder, C. Oblique evaporation and surface diffusion. *Thin Solid Films* **305**, 1-21, doi:Doi 10.1016/S0040-6090(97)00095-3 (1997).
86 Vankranenburg, H. & Lodder, C. Tailoring Growth and Local Composition by Oblique-Incidence Deposition - a Review and New Experimental-Data. *Mat Sci Eng R* **11**, 295-354, doi:Doi 10.1016/0927-796x(94)90021-3 (1994).
87 Beavers, K. R., Marotta, N. E. & Bottomley, L. A. Thermal Stability of Silver Nanorod Arrays. *Chem Mater* **22**, 2184-2189, doi:Doi 10.1021/Cm901791u (2010).
88 Chaney, S. B., Shanmukh, S., Dluhy, R. A. & Zhao, Y. P. Aligned silver nanorod arrays produce high sensitivity surface-enhanced Raman spectroscopy substrates. *Appl Phys Lett* **87**, doi:Artn 031908
Doi 10.1063/1.1988980 (2005).
89 Chu, H. Y., Huang, Y. W. & Zhao, Y. P. Silver nanorod arrays as a surface-enhanced Raman scattering substrate for foodborne pathogenic bacteria detection. *Applied spectroscopy* **62**, 922-931 (2008).
90 Driskell, J. D. *et al.* Infectious agent detection with SERS-active silver nanorod arrays prepared by oblique angle deposition. *Ieee Sens J* **8**, 863-870, doi:Doi 10.1109/Jsen.2008.922682 (2008).
91 Driskell, J. D. & Tripp, R. A. Label-free SERS detection of microRNA based on affinity for an unmodified silver nanorod array substrate. *Chem Commun* **46**, 3298-3300, doi:Doi 10.1039/C002059a (2010).

92 Rivera-Betancourt, O. E. *et al.* Identification of mycobacteria based on spectroscopic analyses of mycolic acid profiles. *The Analyst* **138**, 6774-6785, doi:Doi 10.1039/C3an01157g (2013).
93 Liu, Y. J., Zhang, Z. Y., Dluhy, R. A. & Zhao, Y. P. The SERS response of semiordered Ag nanorod arrays fabricated by template oblique angle deposition. *J Raman Spectrosc* **41**, 1112-1118, doi:Doi 10.1002/Jrs.2567 (2010).
94 Negri, P., Kage, A., Nitsche, A., Naumann, D. & Dluhy, R. A. Detection of viral nucleoprotein binding to anti-influenza aptamers via SERS. *Chem Commun* **47**, 8635-8637, doi:Doi 10.1039/C0cc05433j (2011).
95 Driskell, J. D. *et al.* The use of aligned silver nanorod Arrays prepared by oblique angle deposition as surface enhanced Raman scattering substrates. *J Phys Chem C* **112**, 895-901, doi:Doi 10.1021/Jp075288u (2008).
96 He, Y. Z., Fu, J. X. & Zhao, Y. P. Oblique angle deposition and its applications in plasmonics. *Front Phys-Beijing* **9**, 47-59, doi:DOI 10.1007/s11467-013-0357-1 (2014).
97 Leverette, C. L. *et al.* Aligned silver nanorod arrays as substrates for surface-enhanced infrared absorption spectroscopy. *Appl Spectrosc* **60**, 906-913, doi:Doi 10.1366/000370206778062084 (2006).
98 Zasadzinski, J. A., Viswanathan, R., Madsen, L., Garnaes, J. & Schwartz, D. K. Langmuir-Blodgett-Films. *Science* **263**, 1726-1733, doi:DOI 10.1126/science.8134836 (1994).
99 Iler, R. K. Multilayers of Colloidal Particles. *J Colloid Interf Sci* **21**, 569-&, doi:Doi 10.1016/0095-8522(66)90018-3 (1966).
100 Nicolau, Y. F. & Menard, J. C. Solution Growth of Zns, Cds and Zn1-Xcdxs Thin-Films by the Successive Ionic-Layer Adsorption and Reaction Process - Growth-Mechanism. *J Cryst Growth* **92**, 128-142, doi:Doi 10.1016/0022-0248(88)90443-5 (1988).
101 Nicolau, Y. F., Davied, S., Genoud, F., Nechtschein, M. & Travers, J. P. Polyaniline, Polypyrrole, Poly(3-Methylthiophene) and Polybithiophene Layer-by-Layer Deposited Thin-Films. *Synthetic Met* **42**, 1491-1494, doi:Doi 10.1016/0379-6779(91)91884-D (1991).
102 Stockton, W. B. & Rubner, M. F. Molecular-level processing of conjugated polymers .4. Layer-by-layer manipulation of polyaniline via hydrogen-bonding interactions. *Macromolecules* **30**, 2717-2725, doi:Doi 10.1021/Ma9700486 (1997).
103 Wang, L. Y. *et al.* A new approach for the fabrication of an alternating multilayer film of poly(4-vinylpyridine) and poly(acrylic acid) based on hydrogen bonding. *Macromol Rapid Comm* **18**, 509-514, doi:DOI 10.1002/marc.1997.030180609 (1997).
104 Hao, E. C. & Lian, T. Q. Buildup of polymer/Au nanoparticle multilayer thin films based on hydrogen bonding. *Chem Mater* **12**, 3392-3396, doi:Doi 10.1021/Cm000565u (2000).
105 Schonhoff, M. Self-assembled polyelectrolyte multilayers. *Curr Opin Colloid In* **8**, 86-95, doi:Doi 10.1016/S1359-0294(03)00003-7 (2003).
106 Sukhorukov, G. B. & Mohwald, H. Multifunctional cargo systems for biotechnology. *Trends Biotechnol* **25**, 93-98, doi:DOI 10.1016/j.tibtech.2006.12.007 (2007).

107 Loh, K. J., Kim, J., Lynch, J. P., Kam, N. W. S. & Kotov, N. A. Multifunctional layer-by-layer carbon nanotube-polyelectrolyte thin films for strain and corrosion sensing. *Smart Mater Struct* **16**, 429-438, doi:Doi 10.1088/0964-1726/16/2/022 (2007).

108 Dai, Z. F., Voigt, A., Donath, E. & Mohwald, H. Novel encapsulated functional dye particles based on alternately adsorbed multilayers of active oppositely charged macromolecular species. *Macromol Rapid Comm* **22**, 756-762, doi:Doi 10.1002/1521-3927(20010701)22:10<756::Aid-Marc756>3.0.Co;2-D (2001).

109 Gobi, K. V. & Mizutani, F. Layer-by-layer construction of an active multilayer enzyme electrode applicable for direct amperometric determination of cholesterol. *Sensor Actuat B-Chem* **80**, 272-277 (2001).

110 Rao, S. V., Anderson, K. W. & Bachas, L. G. Controlled layer-by-layer immobilization of horseradish peroxidase. *Biotechnol Bioeng* **65**, 389-396, doi:Doi 10.1002/(Sici)1097-0290(19991120)65:4<389::Aid-Bit3>3.0.Co;2-V (1999).

111 Wood, K. C., Chuang, H. F., Batten, R. D., Lynn, D. M. & Hammond, P. T. Controlling interlayer diffusion to achieve sustained, multiagent delivery from layer-by-layer thin films. *P Natl Acad Sci USA* **103**, 10207-10212, doi:DOI 10.1073/pnas.0602884103 (2006).

112 Veerabadran, N. G., Goli, P. L., Stewart-Clark, S. S., Lvov, Y. M. & Mills, D. K. Nanoencapsulation of stem cells within polyelectrolyte multilayer shells. *Macromol Biosci* **7**, 877-882, doi:DOI 10.1002/mabi.200700061 (2007).

113 Zhou, J. *et al.* Layer by layer chitosan/alginate coatings on poly(lactide-co-glycolide) nanoparticles for antifouling protection and Folic acid binding to achieve selective cell targeting. *J Colloid Interf Sci* **345**, 241-247, doi:DOI 10.1016/j.jcis.2010.02.004 (2010).

114 Angelatos, A. S., Radt, B. & Caruso, F. Light-responsive polyelectrolyte/gold nanoparticle microcapsules. *J Phys Chem B* **109**, 3071-3076, doi:Doi 10.1021/Jp045070x (2005).

115 Chen, W. & McCarthy, T. J. Layer-by-layer deposition: A tool for polymer surface modification. *Macromolecules* **30**, 78-86, doi:Doi 10.1021/Ma961096d (1997).

116 Clark, S. L. & Hammond, P. T. The role of secondary interactions in selective electrostatic multilayer deposition. *Langmuir* **16**, 10206-10214, doi:Doi 10.1021/La000418a (2000).

117 Fu, Y. *et al.* Hydrogen-bonding-directed layer-by-layer multilayer assembly: Reconformation yielding microporous films. *Macromolecules* **35**, 9451-9458, doi:Doi 10.1021/Ma0207881 (2002).

118 Kotov, N. A. Layer-by-layer self-assembly: The contribution of hydrophobic interactions. *Nanostruct Mater* **12**, 789-796, doi:Doi 10.1016/S0965-9773(99)00237-8 (1999).

119 Sano, M., Kamino, A., Okamura, J. & Shinkai, S. Noncovalent self-assembly of carbon nanotubes for construction of "cages". *Nano Lett* **2**, 531-533, doi:Doi 10.1021/Nl025525z (2002).

120 Bantchev, G., Lu, Z. H. & Lvov, Y. Layer-by-Layer Nanoshell Assembly on Colloids Through Simplified Washless Process. *J Nanosci Nanotechno* **9**, 396-403, doi:DOI 10.1166/jnn.2009.J055 (2009).

121 Decher, G. Fuzzy nanoassemblies: Toward layered polymeric multicomposites. *Science* **277**, 1232-1237, doi:DOI 10.1126/science.277.5330.1232 (1997).
122 Ochs, C. J., Such, G. K., Yan, Y., van Koeverden, M. P. & Caruso, F. Biodegradable Click Capsules with Engineered Drug-Loaded Multilayers. *Acs Nano* **4**, 1653-1663, doi:Doi 10.1021/Nn9014278 (2010).
123 Bertrand, P., Jonas, A., Laschewsky, A. & Legras, R. Ultrathin polymer coatings by complexation of polyelectrolytes at interfaces: suitable materials, structure and properties. *Macromol Rapid Comm* **21**, 319-348, doi:Doi 10.1002/(Sici)1521-3927(20000401)21:7<319::Aid-Marc319>3.0.Co;2-7 (2000).
124 Becker, A. L., Johnston, A. P. R. & Caruso, F. Peptide Nucleic Acid Films and Capsules: Assembly and Enzymatic Degradation. *Macromol Biosci* **10**, 488-495, doi:DOI 10.1002/mabi.200900347 (2010).
125 Johnston, A. P. R., Read, E. S. & Caruso, F. DNA multilayer films on planar and colloidal supports: Sequential assembly of like-charged polyelectrolytes. *Nano Lett* **5**, 953-956, doi:Doi 10.1021/Nl050608b (2005).
126 Dhamodharan, R. & McCarthy, T. J. Adsorption of alginic acid and chondroitin sulfate-A to amine functionality introduced on polychlorotrifluoroethylene and glass surfaces. *Macromolecules* **32**, 4106-4112, doi:Doi 10.1021/Ma980991p (1999).
127 Decher, G., Lehr, B., Lowack, K., Lvov, Y. & Schmitt, J. New Nanocomposite Films for Biosensors - Layer-by-Layer Adsorbed Films of Polyelectrolytes, Proteins or DNA. *Biosens Bioelectron* **9**, 677-684, doi:Doi 10.1016/0956-5663(94)80065-0 (1994).
128 Caruso, F., Yang, W. J., Trau, D. & Renneberg, R. Microencapsulation of uncharged low molecular weight organic materials by polyelectrolyte multilayer self-assembly. *Langmuir* **16**, 8932-8936, doi:Doi 10.1021/La000401s (2000).
129 Decher, G., Hong, J. D. & Schmitt, J. Buildup of Ultrathin Multilayer Films by a Self-Assembly Process .3. Consecutively Alternating Adsorption of Anionic and Cationic Polyelectrolytes on Charged Surfaces. *Thin Solid Films* **210**, 831-835, doi:Doi 10.1016/0040-6090(92)90417-A (1992).
130 Ferreira, M. & Rubner, M. F. Molecular-Level Processing of Conjugated Polymers .1. Layer-by-Layer Manipulation of Conjugated Polyions. *Macromolecules* **28**, 7107-7114, doi:Doi 10.1021/Ma00125a012 (1995).
131 Dubas, S. T. & Schlenoff, J. B. Factors controlling the growth of polyelectrolyte multilayers. *Abstr Pap Am Chem S* **219**, U532-U532 (2000).
132 McAloney, R. A., Sinyor, M., Dudnik, V. & Goh, M. C. Atomic force microscopy studies of salt effects on polyelectrolyte multilayer film morphology. *Langmuir* **17**, 6655-6663, doi:Doi 10.1021/La010136q (2001).
133 Schlenoff, J. B. & Dubas, S. T. Mechanism of polyelectrolyte multilayer growth: Charge overcompensation and distribution. *Macromolecules* **34**, 592-598, doi:Doi 10.1021/Ma0003093 (2001).
134 Dubas, S. T. & Schlenoff, J. B. Factors controlling the growth of polyelectrolyte multilayers. *Macromolecules* **32**, 8153-8160, doi:Doi 10.1021/Ma981927a (1999).
135 Harris, K. A. & Hartley, J. C. Development of broad-range 16S rDNA PCR for use in the routine diagnostic clinical microbiology service. *J Med Microbiol* **52**, 685-691, doi:DOI 10.1099/jmm.0.05213-0 (2003).

136 Corless, C. E. *et al.* Contamination and sensitivity issues with a real-time universal 16S rRNA PCR. *J Clin Microbiol* **38**, 1747-1752 (2000).
137 Deisingh, A. K. & Thompson, M. Detection of infectious and toxigenic bacteria. *The Analyst* **127**, 567-581, doi:Doi 10.1039/B109895k (2002).
138 Chao, R. K., Fishaut, M., Schwartzman, J. D. & Mcintosh, K. Detection of Respiratory Syncytial Virus in Nasal Secretions from Infants by Enzyme-Linked Immunosorbent Assay. *J Infect Dis* **139**, 483-486 (1979).
139 Freymuth, F. *et al.* Detection of Respiratory Syncytial Virus by Reverse Transcription Pcr and Hybridization with a DNA Enzyme-Immunoassay. *J Clin Microbiol* **33**, 3352-3355 (1995).
140 Henkel, J. H., Aberle, S. W., Kundi, M. & PopowKraupp, T. Improved detection of respiratory syncytial virus in nasal aspirates by seminested RT-PCR. *J Med Virol* **53**, 366-371 (1997).
141 Connor, E. M. & Loeb, M. R. A Hemadsorption Method for Detection of Colonies of Hemophilus-Influenzae Type-B Expressing Fimbriae. *J Infect Dis* **148**, 855-860 (1983).
142 Patel, I. S., Premasiri, W. R., Moir, D. T. & Ziegler, L. D. Barcoding bacterial cells: a SERS-based methodology for pathogen identification. *J Raman Spectrosc* **39**, 1660-1672, doi:Doi 10.1002/Jrs.2064 (2008).
143 Bao, P. D., Huang, T. Q., Liu, X. M. & Wu, T. Q. Surface-enhanced Raman spectroscopy of insect nuclear polyhedrosis virus. *J Raman Spectrosc* **32**, 227-230, doi:Doi 10.1002/Jrs.665.Abs (2001).
144 Camden, J. P., Dieringer, J. A., Zhao, J. & Van Duyne, R. P. Controlled plasmonic nanostructures for surface-enhanced spectroscopy and sensing. *Accounts of chemical research* **41**, 1653-1661, doi:10.1021/ar800041s (2008).
145 Jarvis, R. M., Brooker, A. & Goodacre, R. Surface-enhanced Raman scattering for the rapid discrimination of bacteria. *Faraday Discuss* **132**, 281-292, doi:Doi 10.1039/B506413a (2006).
146 Jarvis, R. M. *et al.* Surface-enhanced Raman scattering from intracellular and extracellular bacterial locations. *Anal Chem* **80**, 6741-6746, doi:Doi 10.1021/Ac800838v (2008).
147 Stewart, S. & Fredericks, P. M. Surface-enhanced Raman spectroscopy of peptides and proteins adsorbed on an electrochemically prepared silver surface. *Spectrochim Acta A* **55**, 1615-1640, doi:Doi 10.1016/S1386-1425(98)00293-5 (1999).
148 Stewart, S. & Fredericks, P. M. Surface-enhanced Raman spectroscopy of amino acids adsorbed on an electrochemically prepared silver surface. *Spectrochim Acta A* **55**, 1641-1660, doi:Doi 10.1016/S1386-1425(98)00294-7 (1999).
149 Wiley, J. H. & Atalla, R. H. Band Assignments in the Raman-Spectra of Celluloses. *Carbohyd Res* **160**, 113-129, doi:Doi 10.1016/0008-6215(87)80306-3 (1987).
150 Hennigan, S. L. *et al.* Detection and Differentiation of Avian Mycoplasmas by Surface-Enhanced Raman Spectroscopy Based on a Silver Nanorod Array. *Appl Environ Microb* **78**, 1930-1935, doi:Doi 10.1128/Aem.07419-11 (2012).
151 Hennigan, S. L. *et al.* Detection of Mycoplasma pneumoniae in Simulated and True Clinical Throat Swab Specimens by Nanorod Array-Surface-Enhanced Raman Spectroscopy. *Plos One* **5**, doi:ARTN e13633
DOI 10.1371/journal.pone.0013633 (2010).

152 Negri, P., Marotta, N. E., Bottomley, L. A. & Dluhy, R. A. Removal of Surface Contamination and Self-Assembled Monolayers (SAMs) from Silver (Ag) Nanorod Substrates by Plasma Cleaning with Argon. *Appl Spectrosc* **65**, 66-74, doi:Doi 10.1366/10-06037 (2011).
153 Zeiri, L., Bronk, B. V., Shabtai, Y., Czege, J. & Efrima, S. Silver metal induced surface enhanced Raman of bacteria. *Colloid Surface A* **208**, 357-362, doi:Pii S0927-7757(02)00162-0
Doi 10.1016/S0927-7757(02)00162-0 (2002).
154 Zeiri, L., Bronk, B. V., Shabtai, Y., Eichler, J. & Efrima, S. Surface-enhanced Raman spectroscopy as a tool for probing specific biochemical components in bacteria. *Appl Spectrosc* **58**, 33-40, doi:Doi 10.1366/000370204322729441 (2004).
155 Culha, M. *et al.* Characterization of Thermophilic Bacteria Using Surface-Enhanced Raman Scattering. *Appl Spectrosc* **62**, 1226-1232 (2008).
156 Naja, G., Bouvrette, P., Hrapovic, S. & Luong, J. H. T. Raman-based detection of bacteria using silver nanoparticles conjugated with antibodies. *The Analyst* **132**, 679-686, doi:Doi 10.1039/B701160a (2007).
157 Sztainbuch, I. W. The effects of Au aggregate morphology on surface-enhanced Raman scattering enhancement. *Journal of Chemical Physics* **125**, doi:Artn 124707
Doi 10.1063/1.2338029 (2006).
158 Kahraman, M., Keseroglu, K. & Culha, M. On Sample Preparation for Surface-Enhanced Raman Scattering (SERS) of Bacteria and the Source of Spectral Features of the Spectra. *Appl Spectrosc* **65**, 500-506, doi:Doi 10.1366/10-06184 (2011).
159 Kahraman, M., Yazici, M. M., Sahin, F., Bayrak, O. F. & Culha, M. Reproducible surface-enhanced Raman scattering spectra of bacteria on aggregated silver nanoparticles. *Appl Spectrosc* **61**, 479-485, doi:Doi 10.1366/000370207780807731 (2007).
160 Kahraman, M., Yazici, M. M., Sahin, F. & Culha, M. Convective assembly of bacteria for surface-enhanced Raman scattering. *Langmuir* **24**, 894-901, doi:Doi 10.1021/La702240q (2008).
161 Kahraman, M., Zamaleeva, A. I., Fakhrullin, R. F. & Culha, M. Layer-by-layer coating of bacteria with noble metal nanoparticles for surface-enhanced Raman scattering. *Anal Bioanal Chem* **395**, 2559-2567, doi:DOI 10.1007/s00216-009-3159-0 (2009).
162 Berry, V., Gole, A., Kundu, S., Murphy, C. J. & Saraf, R. F. Deposition of CTAB-terminated nanorods on bacteria to form highly conducting hybrid systems. *J Am Chem Soc* **127**, 17600-17601, doi:Doi 10.1021/Ja056428l (2005).
163 Wilson, W. W., Wade, M. M., Holman, S. C. & Champlin, F. R. Status of methods for assessing bacterial cell surface charge properties based on zeta potential measurements. *J Microbiol Meth* **43**, 153-164, doi:Doi 10.1016/S0167-7012(00)00224-4 (2001).
164 Berry, V. & Saraf, R. F. Self-assembly of nanoparticles on live bacterium: An avenue to fabricate electronic devices. *Angew Chem Int Edit* **44**, 6668-6673, doi:DOI 10.1002/anie.200501711 (2005).
165 Zhou, H. B. *et al.* SERS Detection of Bacteria in Water by in Situ Coating with Ag Nanoparticles. *Anal Chem* **86**, 1525-1533, doi:Doi 10.1021/Ac402935p (2014).

166 Efrima, S. & Bronk, B. V. Silver colloids impregnating or coating bacteria. *J Phys Chem B* **102**, 5947-5950, doi:Doi 10.1021/Jp9813903 (1998).
167 Efrima, S. & Zeiri, L. Understanding SERS of bacteria. *J Raman Spectrosc* **40**, 277-288, doi:Doi 10.1002/Jrs.2121 (2009).
168 Tindall, B. J. *et al.* Cultivatable microbial biodiversity: gnawing at the Gordian knot. *Environ Microbiol* **2**, 310-318, doi:DOI 10.1046/j.1462-2920.2000.00108.x (2000).
169 Rinnan, A., van den Berg, F. & Engelsen, S. B. Review of the most common pre-processing techniques for near-infrared spectra. *Trac-Trend Anal Chem* **28**, 1201-1222, doi:DOI 10.1016/j.trac.2009.07.007 (2009).
170 Boulet, J. C. & Roger, J. M. Pretreatments by means of orthogonal projections. *Chemometr Intell Lab* **117**, 61-69, doi:DOI 10.1016/j.chemolab.2012.02.002 (2012).
171 Norris, K. H. & Williams, P. C. Optimization of Mathematical Treatments of Raw near-Infrared Signal in the Measurement of Protein in Hard Red Spring Wheat .1. Influence of Particle-Size. *Cereal Chem* **61**, 158-165 (1984).
172 Savitzky, A. & Golay, M. J. E. Smoothing + Differentiation of Data by Simplified Least Squares Procedures. *Anal Chem* **36**, 1627-&, doi:Doi 10.1021/Ac60214a047 (1964).
173 Brereton, R. G. The evolution of chemometrics. *Anal Methods-Uk* **5**, 3785-3789, doi:Doi 10.1039/C3ay90051g (2013).
174 Bro, R. & Smilde, A. K. Principal component analysis. *Anal Methods-Uk* **6**, 2812-2831, doi:Doi 10.1039/C3ay41907j (2014).
175 Wold, S. Pattern-Recognition by Means of Disjoint Principal Components Models. *Pattern Recogn* **8**, 127-139, doi:Doi 10.1016/0031-3203(76)90014-5 (1976).
176 Ward, J. H. Hierarchical Grouping to Optimize an Objective Function. *J Am Stat Assoc* **58**, 236-&, doi:Doi 10.2307/2282967 (1963).
177 Chevallier, S., Bertrand, D., Kohler, A. & Courcoux, P. Application of PLS-DA in multivariate image analysis. *J Chemometr* **20**, 221-229, doi:Doi 10.1002/Cem.994 (2006).
178 Bino, R. J. *et al.* Potential of metabolomics as a functional genomics tool. *Trends Plant Sci* **9**, 418-425, doi:DOI 10.1016/j.tplants.2004.07.004 (2004).
179 Fiehn, O. Metabolomics - the link between genotypes and phenotypes. *Plant Mol Biol* **48**, 155-171, doi:Doi 10.1023/A:1013713905833 (2002).
180 Bollard, M. E., Stanley, E. G., Lindon, J. C., Nicholson, J. K. & Holmes, E. NMR-based metabonomic approaches for evaluating physiological influences on biofluid composition. *Nmr Biomed* **18**, 143-162, doi:Doi 10.1002/Nbm.935 (2005).
181 van der Greef, J. & Smilde, A. K. Symbiosis of chemometrics and metabolomics: past, present, and future. *J Chemometr* **19**, 376-386, doi:Doi 10.1002/Cem.941 (2005).
182 Westerhuis, J. A. *et al.* Assessment of PLSDA cross validation. *Metabolomics* **4**, 81-89, doi:DOI 10.1007/s11306-007-0099-6 (2008).
183 Anderssen, E., Dyrstad, K., Westad, F. & Martens, H. Reducing over-optimism in variable selection by cross-model validation. *Chemometr Intell Lab* **84**, 69-74, doi:DOI 10.1016/j.chemolab.2006.04.021 (2006).

184 Brereton, R. G. Consequences of sample size, variable selection, and model validation and optimisation, for predicting classification ability from analytical data. *Trac-Trend Anal Chem* **25**, 1103-1111, doi:DOI 10.1016/j.trac.2006.10.005 (2006).
185 Broadhurst, D. I. & Kell, D. B. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics* **2**, 171-196, doi:DOI 10.1007/s11306-006-0037-z (2006).
186 Barker, M. & Rayens, W. Partial least squares for discrimination. *J Chemometr* **17**, 166-173, doi:Doi 10.1002/Cem.785 (2003).
187 Ballabio, D. & Consonni, V. Classification tools in chemistry. Part 1: linear models. PLS-DA. *Anal Methods-Uk* **5**, 3790-3798, doi:Doi 10.1039/C3ay40582f (2013).
188 Cortes, C. & Vapnik, V. Support-Vector Networks. *Mach Learn* **20**, 273-297, doi:Doi 10.1023/A:1022627411411 (1995).
189 Burges, C. J. C. A tutorial on Support Vector Machines for pattern recognition. *Data Min Knowl Disc* **2**, 121-167, doi:Doi 10.1023/A:1009715923555 (1998).
190 Muller, K. R., Mika, S., Ratsch, G., Tsuda, K. & Scholkopf, B. An introduction to kernel-based learning algorithms. *Ieee T Neural Networ* **12**, 181-201, doi:Doi 10.1109/72.914517 (2001).
191 Smola, A. J. & Scholkopf, B. On a kernel-based method for pattern recognition, regression, approximation, and operator inversion. *Algorithmica* **22**, 211-231, doi:Doi 10.1007/Pl00013831 (1998).
192 Sanchez, V. D. Advanced support vector machines and kernel methods. *Neurocomputing* **55**, 5-20, doi:Doi 10.1016/S0925-2312(03)00373-4 (2003).
193 Xiaobo, Z., Jiewen, Z., Povey, M. J., Holmes, M. & Hanpin, M. Variables selection methods in near-infrared spectroscopy. *Analytica chimica acta* **667**, 14-32, doi:10.1016/j.aca.2010.03.048 (2010).
194 Ferreira, M. P., Grondona, A. E. B., Rolim, S. B. A. & Shimabukuro, Y. E. Analyzing the spectral variability of tropical tree species using hyperspectral feature selection and leaf optical modeling. *J Appl Remote Sens* **7**, doi:Artn 073502 Doi 10.1117/1.Jrs.7.073502 (2013).
195 Fenn, M. B., Pappu, V., Georgeiv, P. G. & Pardalos, P. M. Raman spectroscopy utilizing Fisher-based feature selection combined with Support Vector Machines for the characterization of breast cell lines. *J Raman Spectrosc* **44**, 939-948, doi:Doi 10.1002/Jrs.4309 (2013).
196 Owen, C. A. *et al.* In vitro toxicology evaluation of pharmaceuticals using Raman micro-spectroscopy. *J Cell Biochem* **99**, 178-186, doi:Doi 10.1002/Jcb.20884 (2006).
197 Pyrgiotakis, G. *et al.* Cell Death Discrimination with Raman Spectroscopy and Support Vector Machines. *Ann Biomed Eng* **37**, 1464-1473, doi:DOI 10.1007/s10439-009-9688-z (2009).
198 Notingher, I., Selvakumaran, J. & Hench, L. L. New detection system for toxic agents based on continuous spectroscopic monitoring of living cells. *Biosens Bioelectron* **20**, 780-789, doi:DOI 10.1016/j.bios.2004.04.008 (2004).
199 Jain, A. & Zongker, D. Feature selection: Evaluation, application, and small sample performance. *Ieee T Pattern Anal* **19**, 153-158, doi:Doi 10.1109/34.574797 (1997).

200 Wang, L., Slayden, R. A., Barry, C. E. & Liu, J. Cell wall structure of a mutant of Mycobacterium smegmatis defective in the biosynthesis of mycolic acids. *J Biol Chem* **275**, 7224-7229 (2000).
201 WHO. (WHO/TB/98.258) (World Health Organization, Geneva, 1998).
202 Boehme, C. C. *et al.* Rapid Molecular Detection of Tuberculosis and Rifampin Resistance. *N. Engl. J. Med.* **363**, 1005-1015, doi:doi:10.1056/NEJMoa0907847 (2010).
203 Evans, C. A. GeneXpert—A Game-Changer for Tuberculosis Control? *PLoS Med* **8**, e1001064, doi:10.1371/journal.pmed.1001064 (2011).
204 Butler, W. R. & Guthertz, L. S. Mycolic Acid Analysis by High-Performance Liquid Chromatography for Identification of Mycobacterium Species. *Clin Microbiol Rev* **14**, 704-726, doi:10.1128/cmr.14.4.704-726.2001 (2001).
205 Fujita, Y., Naka, T., McNeil, M. R. & Yano, I. Intact molecular characterization of cord factor (trehalose 6,6 '-dimycolate) from nine species of mycobacteria by MALDI-TOF mass spectrometry. *Microbiol-Sgm* **151**, 3403-3416, doi:DOI 10.1099/mic.0.28158-0 (2005).
206 Laval, F., Laneelle, M. A., Deon, C., Monsarrat, B. & Daffe, M. Accurate molecular mass determination of mycolic acids by MALDI-TOF mass spectrometry. *Anal Chem* **73**, 4537-4544 (2001).
207 Negri, P. & Dluhy, R. A. Ag nanorod based surface-enhanced Raman spectroscopy applied to bioanalytical sensing. *J Biophotonics* **6**, 20-35, doi:DOI 10.1002/jbio.201200133 (2013).
208 Liu, Y. J., Fan, J. G., Zhao, Y. P., Shanmukh, S. & Dluhy, R. A. Angle dependent surface enhanced Raman scattering obtained from a Ag nanorod array substrate. *Appl Phys Lett* **89**, doi:Artn 173134
Doi 10.1063/1.2369644 (2006).
209 Kates, M. *Techniques of Lipidology. Isolation, Analysis, and Identification of Lipids*. Vol. 3, Part 2 (Elsevier, 1986).
210 Ichihara, K. & Fukubayashi, Y. Preparation of fatty acid methyl esters for gas-liquid chromatography. *J Lipid Res* **51**, 635-640, doi:Doi 10.1194/Jlr.D001065 (2010).
211 Olivier, I. & Loots, D. T. A metabolomics approach to characterise and identify various Mycobacterium species. *J Microbiol Meth* **88**, 419-426, doi:DOI 10.1016/j.mimet.2012.01.012 (2012).
212 Chaney, S. B., Shanmukh, S., Dluhy, R. A. & Zhao, Y. P. Aligned silver nanorod arrays produce high sensitivity surface-enhanced Raman spectroscopy substrates. *Appl Phys Lett* **87**, doi:Artn 031908
Doi 10.1063/1.1988980 (2005).
213 Driskell, J. D. *et al.* The use of aligned silver nanorod Arrays prepared by oblique angle deposition as surface enhanced Raman scattering substrates. *J Phys Chem C* **112**, 895-901, doi:10.1021/jp075288u|ISSN 1932-7447 (2008).
214 Lambert, M. A., Moss, C. W., Silcox, V. A. & Good, R. C. Analysis of Mycolic Acid Cleavage Products and Cellular Fatty-Acids of Mycobacterium Species by Capillary Gas-Chromatography. *J Clin Microbiol* **23**, 731-736 (1986).
215 Karls, R. K., Guarner, J., McMurray, D. N., Birkness, K. A. & Quinn, F. D. Examination of Mycobacterium tuberculosis sigma factor mutants using low-dose

aerosol infection of guinea pigs suggests a role for SigC in pathogenesis. *Microbiol-Sgm* **152**, 1591-1600, doi:Doi 10.1099/Mic.028591-0 (2006).

216 Watanabe, M., Aoyagi, Y., Ridell, M. & Minnikin, D. E. Separation and characterization of individual mycolic acids in representative mycobacteria. *Microbiol-Sgm* **147**, 1825-1837 (2001).

217 Mederos, L., Valdivia, J. A. & Valero-Guillen, P. L. Analysis of the structure of mycolic acids of Mycobacterium simiae reveals a particular composition of alpha-mycolates in strain 'habana' TMC 5135, considered as immunogenic in tuberculosis and leprosy. *Microbiol-Sgm* **153**, 4159-4165, doi:DOI 10.1099/mic.0.2007/011262-0 (2007).

218 Uenishi, Y., Fujita, Y., Kusunose, N., Yano, I. & Sunagawa, M. Comprehensive analysis of mycolic acid subclass and molecular species composition of Mycobacterium bovis BCG Tokyo 172 cell wall skeleton (SMP-105). *J Microbiol Meth* **72**, 149-156, doi:DOI 10.1016/j.mimet.2007.11.016 (2008).

219 Glickman, M. S., Cahill, S. M. & Jacobs, W. R. The Mycobacterium tuberculosis cmaA2 gene encodes a mycolic acid trans-cyclopropane synthetase. *J Biol Chem* **276**, 2228-2233 (2001).

220 Laval, F. *et al.* Investigating the function of the putative mycolic acid methyltransferase UmaA - Divergence between the Mycobacterium smegmatis and Mycobacterium tuberculosis proteins. *J Biol Chem* **283**, 1419-1427, doi:DOI 10.1074/jbc.M708859200 (2008).

221 Shanmukh, S. *et al.* Rapid and sensitive detection of respiratory virus molecular signatures using a silver nanorod array SERS substrate. *Nano Lett* **6**, 2630-2636, doi:Doi 10.1021/Nl061666f (2006).

222 Maquelin, K. *et al.* Identification of medically relevant microorganisms by vibrational spectroscopy. *J Microbiol Meth* **51**, 255-271, doi:Pii S0167-7012(02)00127-6
Doi 10.1016/S0167-7012(02)00127-6 (2002).

223 Liu, T. T. *et al.* A High Speed Detection Platform Based on Surface-Enhanced Raman Scattering for Monitoring Antibiotic-Induced Chemical Changes in Bacteria Cell Wall. *Plos One* **4**, doi:Artn E5470
Doi 10.1371/Journal.Pone.0005470 (2009).

224 Naumann, D. FT-infrared and FT-Raman spectroscopy in biomedical research. *Appl Spectrosc Rev* **36**, 239-298 (2001).

225 Brennan, P. J. & Nikaido, H. The Envelope of Mycobacteria. *Annu Rev Biochem* **64**, 29-63, doi:DOI 10.1146/annurev.biochem.64.1.29 (1995).

226 Negri, P. & Dluhy, R. A. Ag nanorod based surface-enhanced Raman spectroscopy applied to bioanalytical sensing. *J Biophotonics* **6**, 20-35, doi:10.1002/jbio.201200133 (2013).

227 Jarvis, R. M. & Goodacre, R. Characterisation and identification of bacteria using SERS. *Chem Soc Rev* **37**, 931-936, doi:Doi 10.1039/B705973f (2008).

228 Driskell, J. D. *et al.* Infectious agent detection with SERS-active silver nanorod arrays prepared by oblique angle deposition. *Ieee Sens J* **8**, 863-870, doi:Doi 10.1109/Jsen.2008.922682 (2008).

229 Negri, P. *et al.* Direct Optical Detection of Viral Nucleoprotein Binding to an Anti-Influenza Aptamer. *Anal Chem* **84**, 5501-5508, doi:10.1021/ac202427e (2012).

230 Negri, P., Kage, A., Nitsche, A., Naumann, D. & Dluhy, R. A. Detection of viral nucleoprotein binding to anti-influenza aptamers via SERS. *Chem Comm* **47**, 8635-8637 (2011).
231 Buijtels, P. C. A. M. *et al.* Rapid identification of mycobacteria by Raman spectroscopy. *J Clin Microbiol* **46**, 961-965, doi:Doi 10.1128/Jcm.01763-07 (2008).
232 Barker, M. & Rayens, W. Partial least squares for discrimination. *J Chemometr* **17**, 166-173, doi:Doi 10.1002/Cem.785 (2003).
233 Musumarra, G., Barresi, V., Condorelli, D. F., Fortuna, C. G. & Scire, S. Potentialities of multivariate approaches in genome-based cancer research: identification of candidate genes for new diagnostics by PLS discriminant analysis. *J. Chemometr.* **18**, 125-132, doi:Doi 10.1002/Cem.846 (2004).
234 Hoang, V., Tripp, R. A., Rota, P. & Dluhy, R. A. Identification of individual genotypes of measles virus using surface enhanced Raman spectroscopy. *Analyst* **135**, 3103-3109, doi:Doi 10.1039/C0an00453g (2010).
235 Driskell, J. D. *et al.* Rapid and sensitive detection of rotavirus molecular signatures using surface enhanced Raman spectroscopy. *PLoS One* **5**, e10222, doi:doi:10.1371/journal.pone.0010222 (2010).
236 Bylesjö, M. *et al.* OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *J Chemometr* **20**, 341-351, doi:10.1002/cem.1006 (2006).
237 Brereton, R. G. *Chemometrics for Pattern Recognition*. (J. Wiley & Sons, Ltd., 2009).
238 Waitest, K. B., Balish, M. F. & Atkinson, T. P. New insights into the pathogenesis and detection of Mycoplasma pneumoniae infections. *Future Microbiol* **3**, 635-648, doi:Doi 10.2217/17460913.3.6.635 (2008).
239 Atkinson, T. P., Balish, M. F. & Waites, K. B. Epidemiology, clinical manifestations, pathogenesis and laboratory detection of Mycoplasma pneumoniae infections. *Fems Microbiology Reviews* **32**, 956-973, doi:DOI 10.1111/j.1574-6976.2008.00129.x (2008).
240 Shanmukh, S. *et al.* Identification and classification of respiratory syncytial virus (RSV) strains by surface-enhanced Raman spectroscopy and multivariate statistical techniques. *Anal Bioanal Chem* **390**, 1551-1555, doi:DOI 10.1007/s00216-008-1851-0 (2008).
241 Shanmukh, S. *et al.* Rapid and sensitive detection of respiratory virus molecular signatures using a silver nanorod array SERS substrate. *Nano Lett* **6**, 2630-2636, doi:Doi 10.1021/Nl061666f (2006).
242 Dong, W. F., Sukhorukov, G. B. & Mohwald, H. Enhanced Raman imaging and optical spectra of gold nanoparticle doped microcapsules. *Phys Chem Chem Phys* **5**, 3003-3012, doi:Doi 10.1039/B301591b (2003).
243 Hu, X. G. *et al.* Fabrication, characterization, and application in SERS of self-assembled polyelectrolyte-gold nanorod multilayered films. *J Phys Chem B* **109**, 19385-19389, doi:Doi 10.1021/Jp052706r (2005).
244 Fakhrullin, R. F. & Minullina, R. T. Hybrid Cellular-Inorganic Core-Shell Microparticles: Encapsulation of Individual Living Cells in Calcium Carbonate Microshells. *Langmuir* **25**, 6617-6621, doi:Doi 10.1021/La901395z (2009).

245 de Villiers, M. M., Otto, D. P., Strydom, S. J. & Lvov, Y. M. Introduction to nanocoatings produced by layer-by-layer (LbL) self-assembly. *Advanced Drug Delivery Reviews* **63**, 701-715, doi:DOI 10.1016/j.addr.2011.05.011 (2011).
246 Decher, G. & Hong, J. D. Buildup of Ultrathin Multilayer Films by a Self-Assembly Process .1. Consecutive Adsorption of Anionic and Cationic Bipolar Amphiphiles on Charged Surfaces. *Makromol Chem-M Symp* **46**, 321-327, doi:DOI 10.1002/masy.19910460145 (1991).
247 Diaspro, A., Silvano, D., Krol, S., Cavalleri, O. & Gliozzi, A. Single living cell encapsulation in nano-organized polyelectrolyte shells. *Langmuir* **18**, 5047-5050, doi:Doi 10.1021/La025646e (2002).
248 Ai, H., Jones, S. A. & Lvov, Y. M. Biomedical applications of electrostatic layer-by-layer nano-assembly of polymers, enzymes, and nanoparticles. *Cell Biochem Biophys* **39**, 23-43, doi:Doi 10.1385/Cbb:39:1:23 (2003).
249 Liu, Y. J., Wang, Y. X. & Claus, R. O. Layer-by-layer ionic self-assembly of Au colloids into multilayer thin-films with bulk metal conductivity. *Chemical Physics Letters* **298**, 315-319, doi:Doi 10.1016/S0009-2614(98)01209-3 (1998).
250 Cant, N. E. *et al.* Fabrication and characterization of self-assembled nanoparticle/polyelectrolyte multilayer films. *Journal of Physical Chemistry B* **107**, 13557-13562, doi:Doi 10.1021/Jp035388+ (2003).
251 Fakhrullin, R. F. & Lvov, Y. M. "Face-Lifting" and "Make-Up" for Microorganisms: Layer-by-Layer Polyelectrolyte Nanocoating. *Acs Nano* **6**, 4557-4564, doi:Doi 10.1021/Nn301776y (2012).
252 Fakhrullin, R. F. *et al.* Living Fungi Cells Encapsulated in Polyelectrolyte Shells Doped with Metal Nanoparticles. *Langmuir* **25**, 4628-4634, doi:Doi 10.1021/La803871z (2009).
253 Kahraman, M., Yazici, M. M., Sahin, F. & Culha, M. Experimental parameters influencing surface-enhanced Raman scattering of bacteria. *J Biomed Opt* **12**, doi:Artn 054015
Doi 10.1117/1.2798640 (2007).
254 Kenri, T. *et al.* Genotyping analysis of Mycoplasma pneumoniae clinical strains in Japan between 1995 and 2005: type shift phenomenon of M. pneumoniae clinical strains. *Journal of medical microbiology* **57**, 469-475, doi:10.1099/jmm.0.47634-0 (2008).
255 Krause, D. C., Leith, D. K., Wilson, R. M. & Baseman, J. B. Identification of Mycoplasma pneumoniae proteins associated with hemadsorption and virulence. *Infection and immunity* **35**, 809-817 (1982).
256 Dallo, S. F., Lazzell, A. L., Chavoya, A., Reddy, S. P. & Baseman, J. B. Biofunctional domains of the Mycoplasma pneumoniae P30 adhesin. *Infection and immunity* **64**, 2595-2601 (1996).
257 Tully, J. G., Rose, D. L., Whitcomb, R. F. & Wenzel, R. P. Enhanced Isolation of Mycoplasma-Pneumoniae from Throat Washings with a Newly Modified Culture-Medium. *J Infect Dis* **139**, 478-482 (1979).
258 Somerson, N. L., Kocka, J. P., Rose, D. & Delgiudice, R. A. Isolation of Acholeplasmas and a Mycoplasma from Vegetables. *Appl Environ Microb* **43**, 412-417 (1982).

259 Razin, S. Time-line of significant contributions to mycoplasmology. *Biologicals* **38**, 191-192, doi:DOI 10.1016/j.biologicals.2009.12.002 (2010).
260 Razin, S. & Hayflick, L. Highlights of mycoplasma research-An historical perspective. *Biologicals* **38**, 183-190, doi:DOI 10.1016/j.biologicals.2009.11.008 (2010).
261 Collins, S. P. *et al.* Advantages of Environmental Scanning Electron-Microscopy in Studies of Microorganisms. *Microsc Res Techniq* **25**, 398-405, doi:DOI 10.1002/jemt.1070250508 (1993).
262 Mendelsohn, J. D. *et al.* Fabrication of microporous thin films from polyelectrolyte multilayers. *Langmuir* **16**, 5017-5023, doi:Doi 10.1021/La000075g (2000).
263 Lavalle, P. *et al.* Comparison of the structure of polyelectrolyte multilayer films exhibiting a linear and an exponential growth regime: An in situ atomic force microscopy study. *Macromolecules* **35**, 4458-4465, doi:Doi 10.1021/Ma0119833 (2002).
264 Han, L. L. *et al.* Directional cell migration through cell-cell interaction on polyelectrolyte multilayers with swelling gradients. *Biomaterials* **34**, 975-984, doi:DOI 10.1016/j.biomaterials.2012.10.041 (2013).
265 Kapuscinski, J. & Szer, W. Interactions of 4', 6-diamidine-2-phenylindole with synthetic polynucleotides. *Nucleic Acids Res* **6**, 3519-3534 (1979).
266 Kapuscinski, J. & Yanagi, K. Selective staining by 4', 6-diamidine-2-phenylindole of nanogram quantities of DNA in the presence of RNA on gels. *Nucleic Acids Res* **6**, 3535-3542 (1979).
267 Estrela-Lopis, I. *et al.* SANS studies of polyelectrolyte multilayers on colloidal templates. *Langmuir* **18**, 7861-7866, doi:Doi 10.1021/La0258603 (2002).
268 Puppels, G. J. *et al.* Studying single living cells and chromosomes by confocal Raman microspectroscopy. *Nature* **347**, 301-303, doi:10.1038/347301a0 (1990).
269 Negri, P. *et al.* Direct Optical Detection of Viral Nucleoprotein Binding to an Anti-Influenza Aptamer. *Anal Chem* **84**, 5501-5508, doi:Doi 10.1021/Ac202427e (2012).
270 Premasiri, W. R. *et al.* Characterization of the surface enhanced raman scattering (SERS) of bacteria. *The journal of physical chemistry. B* **109**, 312-320, doi:10.1021/jp040442n (2005).
271 Patel, I. S., Premasiri, W. R., Moir, D. T. & Ziegler, L. D. Barcoding bacterial cells: A SERS based methodology for pathogen identification. *J Raman Spectrosc* **39**, 1660-1672, doi:10.1002/jrs.2064 (2008).
272 Maquelin, K. *et al.* Identification of medically relevant microorganisms by vibrational spectroscopy. *J Microbiol Meth* **51**, 255-271, doi:Pii S0167-7012(02)00127-6
Doi 10.1016/S0167-7012(02)00127-6 (2002).
273 Goeller, L. J. & Riley, M. R. Discrimination of bacteria and bacteriophages by Raman spectroscopy and surface-enhanced Raman spectroscopy. *Applied spectroscopy* **61**, 679-685, doi:Doi 10.1366/000370207781393217 (2007).
274 Jarvis, R. M. & Goodacre, R. Characterisation and identification of bacteria using SERS. *Chem Soc Rev* **37**, 931-936, doi:Doi 10.1039/B705973f (2008).

275 Negri, P., Kage, A., Nitsche, A., Naumann, D. & Dluhy, R. A. Detection of viral nucleoprotein binding to anti-influenza aptamers via SERS. *Chem Commun (Camb)* **47**, 8635-8637, doi:10.1039/c0cc05433j (2011).
276 Sengupta, A., Mujacic, M. & Davis, E. J. Detection of bacteria by surface-enhanced Raman spectroscopy. *Anal Bioanal Chem* **386**, 1379-1386, doi:DOI 10.1007/s00216-006-0711-z (2006).
277 Musumarra, G., Barresi, V., Condorelli, D. F., Fortuna, C. G. & Scire, S. Potentialities of multivariate approaches in genome-based cancer research: identification of candidate genes for new diagnostics by PLS discriminant analysis. *Journal of Chemometrics* **18**, 125-132, doi:Doi 10.1002/Cem.846 (2004).
278 Dorigo-Zetsma, J. W. *et al.* Comparison of PCR, culture, and serological tests for diagnosis of Mycoplasma pneumoniae respiratory tract infection in children. *J Clin Microbiol* **37**, 14-17 (1999).
279 Sluijter, M., Spuesens, E. B., Hartwig, N. G., van Rossum, A. M. & Vink, C. The Mycoplasma pneumoniae MPN490 and Mycoplasma genitalium MG339 genes encode reca homologs that promote homologous DNA strand exchange. *Infection and immunity* **77**, 4905-4911, doi:10.1128/IAI.00747-09 (2009).
280 Atkinson, T. P., Balish, M. F. & Waites, K. B. Epidemiology, clinical manifestations, pathogenesis and laboratory detection of Mycoplasma pneumoniae infections. *FEMS microbiology reviews* **32**, 956-973, doi:10.1111/j.1574-6976.2008.00129.x (2008).
281 Wilson, M. H. & Collier, A. M. Ultrastructural study of Mycoplasma pneumoniae in organ culture. *Journal of bacteriology* **125**, 332-339 (1976).
282 Pereyre, S., Charron, A., Renaudin, H., Bebear, C. & Bebear, C. M. First report of macrolide-resistant strains and description of a novel nucleotide sequence variation in the P1 adhesin gene in Mycoplasma pneumoniae clinical strains isolated in France over 12 years. *Journal of clinical microbiology* **45**, 3534-3539, doi:10.1128/JCM.01345-07 (2007).
283 Dallo, S. F., Chavoya, A. & Baseman, J. B. Characterization of the gene for a 30-kilodalton adhesion-related protein of Mycoplasma pneumoniae. *Infection and immunity* **58**, 4163-4165 (1990).
284 Dumke, R., Catrein, I., Pirkil, E., Herrmann, R. & Jacobs, E. Subtyping of Mycoplasma pneumoniae isolates based on extended genome sequencing and on expression profiles. *International journal of medical microbiology : IJMM* **292**, 513-525 (2003).
285 Ai, H. *et al.* Coating and selective deposition of nanofilm on silicone rubber for cell adhesion and growth. *Cell Biochem Biophys* **38**, 103-114, doi:Doi 10.1385/Cbb:38:2:103 (2003).
286 Hillberg, A. L. & Tabrizian, M. Biorecognition through layer-by-layer polyelectrolyte assembly: In-situ hybridization on living cells. *Biomacromolecules* **7**, 2742-2750, doi:Doi 10.1021/Bm060266j (2006).
287 Matsuzawa, A., Matsusaki, M. & Akashi, M. Effectiveness of nanometer-sized extracellular matrix layer-by-layer assembled films for a cell membrane coating protecting cells from physical stress. *Langmuir* **29**, 7362-7368, doi:10.1021/la303459v (2013).

288 Svaldo Lanero, T., Cavalleri, O., Krol, S., Rolandi, R. & Gliozzi, A. Mechanical properties of single living cells encapsulated in polyelectrolyte matrixes. *J Biotechnol* **124**, 723-731, doi:DOI 10.1016/j.jbiotec.2006.02.016 (2006).
289 Orive, G. *et al.* Cell encapsulation: Promise and progress. *Nat Med* **9**, 104-107, doi:Doi 10.1038/Nm0103-104 (2003).
290 Batorsky, A., Liao, J. H., Lund, A. W., Plopper, G. E. & Stegemann, J. P. Encapsulation of adult human mesenchymal stem cells within collagen-agarose microenvironments. *Biotechnol Bioeng* **92**, 492-500, doi:Doi 10.1002/Bit.20614 (2005).
291 Decher, G. & Hong, J. D. Buildup of Ultrathin Multilayer Films by a Self-Assembly Process .2. Consecutive Adsorption of Anionic and Cationic Bipolar Amphiphiles and Polyelectrolytes on Charged Surfaces. *Ber Bunsen Phys Chem* **95**, 1430-1434 (1991).
292 Kenri, T. *et al.* Genotyping analysis of Mycoplasma pneumoniae clinical strains in Japan between 1995 and 2005: type shift phenomenon of M-pneumoniae clinical strains. *J Med Microbiol* **57**, 469-475, doi:DOI 10.1099/jmm.0.47634-0 (2008).
293 Krause, D. C., Leith, D. K., Wilson, R. M. & Baseman, J. B. Identification of Mycoplasma-Pneumoniae Proteins Associated with Hemadsorption and Virulence. *Infection and immunity* **35**, 809-817 (1982).
294 Mabbott, S. *et al.* The optimisation of facile substrates for surface enhanced Raman scattering through galvanic replacement of silver onto copper. *Analyst* **137**, 2791-2798, doi:Doi 10.1039/C2an35323g (2012).
295 Starchenko, V., Muller, M. & Lebovka, N. Sizing of PDADMAC/PSS Complex Aggregates by Polyelectrolyte and Salt Concentration and PSS Molecular Weight. *J Phys Chem B* **116**, 14961-14967, doi:Doi 10.1021/Jp3095243 (2012).
296 Wang, L. X. & Yu, H. Chain Conformation of Linear Poly-Electrolyte in Salt-Solutions - Sodium Poly(Styrene Sulfonate) in Potassium-Chloride and Sodium-Chloride Solutions. *Macromolecules* **21**, 3498-3501, doi:Doi 10.1021/Ma00190a026 (1988).
297 Peitzsch, R. M., Burt, M. J. & Reed, W. F. Evidence of Partial Draining for Linear Polyelectrolytes - Heparin, Chondroitin 6-Sulfate, and Poly(Styrenesulfonate). *Macromolecules* **25**, 806-815, doi:Doi 10.1021/Ma00028a047 (1992).
298 Beer, M., Schmidt, M. & Muthukumar, M. The electrostatic expansion of linear polyelectrolytes: Effects of gegenions, co-ions, and hydrophobicity. *Macromolecules* **30**, 8375-8385, doi:Doi 10.1021/Ma9709821 (1997).
299 Bonvin-Klotz, L. *et al.* Domain analysis of lipoprotein LppQ in Mycoplasma mycoides subsp. mycoides SC. *Antonie van Leeuwenhoek* **93**, 175-183, doi:10.1007/s10482-007-9191-1 (2008).
300 Mueller, R., Kohler, K., Weinkamer, R., Sukhorukov, G. & Fery, A. Melting of PDADMAC/PSS capsules investigated with AFM force spectroscopy. *Macromolecules* **38**, 9766-9771, doi:Doi 10.1021/Ma0513057 (2005).
301 Irigoyen, J. *et al.* Responsive Polyelectrolyte Multilayers Assembled at High Ionic Strength with an Unusual Collapse at Low Ionic Strength. *Macromol Rapid Comm* **33**, 1964-1969, doi:DOI 10.1002/marc.201200471 (2012).
302 Ramos, J. J. I. & Moya, S. E. Water Content of Hydrated Polymer Brushes Measured by an In Situ Combination of a Quartz Crystal Microbalance with

Dissipation Monitoring and Spectroscopic Ellipsometry. *Macromol Rapid Comm* **32**, 1972-1978, doi:DOI 10.1002/marc.201100455 (2011).

303 Puppels, G. J. *et al.* Studying Single Living Cells and Chromosomes by Confocal Raman Microspectroscopy. *Nature* **347**, 301-303, doi:Doi 10.1038/347301a0 (1990).

304 Lu, X. N. *et al.* Comprehensive Detection and Discrimination of Campylobacter Species by Use of Confocal Micro-Raman Spectroscopy and Multilocus Sequence Typing. *Journal of clinical microbiology* **50**, 2932-2946, doi:Doi 10.1128/Jcm.01144-12 (2012).

305 Jarvis, R. M., Blanch, E. W., Golovanov, A. P., Screen, J. & Goodacre, R. Quantification of casein phosphorylation with conformational interpretation using Raman spectroscopy. *Analyst* **132**, 1053-1060, doi:Doi 10.1039/B702944f (2007).

306 Mrozek, M. F. & Weaver, M. J. Detection and identification of aqueous saccharides by using surface-enhanced Raman spectroscopy. *Anal Chem* **74**, 4069-4075, doi:Doi 10.1021/Ac020115g (2002).

307 Dorigo-Zetsma, J. W. *et al.* Comparison of PCR, culture, and serological tests for diagnosis of Mycoplasma pneumoniae respiratory tract infection in children. *Journal of clinical microbiology* **37**, 14-17 (1999).

308 Loens, K., Ursi, D., Goossens, H. & Ieven, M. Molecular diagnosis of Mycoplasma pneumoniae respiratory tract infections. *Journal of clinical microbiology* **41**, 4915-4923, doi:Doi 10.1128/Jcm.41.11.4915-4923.2003 (2003).

309 Waring, A. L. *et al.* Development of a genomics-based PCR assay for detection of Mycoplasma pneumoniae in a large outbreak in New York State. *Journal of clinical microbiology* **39**, 1385-1390, doi:Doi 10.1128/Jcm.39.4.1385-1390.2001 (2001).

310 Su, C. J., Chavoya, A., Dallo, S. F. & Baseman, J. B. Sequence Divergency of the Cytadhesin Gene of Mycoplasma-Pneumoniae. *Infection and immunity* **58**, 2669-2674 (1990).

311 Chang, H. Y., Prince, O. A., Sheppard, E. S. & Krause, D. C. Processing Is Required for a Fully Functional Protein P30 in Mycoplasma pneumoniae Gliding and Cytadherence. *Journal of bacteriology* **193**, 5841-5846, doi:Doi 10.1128/Jb.00104-11 (2011).

312 Taylor-Robinson, D. & Bebear, C. Antibiotic susceptibilities of mycoplasmas and treatment of mycoplasmal infections. *J Antimicrob Chemoth* **40**, 622-630, doi:Doi 10.1093/Jac/40.5.622 (1997).

313 Taylor-Robinson, D. Infections due to species of Mycoplasma and Ureaplasma: an update. *Clin Infect Dis* **23**, 671-682; quiz 683-674 (1996).

314 Korppi, M., Heiskanen-Kosma, T. & Kleemola, M. Incidence of community-acquired pneumonia in children caused by Mycoplasma pneumoniae: Serological results of a prospective, population-based study in primary health care. *Respirology* **9**, 109-114, doi:DOI 10.1111/j.1440-1843.2003.00522.x (2004).

315 Daxboeck, F., Krause, R. & Wenisch, C. Laboratory diagnosis of Mycoplasma pneumoniae infection. *Clin Microbiol Infec* **9**, 263-273, doi:DOI 10.1046/j.1469-0691.2003.00590.x (2003).

316 Mackay, I. M. Real-time PCR in the microbiology laboratory. *Clin Microbiol Infec* **10**, 190-212, doi:DOI 10.1111/j.1198-743X.2004.00722.x (2004).

317 Hauck, T. S., Giri, S., Gao, Y. L. & Chan, W. C. W. Nanotechnology diagnostics for infectious diseases prevalent in developing countries. *Adv Drug Deliver Rev* **62**, 438-448, doi:DOI 10.1016/j.addr.2009.11.015 (2010).
318 Mohr, C., Hofmeister, H., Radnik, J. & Claus, P. Identification of active sites in gold-catalyzed hydrogenation of acrolein. *J Am Chem Soc* **125**, 1905-1911, doi:Doi 10.1021/Ja027321q (2003).
319 Hone, D. C. *et al.* Generation of cytotoxic singlet oxygen via phthalocyanine-stabilized gold nanoparticles: A potential delivery vehicle for photodynamic therapy. *Langmuir* **18**, 2985-2987, doi:Doi 10.1021/La0256230 (2002).
320 Cao, Y. W. C., Jin, R. C. & Mirkin, C. A. Nanoparticles with Raman spectroscopic fingerprints for DNA and RNA detection. *Science* **297**, 1536-1540, doi:DOI 10.1126/science.297.5586.1536 (2002).
321 Constantino, C. J. L., Lemma, T., Antunes, P. A. & Aroca, R. Single-molecule detection using surface-enhanced resonance Raman scattering and Langmuir-Blodgett monolayers. *Anal Chem* **73**, 3674-3678, doi:Doi 10.1021/Ac0101961 (2001).
322 Lee, S. J., Morrill, A. R. & Moskovits, M. Hot spots in silver nanowire bundles for surface-enhanced Raman spectroscopy. *J Am Chem Soc* **128**, 2200-2201, doi:Doi 10.1021/Ja0578350 (2006).
323 Owen, C. A. *et al.* In vitro toxicology evaluation of pharmaceuticals using Raman micro-spectroscopy. *Journal of cellular biochemistry* **99**, 178-186, doi:10.1002/jcb.20884 (2006).
324 Pyrgiotakis, G. *et al.* Cell death discrimination with Raman spectroscopy and support vector machines. *Annals of biomedical engineering* **37**, 1464-1473, doi:10.1007/s10439-009-9688-z (2009).
325 Pyrgiotakis, G., Kundakcioglu, O. E., Pardalos, P. M. & Moudgil, B. M. Raman spectroscopy and support vector machines for quick toxicological evaluation of titania nanoparticles. *J Raman Spectrosc* **42**, 1222-1231, doi:Doi 10.1002/Jrs.2839 (2011).
326 Tully, J. G. & Taylor-Robinson, D. Taxonomy and host distribution of the ureaplasmas. *Pediatric infectious disease* **5**, S292-295 (1986).
327 Kim, K. S. *et al.* Use of rpoB sequences for phylogenetic study of Mycoplasma species. *Fems Microbiol Lett* **226**, 299-305, doi:Doi 10.1016/S0378-1097(03)00618-9 (2003).
328 Kahraman, M., Yazici, M. M., Sahin, F., Bayrak, O. F. & Culha, M. Reproducible surface-enhanced Raman scattering spectra of bacteria on aggregated silver nanoparticles. *Applied spectroscopy* **61**, 479-485, doi:10.1366/000370207780807731 (2007).
329 Horowitz, S., Evinson, B., Borer, A. & Horowitz, J. Mycoplasma fermentans in rheumatoid arthritis and other inflammatory arthritides. *The Journal of rheumatology* **27**, 2747-2753 (2000).
330 Sohaeverbeke, T., Gilroy, C. B., Bebear, C., Dehais, J. & Taylor-Robinson, D. Mycoplasma fermentans in joints of patients with rheumatoid arthritis and other joint disorders. *Lancet* **347**, 1418 (1996).

APPENDIX 1

Supplementary Materials

Spectroscopic Analyses of Lipid Profiles from Mycobacteria

Omar E. Rivera-Betancourt[1], Russell Karls[2], Benjamin Grosse-Siestrup[2] Frederick Quinn[2], and Richard Dluhy[1*]

[1]Department of Chemistry, University of Georgia, Athens, GA 30602

[2]Department of Infectious Diseases, University of Georgia, Athens, GA 30602

Published in: *Analyst*, 138, 2013, 6774-6785

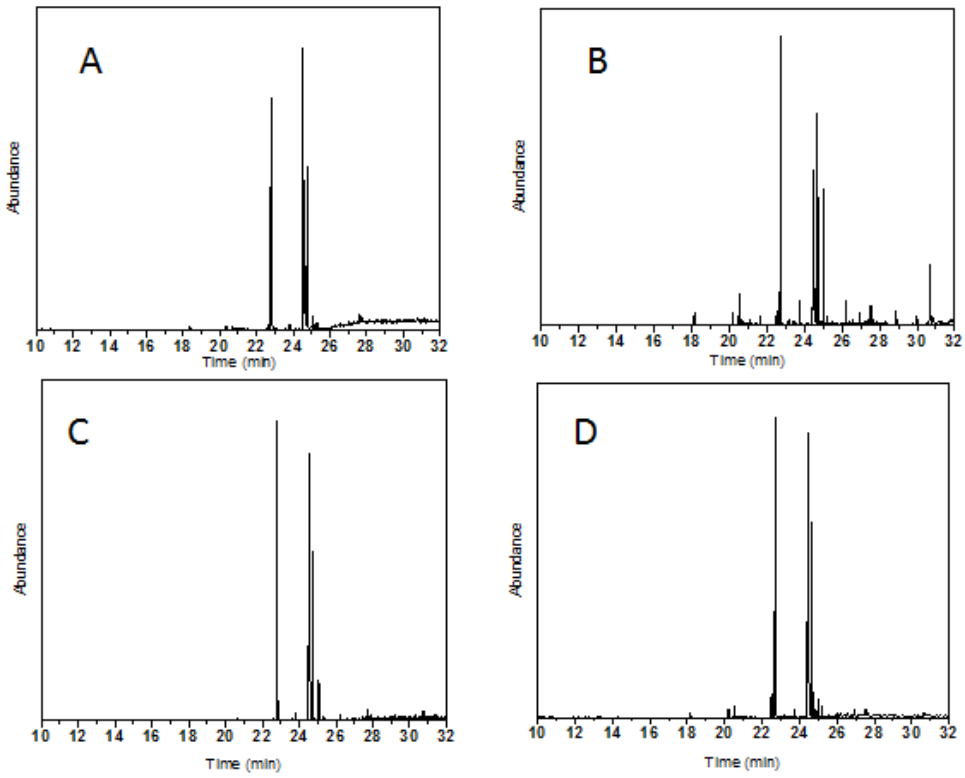

Figure A1.1 Representative total ion chromatograms of four of the *Mycobacterium* species investigated by GC-MS. (a) *M. smegmatis* ; (b) *M. avium* ; (c) *M. tuberculosis* H37Rv ; (d) *M. tuberculosis* Erdman.

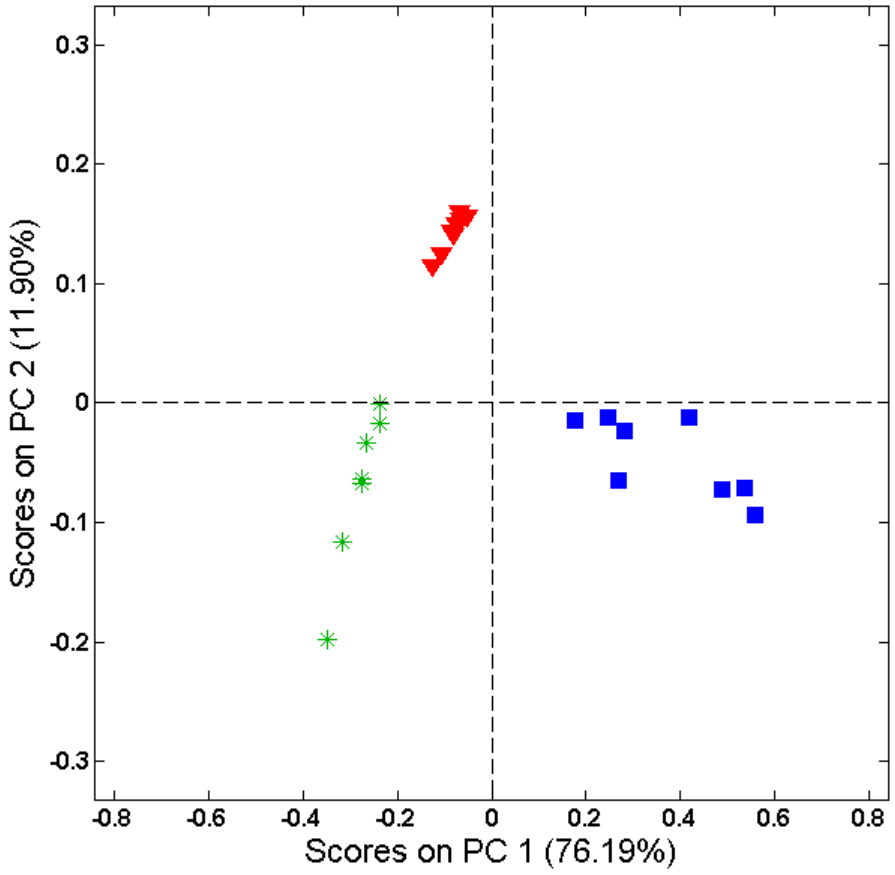

Figure A1.2 PCA scores plot corresponding to *M. smegmatis* (❶), *M. avium* (⑤), *M. bovis* BCG (☒) species. The PC model was constructed from the SERS spectra of the corresponding species using the spectral range 700-1700 cm^{-1}.

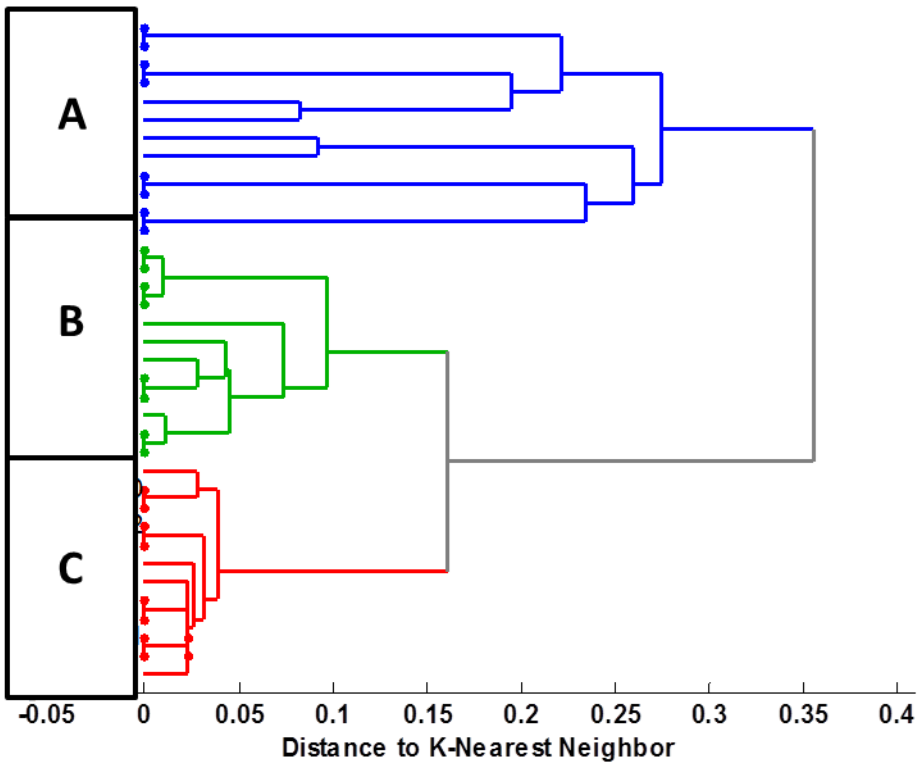

Figure A1.3 A hierarchical cluster analysis dendrogram derived from the PC scores of the NTM species. The nodes group into three recognized clusters and are labeled according to the samples. (A) *M. avium* (blue) ; (B) *M. bovis* BCG (green) ; (C) *M. smegmatis* (red).

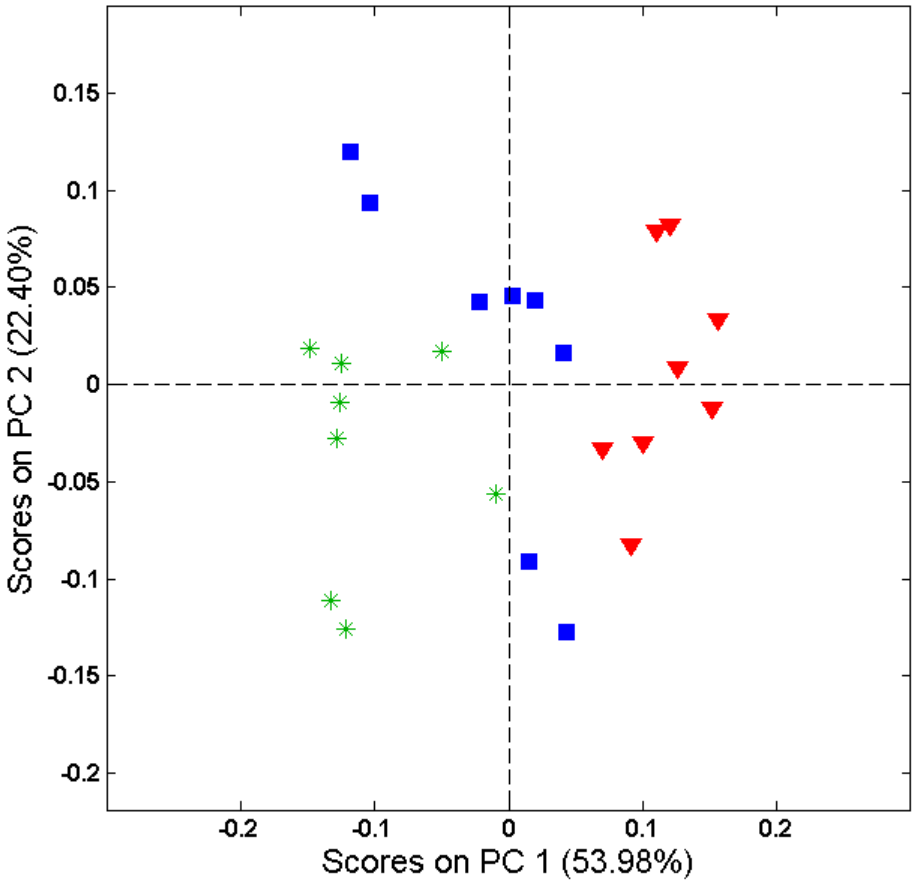

Figure A1.4 PCA scores plot corresponding to *M. tuberculosis* Erdman strains. Parent (⑤), *sigC* mutant (⓿), and *sigC* complement (☒) strains. The PC model was constructed from the SERS spectra of the corresponding species using the spectral range 700 - 1700 cm^{-1}.

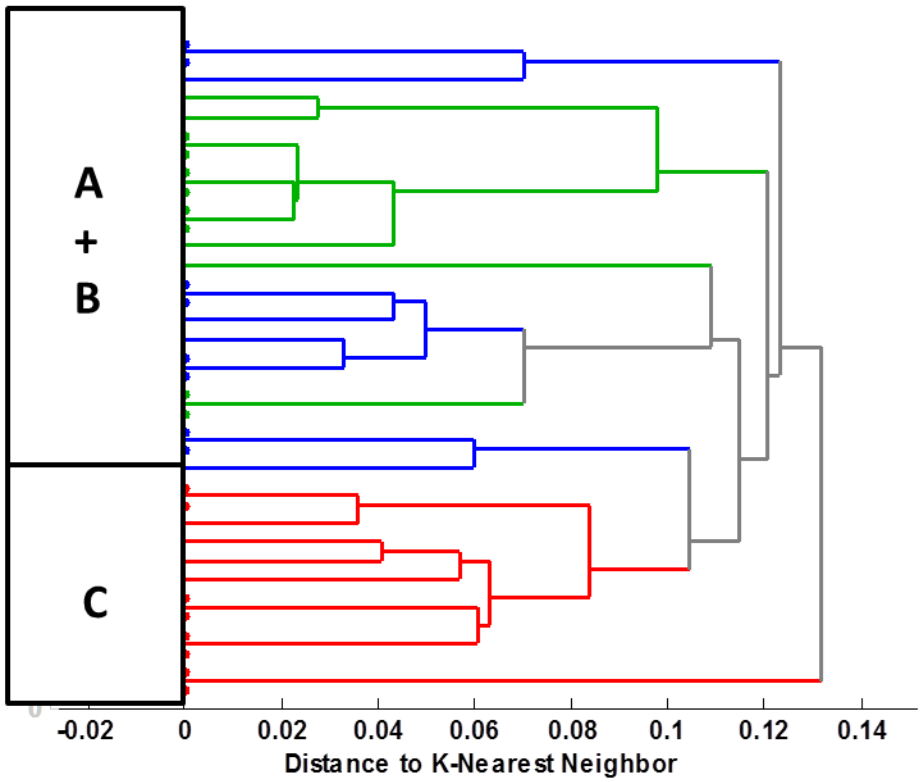

Figure A1.5 A hierarchical cluster analysis dendrogram derived from the PC scores of the *M. tuberculosis* Erdman strains. The nodes group into two distinct clusters and are labeled according to the samples. (A + B) wild-type (blue) and *sigC* complement (green) ; (C) *sig C* mutant (red).

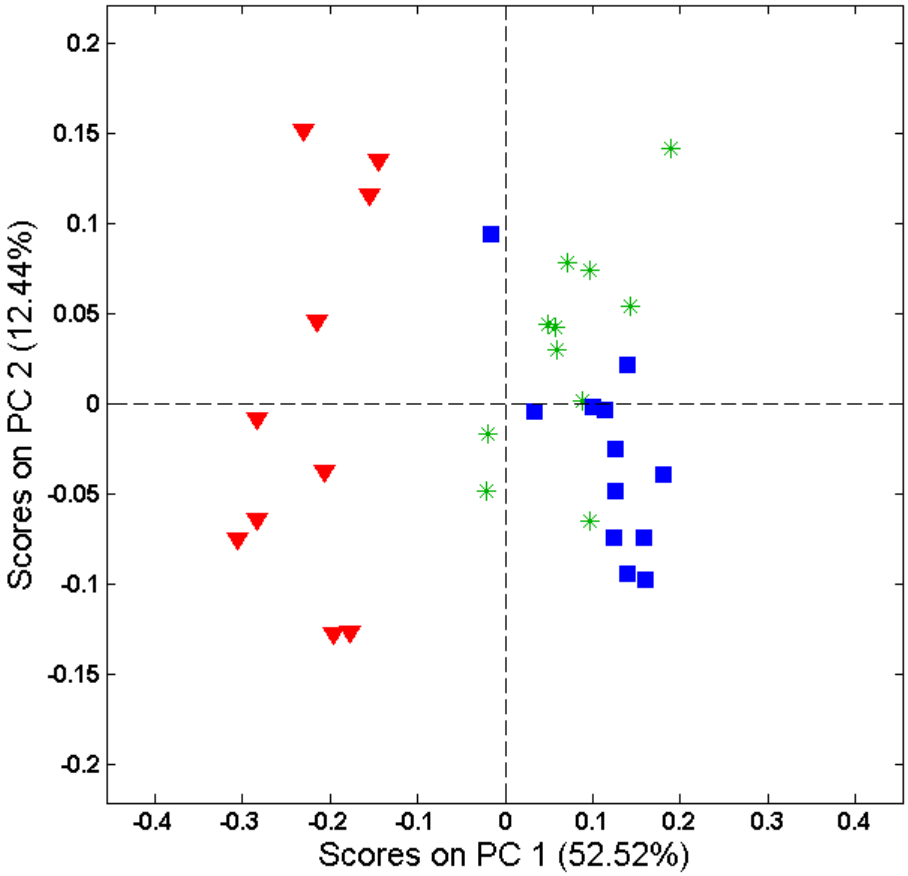

Figure A1.6 PCA scores plot corresponding to *M. tuberculosis* H37Rv strains. Parent (⑤), *sigC* mutant (⓿), and *sigC* complement (☒) strains. The PC model was constructed from the SERS spectra of the corresponding species using the spectral range 700 - 1700 cm^{-1}.

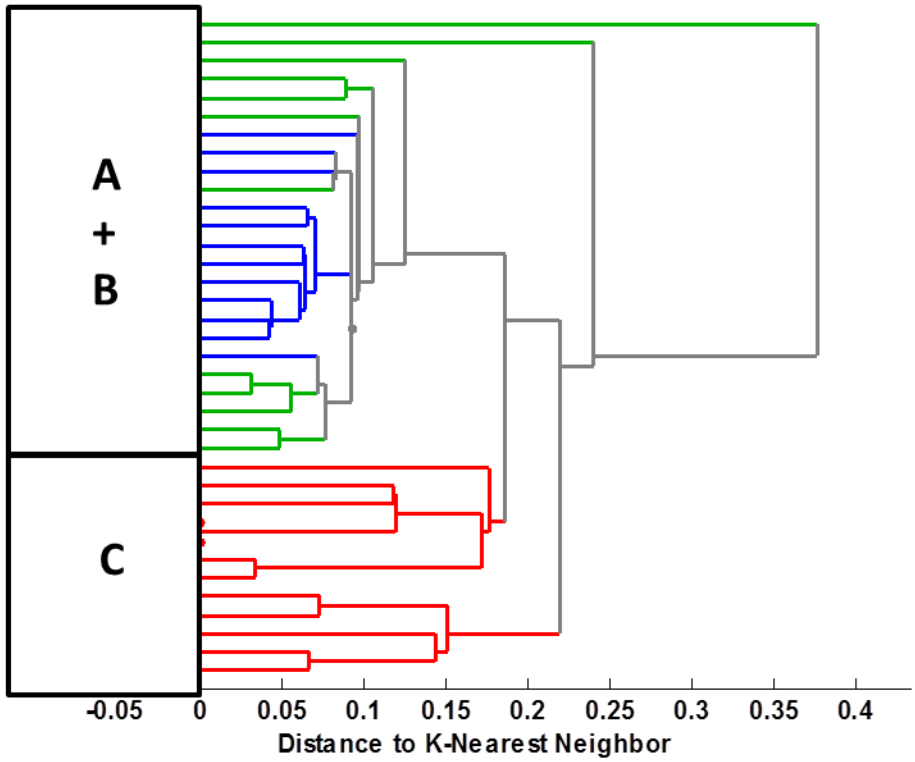

Figure A1.7 A hierarchical cluster analysis dendrogram derived from the PC scores of the *M. tuberculosis* H37Rv strains. The nodes group into two distinct clusters and are labeled according to the samples. (A + B) parent (blue) and *sigC* complement (green) ; (C) *sigC* mutant (red).

Table A1.1 Representative ^{1}H-NMR chemical shifts appearing in the spectra of the non-tuberculous mycobacteria *M. smegmatis*, *M. bovis* and *M. avium*.

Species		***M.Smegmatis***	***BCG***	***M. Avium***
	Mycolic Acid Type Sub Class	keto-mycolate k4	keto- & alpha k3, α1	keto-mycolate k3,k4
Functional Group			**Signal (ppm)**	
$CHCH_3$		0.85	0.85	0.85
CHOR			2.96	2.96
$CHOCH_3$			3.32	3.32
$(CH_2)_n CH_3$		0.85	0.85	0.85
$CH_2 (CH_2)_n CH_2$		1.0-1.5	1.0-1.5	1.0-1.5
$CH(CO_2CH_3)$		2.45	2.45	2.45
CO_2CH_3		3.7	3.7	3.7
trans-CH=CH			5.39,5.35	5.39,5.35
cis-CH=CH		5.3		
cis-cyclopropane			0.6,-0.3	

Table A1.2 Representative ^{1}H-NMR chemical shifts appearing in the spectra of the *M. tuberculosis* clinical Erdman strain.

Species		**Wildtype**	**Complimen t**	**Mutant**
	Mycolic Acid Type	keto-mycolate	keto-mycolate	keto-mycolate
Functional Group			**Signal (ppm)**	
$CHCH_3$		0.85	0.85	0.85
CHOR			2.96	2.96
$CHOCH_3$			3.32	3.32
$(CH_2)_n CH_3$		0.85	0.85	0.85
$CH_2 (CH_2)_n CH_2$		1.0-1.5	1.0-1.5	1.0-1.5
CO_2CH_3		3.7	3.7	3.7
trans-CH=CH				5.39,5.35
cis-CH=CH		5.3	5.3	
cis-cyclopropane		0.64,-0.36	0.64,-0.36	0.64,-0.36
trans-cyclopropane		0.04	0.04	0.04

Table A1.3 Representative ^{1}H-NMR chemical shifts appearing in the spectra of the *M. tuberculosis* clinical H37Rv strain.

Species		**Wildtype**	**Complimen t**	**Mutant**
Functional Group	Mycolic Acid Type	keto-mycolate	keto-mycolate **Signal (ppm)**	keto-mycolate
$CHCH_3$		0.85	0.85	0.85
CHOR			2.96	2.96
$CHOCH_3$			3.32	3.32
$(CH_2)_n CH_3$		0.85	0.85	0.85
$CH_2 (CH_2)_n CH_2$		1.0-1.5	1.0-1.5	1.0-1.5
CO_2CH_3		3.7	3.7	3.7
trans-CH=CH				5.39,5.35
cis-CH=CH		5.3	5.3	
cis-cyclopropane		0.64,-0.36	0.64,-0.36	0.64,-0.36

Table A1.4 Representative Raman bands appearing in the SERS spectra of the non-tuberculous mycobacteria *M. smegmatis*, *M. bovis* and *M. avium*.

Raman Shift, cm^{-1}			Vibrational Mode Assignment
M.smegmatis	*M.avium*	*M. bovis* BCG	
1654	1612	1659	(C=C); C=O-β conjugated; v_s(C=O) carboxylic acid
1598	1594	1593	
		1500	v(C=C)
1442	1442	1441	δ(C-H_2) sci.; CH_3 antisym. bend
	1393		C-O-H bend 1° alcohol
	1350	1366	C-O-H bend; $(CH_2)_n$ in-phase twist
1299	1289	1298	C-O-H bend
		1265	cis dialkyl C-H sym. rock
1240	1242		CH-O epoxy ring breathing mode
1210		1210	C-O-H bend
1164	1170	1164	v_{as}(COC); δ(CH)
1129	1135	1135	C-C skel. str in alkane
	1101	1097	v_{as}(COC); C-C skeletal
1058		1065	v(CHR_2) C-C skel. str ; C-C skel. str in alkane; v_{as}(COC)
1028		1030	oop C-C-O stretch
1001	1003	1004	v(C-C-O) out-of-phase stretch of primary alcohol
	970	962	O-H-O wagging; trans dialkyl wag
	931		v(CHR_2) C-C skel. str where R≠CH_3; COO^- str carboxylic acid
893			v(C-C-O) in-phase stretch of primary alcohol, v_S(C-O-C); C-C skel. str in alkane; COO^- str carboxylic acid
841	855		v(C-C-O) in-phase stretch of primary alcohol; C-C skel. str in alkane; COO^- str carboxylic acid
		832	v(C-C-O) in-phase stretch of primary alcohol; v(CHR_2) C-C skel. str where R≠CH_3
773		778	CH_2 in-phase rock; v(CHR_2) -C-C skel. str where R≠CH_3
698	672	696	cis dialkyl C-H wag

Table A1.5 Representative Raman bands appearing in the SERS spectra of the *M. tuberculosis* clinical Erdman strains.

Raman Shift, cm^{-1}			Vibrational Mode Assignment
Wild Type	EΔsigC	EΔCcomp	
1592	1586	1592	C=O, alkyl ketone; (C=C); C=O-β conjugated; v_s(C=O) carboxylic acid
1564	1564	1564	C=C
	1458	1492	CH_3 antisym. bend.
1458		1457	δ(C-H_2) sci.
1391	1387	1392	C-O-H bend
1319	1315	1314	C-O-H bend; $(CH_2)_n$ in-phase twist
1275	1280	1280	C-O-H bend
1241	1244	1241	C-O-H bend
1158	1157	1163	
1130		1130	C-C skel. str in alkane
1075	1076	1080	v_{as}(COC)
1008	1005	1008	v(C-C-O) out-of-phase stretch of primary alcohol
963			OHO wagging; trans dialkyl wag
	928		v(CHR_2) C-C skel. str where R≠CH_3;
	892		v(C-C-O) in-phase stretch of primary alcohol, v_S(C-O-C); C-C skel. str in alkane
857	856	852	v(C-C-O) in-phase stretch of primary alcohol; C-C skel. str in alkane
786			v(C-C-O) in-phase stretch of primary alcohol; v(CHR_2) C-C skel. str where R≠CH_3
	714	713	CH_2 in-phase rock; v(CHR_2) C-C skel. str where R≠CH_3
707			cis dialkyl C-H wag

Table A1.6 Representative Raman bands appearing in the SERS spectra of the *M. tuberculosis* laboratory-passaged H37Rv strains.

Raman Shift, cm^{-1}			Vibrational Mode Assignment
Wild Type	RvΔsigC	RvΔCcomp	
1592	1586	1592	C=O, alkyl ketone; (C=C); C=O-β conjugated; v_s(C=O) carboxylic acid
1564	1564	1564	C=C
	1491	1497	CH_3 antisym. bend
1457	1460	1458	δ(C-H_2) sci.
1392	1387	1391	C-O-H bend
1314	1315	1319	C-O-H bend; -$(CH_2)_n$- in-phase twist
1280	1280	1275	C-O-H bend
1241	1244	1241	C-O-H bend
1163	1157	1158	
1130	1130	1130	C-C skel. str in Alkane
1080	1076	1075	v_{as}(COC)
1008	1005	1008	v(C-C-O) out-of-phase stretch of primary alcohol
924			v(CHR_2) C-C skel. str where R≠CH_3
		869	v(C-C-O) in-phase stretch of primary alcohol, v_S(C-O-C); C-C skel. str in alkane
852	856	857	v(C-C-O) in-phase stretch of primary alcohol; C-C skel. str in alkane
	816	786	v(C-C-O) in-phase stretch of primary alcohol; v(CHR_2) C-C skel. str where R≠CH_3
713	714		CH_2 in-phase rock; v(CHR_2) C-C skel. str where R≠CH_3
707	713	707	cis dialkyl -CH wag

Table A1.7 Quantitative statistics calculated from the PLS-DA model developed from the SERS spectra[a] of the three NTM species *M. avium*, *M. bovis* BCG, and *M. smegmatis*.

Modeled Class[b]	*M. avium*	*M. bovis* BCG	*M. smegmatis*
Sensitivity (CV)[c]	1.000	1.000	1.000
Specificity (CV)	1.000	1.000	1.000
Class. Error (CV)[d]	0.000	0.000	0.000
RMSECV[e]	0.104	0.141	0.140

[a]Thirty-six spectra used, 12 for each NTM species. Before calculation, spectra were pre-processed using Savitzky-Golay 1st derivatives, vector normalized, and mean-centered.

[b]Five latent variables, accounting for 94.86% of the captured variance, was used in this regression model.

[c]CV, cross-validation based on Venetian blinds method with 6 splits

[d]Class. Error, classification error after cross-validation

[e]RMSECV, root-mean square error after cross-validation

Table A1.8 Quantitative statistics calculated from the PLS-DA model developed from the SERS spectra[a] of the three *M. tuberculosis* Erdman strains.

Modeled Class[b]	Wild Type	EΔCsigC	EΔCcomp
Sensitivity (CV)[c]	1.000	1.000	1.000
Specificity (CV)	1.000	1.000	1.000
Class. Error (CV)[d]	0.000	0.000	0.000
RMSECV[e]	0.203	0.188	0.142

[a]Thirty-six total spectra used, 12 for each MTB Erdman strain. Before calculation, spectra were pre-processed using Savitzky-Golay 1st derivatives, vector normalization, and mean-centering.

[b]Two latent variables, accounting for 87.88% of the captured variance, was used in this regression model.

[c]CV, cross-validation based on Venetian blinds method with 6 splits

[d]Class. Error, classification error after cross-validation

[e]RMSECV, root-mean square error after cross-validation

Table A1.9 Quantitative statistics calculated from the PLS-DA model developed from the SERS spectra[a] of the three *M. tuberculosis* H37Rv strains.

Modeled Class[b]	Wild Type	RvΔCsigC	RvΔCcomp
Sensitivity (CV)[c]	1.000	1.000	0.833
Specificity (CV)	1.000	0.917	0.792
Class. Error (CV)[d]	0.000	0.042	0.188
RMSECV[e]	0.154	0.351	0.352

[a]Thirty-six total spectra used, 12 for each MTB H37Rv strain. Before calculation, spectra were pre-processed using Savitzky-Golay 1st derivatives, vector normalized, and mean-centered.

[b]Three latent variables, accounting for 72.59% of the captured variance, was used in this regression model.

[c]CV, cross-validation based on Venetian blinds method with 6 splits

[d]Class. Error, classification error after cross-validation

[e]RMSECV, root-mean square error after cross-validation