universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

"Raman Microscopy Investigations of Bacteria-imprinted Polymers"

verfasst von / submitted by

Marius Bittermann, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2018 / Vienna 2018

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 862
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Chemie
Betreut von / Supervisor:	Univ.-Prof. Mag. Dr. Peter Lieberzeit

Contents

List of Abbreviations

AFM	Atomic Force Microscopy
APTES	(3-Aminopropyl) triethoxysilane
B. cereus	*Bacillus cereus*
CCD	Charged Coupled Device
CR	Cosmic Ray
DSS	Disuccinimidyl suberate
DVB	Divinylbenzene
E. coli	*Escherichia coli*
kNN	k- Nearest Neighbor
L. lactis	*Lactococcus lactis*
LV	Latent Variable
MIP	Molecularly Imprinted Polymer
NIP	Non-Imprinted Polymer
PC	Principal Component
PCA	Principal Component Analysis
PLS	Partial Least Squares
DA	Discriminant Analysis
QCM	Quartz Crystal Microbalance
S. epidermidis	*Staphylococcus epidermidis*
TCA	True Component Analysis

1 Introduction

The detection of microorganisms and their rapid qualitative and quantitative characterization is of great importance in everyday life, due to their potential impact on human health. Contamination of water resources by pathogenic bacteria is a major issue, especially in developing countries with lack of access to clean water. In 2011, the United Nations mentioned the improvement of water quality as one of the eight Millennium Development Goals.[1] The WHO reports that over 2.6 billion people nowadays lack access to non-contaminated water, a fact that leads to about 2.2 million deaths annually.[2] Water-borne diseases, such as diarrhea or gastrointestinal illness (GI) are caused by various bacteria, viruses and protozoa.[3] To protect public health, the U.S. Environmental Protection Agency's National Primary Drinking Water Regulations contain the Maximum Contaminant Level (MCL) standard, that describes the highest level of contaminate allowable in drinking water.[4] The MCL level has been defined for several microorganisms, including Total Coliforms, such as the fecal coliform *Escherichia coli* (*E. coli*). Generally, *E. coli* function as a bio indicator for water contamination, since multiple studies have identified trends between indicator organisms in water and GI illness in humans.[5] *E. coli* contaminated drinking water violates MCL regulations. To reliably detect and characterize microorganisms, many methods are already fully established, such as polymerase chain reactions and flow cytometry.[6] Even though these methods are sensitive, they still consume a lot of time and require expertise for carrying out the measurements and interpreting the data. Especially in developing countries, there is a growing need in rapid and cheap analysis tools, which are also easy to handle. Sensor systems, which can reliably monitor bacteria on a very short timescale (seconds-minutes) would address the need for the fast analysis of pathogenic microorganisms. The Sensor group at the Department of Physical Chemistry of the University of Vienna researches Quartz Crystal Microbalance (QCM) Sensors for Rapid analysis.[7] QCM sensors rely on an efficient combination of biological recognition principles with artificial soft matter.[8] A technique called Molecular Imprinting provides the recognition of an analyte of interest. Molecularly Imprinted Polymers (MIPs) are highly selective, sensitive and robust

materials, which can be applied as receptor systems for bacteria, viruses or proteins.[9] MIPs mimic natural recognition patterns, since they maintain the binding sites and affinities that are complementary to those of the analyte.

The sensitive layers developed in the group of sensors and rapid analysis are based on different polyacrylamides/acrylates, polyurethanes and Divinylbenzene/Styrene copolymers. Synthesis, coating and subsequent analyte-imprinting of the MIPs can be difficult to reproduce due to a complex manufacturing process and therefore reliable characterization of the QCM surface proves to be valuable. Nevertheless, it can be a difficult task to characterize MIPs, since the sensitive layers on QCM surfaces are ideally only a few hundred nanometers thick. Among others, optical microscopy and Atomic Force Microscopy (AFM) are techniques of choice to characterize the QCM surface and assess the quality of the imprinting process. However, the imprinted structures on the MIP surfaces are difficult to analyze optically; AFM and white light images offer wide room for interpretation. Another novel technique to characterize surfaces is Raman Microscopy, in which a confocal Microscope is coupled to a Raman spectrometer.[10] Raman vibrational spectroscopy is especially useful when it comes to analytics of bacteria, because every single bacteria cell exhibits its unique Raman spectrum.[11] Nevertheless, the information of the surface obtained by Raman Microscopy should be interpreted by sophisticated data analysis techniques, because the Raman signal is very weak and a lot of spectra vary only minimally and cannot be separated by eye.[12] The main goal of this thesis has been to characterize bacteria (i.e. *E. coli*) molecularly imprinted surfaces via Raman Microscopy and to assist in the further development and optimization of sensor systems, that can be used to rapidly assess the quality of drinking water. Due to the inherent complexity of Raman spectra, multivariate data analysis techniques, such as Principle Component Analysis (PCA) and Partial Least Squares-Discriminant Analysis (PLS-DA) methods were applied to evaluate obtained data sets. Additional to *E. coli*, different bacteria species, namely *B. cereus*, *S. epidermidis* and *L. lactis*, where investigated via Raman Microscopy and multivariate data analysis in an attempt to separate them from each other by their unique spectral fingerprints.

2 Methodology

2.1. Mass-sensitive Sensors

A sensor is a system that converts chemical information, such as concentrations, into an analytically useful signal. An ideal sensor can operate continuously, has high selectivity and sensitivity, has a decent signal to noise ratio, is easy to calibrate and is compact. Figure 2.1 schematically shows a typical sensor setup.[13]

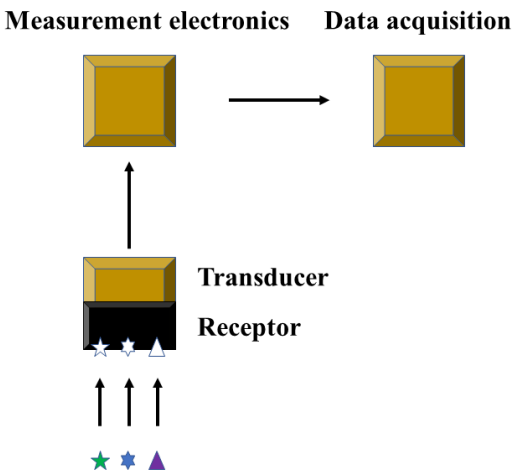

Figure 2.1: Block diagram of a sensor system; the sensitive layer (Receptor) is coupled to a transducer, which converts chemical information into measurable electrical signals. Data acquisition and evaluation extract and interpret valuable analytic information.

The chemically selective layer recognizes the analyte and acts as the Receptor. The transducer then transforms the change on the sensitive layer into an electrically processable signal.[14] Eventually, data acquisition helps to extract valuable chemical information. Sensors can be classified according to the working principle of transducers (i.e. electrochemical, optical, mass-sensitive etc.) Mass-sensitive transducers are based on the piezoelectric effect, which is illustrated in Figure 2.2.[15]

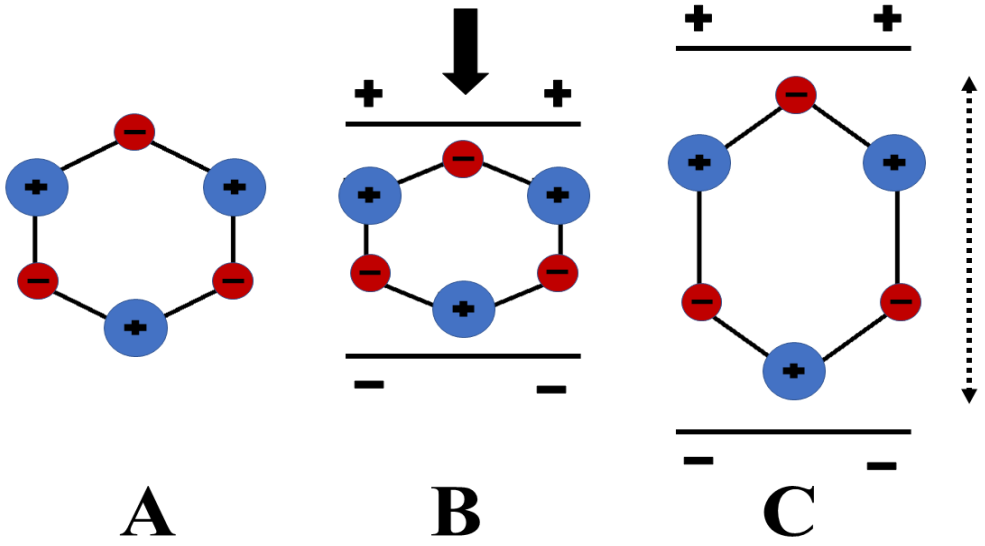

Figure 2.2: Schematic representation of the *piezoelectric* effect, adapted from reference 15; A: Quartz crystals consist of positively charged silicon atoms (blue circles) and negatively charged oxygen atoms (red circles). B: Application of pressure deforms a quartz crystal, which results in a surface charge of the electrodes. C: The inverse-*piezoelectric* effect means that the crystal becomes strained after the application of an electric field.

Certain crystals, such as quartz create an electric charge after being submitted to mechanical stress. Vice versa, quartz also responds to an applied electrical field by internal generation of mechanical strain. Generally, the *piezoelectric* and the *inverse-piezoelectric* effect are understood to occur in crystalline materials with no inversion symmetry along one of the crystallographic axes.[16] A very common application of the reverse piezoelectric effect is the Quartz Crystal Microbalance (QCM), a mass-sensitive device, that can detect quantities as low as 100 pg of analyte per square centimeter.[17] In 1959 M. Sauerbrey discovered the relation between the oscillation frequency of the crystal and the change of mass on the crystal's surface. This proportionality is shown in equation 2.1, which is the reduced form of the Sauerbrey equation.

$$\Delta f = -2f_0^2 \Delta m \cdot c \qquad (2.1)$$

The change of frequency Δf is proportionally related to the change of mass Δm. The proportionality factor c depends on the mechanical properties of the sensor material (i.e. density, shear modulus etc.) and can be considered constant. The resonant frequency f_0 of the quartz has high impact on the sensitivity of the QCM and determines the operating frequency.[18] Any change of mass will result in a change of frequency that can be measured. Therefore, QCMs provide qualitative, as well as quantitative information on the analyte of interest.

2.2. Molecularly Imprinted Polymers

In order to provide the QCM systems with selectivity, one has to modify and functionalize the surface of the sensor. Molecularly Imprinted Polymers act efficiently as sensitive layers and give rise to high selectivity. MIPs mimic the highly selective recognition entities of biomolecules, such as antibodies, that recognize antigens, or the interaction between biological receptors and respective target molecules.[19] As compared to their biological counterparts, imprinted polymers have higher physical robustness, strength, resistance to elevated temperatures and inertness to acids, bases and organic solvents. Additionally, they are relatively cheap can be stored over extended periods of time. The imprinting process is summarized in Figure 2.3.[20]

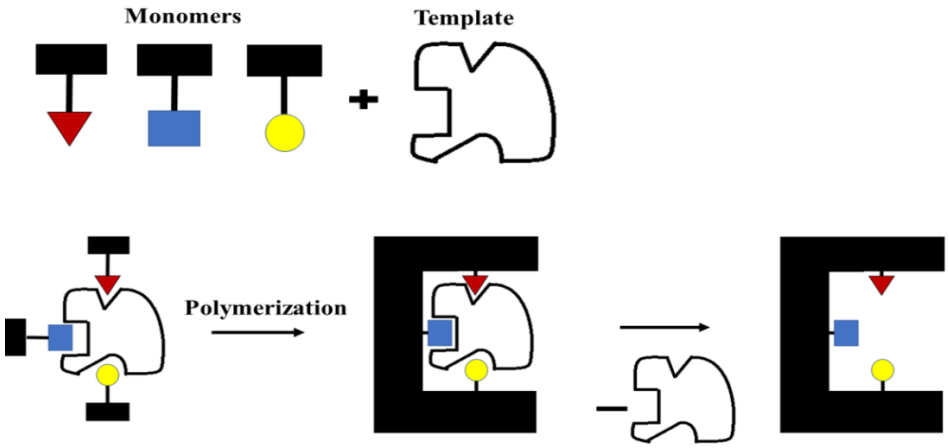

Figure 2.3: During molecular imprinting, a polymeric network is being established under the presence of a template molecule (i.e. the target molecule/analyte). After polymerization, the template is being washed out of the system, thus leaving cavities complementary to the shape of the target molecule behind. The respective analyte now selectively fits into these cavities; adapted from reference 20.

Polymerization occurs in the presence of the target molecule (template), which is incorporated into the polymer matrix.[21] Within this process, monomers spatially surround the template and their position is fixed by cross-linking via copolymerization. The resulting polymer-template complex consists of many microcavities. Specific interactions between template and polymer determine the selectivity of the system. After full polymerization, the templates are removed by a washing step and leave behind a polymer network with cavities containing binding sites to incorporate the target molecule. The MIP now selectively binds the analyte upon contact. Due to the weak interaction between

analyte and polymer the binding process is reversible. Nowadays, MIPs can be synthesized to target nearly any analyte, from inorganic compounds[22] to large biopolymers, such as proteins[23] and cells.[24] In contrast to their biological counterparts, molecularly imprinted polymers are arranged in chaotic networks containing up to millions of binding sites, depending on size and target molecule. A major problem with MIPs is homogeneity, which can barely be achieved and controlled, due to the complexity of the systems. Heterogenic distribution of the cavities within the system leads to underwhelming signal to noise ratios, as the binding sites are not evenly distributed on the sensor surface.[25] Sensitivity of the sensor is also affected by the thickness of the layer on the transducer. Generally, a thinner layer will lead to a better signal-to-noise ratio. MIPs can be produced in different shapes and states, wherefore nanogels, nanospheres and thin films would be typical examples.[26] The shape and the function of the MIP depend on the synthesis of the polymer, as well as on the imprinting process. Several techniques are available for the preparation of MIPs, such as bulk polymerization, electro polymerization, suspension polymerization, emulsion polymerization and precipitation polymerization.[27] The choice of the optimal polymer for the development of MIPs can be a challenging task and depends on the choice of the imprinting technique. The quality of MIPs can be estimated by considering the imprinting factor, which is a measure of the strength of interaction of the imprinted polymer towards the template molecule.[28] The imprinting factor IF is obtained from the ratio of retention factor k' of MIP to the non-imprinted (NIP) side.

$$IF = \frac{k'_{MIP}}{k'_{NIP}} \tag{2.2}$$

High performance imprinted polymers interact strongly with the template molecule, thus lead to higher chromatographic retention factor and higher IF. It was demonstrated that polymers based on divinylbenzene (DVB) lead to relatively high IF of 1.85.[29] In this thesis, DVB monomers were used as crosslinking agents in a styrene matrix to produce *E. coli* imprinted polymer networks.

2.2.2 Imprinting Techniques

Two of the most common imprinting techniques are bulk and surface imprinting. In bulk imprinting, a template is wholly imprinted in the matrix.[30] The process starts by dissolving template and monomers in a suitable solvent and subsequent initiation of polymerization. After polymerization, templates are removed by washing so that a "molecular memory" remains within the polymer network.[21a] Bulk or sedimentation imprinting is usually used for smaller molecules, since theoretically adsorption and release of the template are usually faster and fully reversibly in contrast to other techniques. In case of larger structures, such as proteins, cells, or microorganisms, the method has some drawbacks: First and foremost, larger structures will leave larger cavities, which can cause cross-selectivity due to enclosure of smaller molecules (i.e. polypeptides, amino acids).[31] Also, maintaining the conformational stability of larger structures during the imprinting process is a big challenge. Furthermore, some template species can be embedded too deeply in the matrices, leading to much longer response times, drift problems and poor regeneration.[32] In worst case scenarios, the templates cannot be removed at all, because they are stuck in the bulk of the polymer network. To address common issues with bulk imprinting techniques, other methods, such as surface imprinting, have been developed. In surface imprinting, high affinity recognition sites are formed at the surface of a substrate.[33] A major advantage would be that the binding sites are more easily accessible for the polymer matrix and the template/polymer interactions are not limited by diffusion only.[34] Due to the fact that the template species are concentrated on surfaces, disadvantages encountered with bulk imprinting, such as poor reversibility and longer response times can be avoided. Nevertheless, surface imprinting still causes problems. First of all, the technique requires more effort as the templates have to be immobilized on substrates before the imprinting step. Secondly, in surface imprinting sensitivity will be lower, due to a lower number of potential binding sites. A popular surface-based technique, employed in this thesis is stamp imprinting, in which the analyte is immobilized on a glass slide prior to imprinting.[35] Immobilization can be carried out either by binding the target molecules covalently or noncovalently to the surface of the substrate. The stamp is then pressed onto the pre-polymer and removed after full polymerization.

Covalently bound stamp imprinting usually does not require additional washing after polymerization, due to strong attractive forces between template and matrix. Both, bulk imprinting and surface imprinting are illustrated in Figure 2.4 and Figure 2.5, respectively.

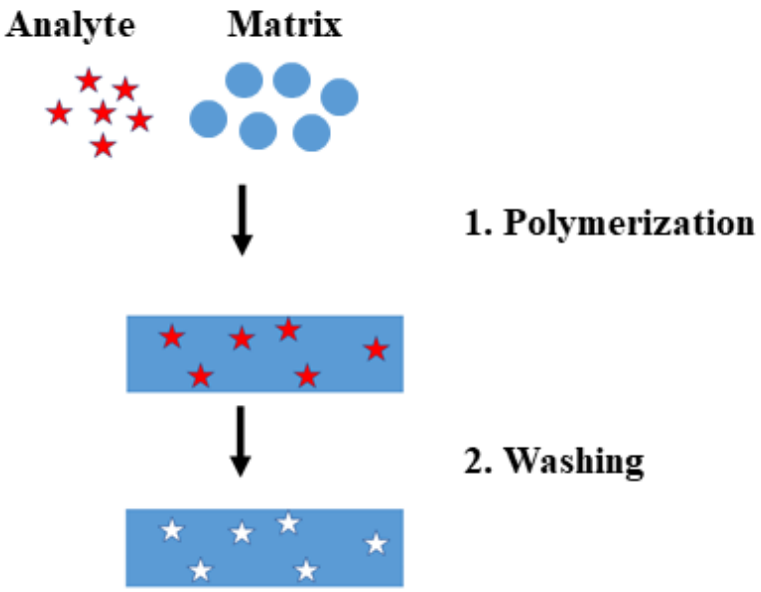

Figure 2.4: In bulk imprinting, matrix and template molecule are being polymerized concertedly.

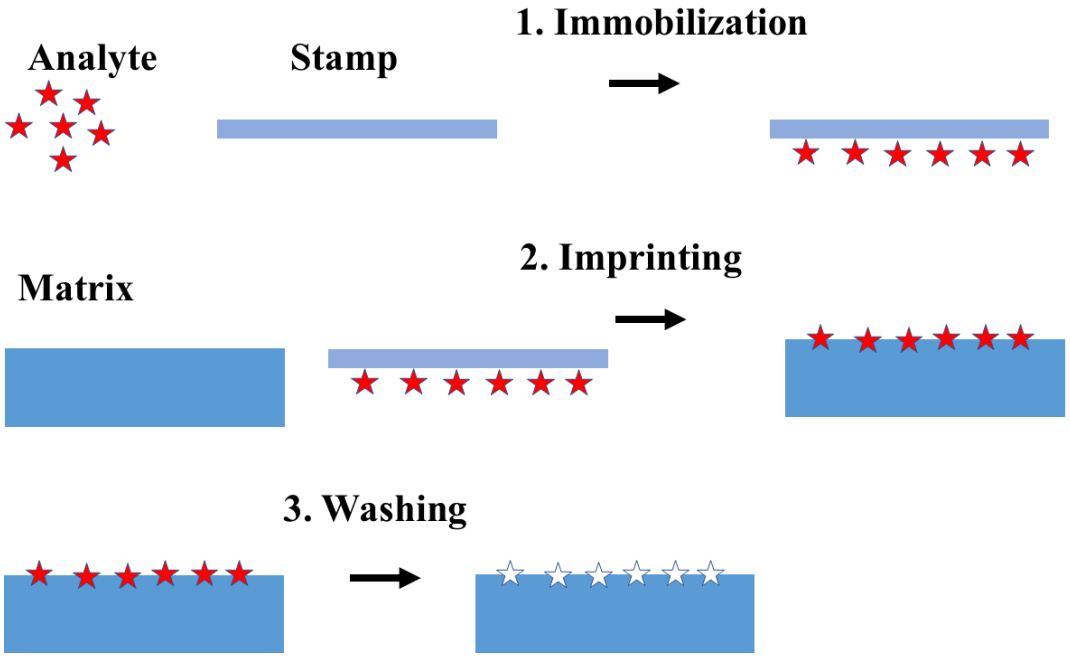

Figure 2.5: In stamp imprinting, the target molecule is immobilized on a stamp prior to the imprinting process. Removal of the templates by washing leaves cavities, which are specific to the analyte of interest.

2.2.1 MIP Analyte: Escherichia Coli

Escherichia Coli (*E. coli*) is regarded the most intensively studied organism of the planet.[36] In 1884, during a study of infant gut microbes and their role in digestion and disease, the German biologist Theodor Escherich discovered the fast-growing bacterium.[37] Basically, sturdy and non-pathogenic *E. coli* strains can be extracted from any human being. Due to this trait, and the fact that the microorganism is very easy to work with and grows rapidly in many different nutrients, *E. coli* underwent a tremendous rise in microbiology and can now be considered the most popular model organism. The rod-shaped *E. coli* cells have a length of about 2-6 μm and a diameter of 1-1.15 μm. White light pictures of E. coli cells can be seen in Section 4. The gram-negative bacteria are excellent templates for biosensors, since they have been studied as model organisms for the detection of contamination since the 1980s.[5b] Typical biochemical techniques, such as PCR can reliably detect the presence of *E. coli* in water samples. However, these methods take usually a few days, since culturing microbial species can be arduous and time-consuming. QCM-based techniques could sense bacteria on a much shorter and feasible time-scale, while maintaining decent sensitivity and reproducibility.[38]

2.3 MIP characterization

In general, due to their insolubility, MIPs are difficult to characterize. For surface investigation light microscopy and atomic force microscopy (AFM) can be applied.[39] To obtain qualitative spectral information about the systems, vibrational spectroscopy, such as Raman spectroscopy can be used to extract more sophisticated information. In this thesis, Raman microscopy, a technique that combines confocal light microscopy with Raman spectroscopy, has been employed to study MIP surfaces.

2.3.1 Atomic Force Microscopy

Since its invention in 1986, atomic force microscopy has evolved to become one of the most important tools for imaging surface structures down to nanometer scale resolutions.[40] The atomic force microscope can be operated in air, as well as in liquid phase, so that samples can be investigated in their natural environment.[41] Generally,

atomic force microscopes consist of four major parts: a cantilever usually made of silicon nitride, a piezo-scanner that drives the cantilever, a laser diode and a position-sensitive detector (Figure 2.6).

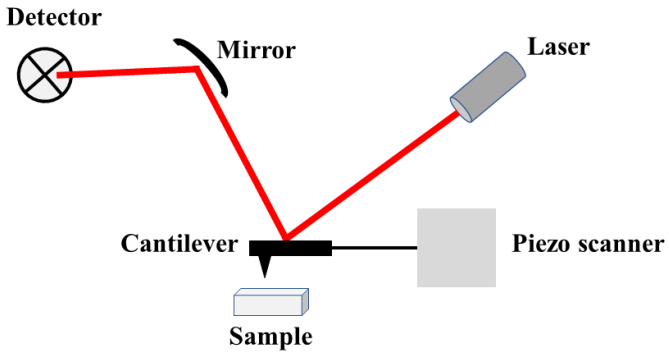

Figure 2.6: AFM working principle; the piezo scanning element steers the cantilever over the sample surface. A laser beam, emitted by a diode laser, registers displacement of the cantilever upon scanning the sample. In a feedback loop system, the laser position on the photodetector is used to track the surface and extract topographical information.

The cantilever tip scans over the surface; Interactions between the AFM tip and the surface features cause displacement of the cantilever, which is measured by detecting the deflection of a weak laser beam. Applying this method, topographical images of surfaces can be obtained with close to no sample preparation. Atomic force microscopes can be used in different modes, of which contact mode and non-contact mode are the most popular ones. In contact mode, force is applied by the AFM to the sample and the cantilever tip bends to react to changes in sample topography. The resulting friction between tip and sample can cause sample damage and therefore, the force applied must be chosen carefully. Non-contact mode almost guarantees non-damaging surface analysis by oscillation of the cantilever at ~5-15 nm above the sample surface. In contrast to contact mode, non-contact AFM utilizes weak attractive interactions (i.e. van der Waals forces, Coloumb and dipole interactions) between the lever and the sample.[42] These forces, in turn, induce alterations in the resonant behavior of the oscillating cantilever. Eventually, frequency shifts, phase shifts or amplitude changes are used to generate topographical images.

2.3.2 Confocal Microscopy

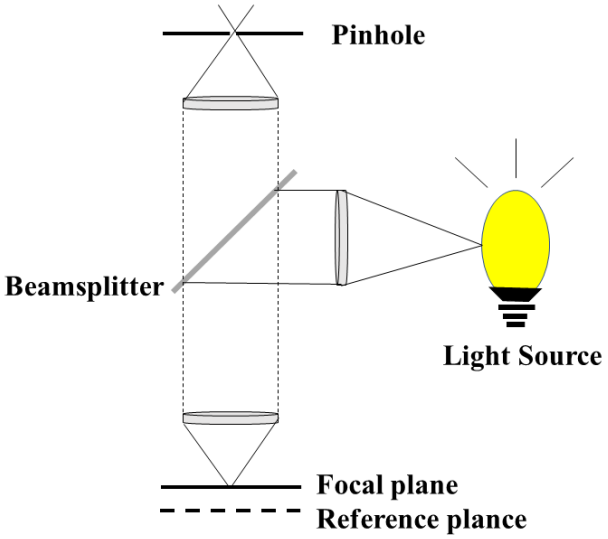

Figure 2.7: Confocal microscope working principle; the point-like light source is focused on the sample after passing a lens and a beam splitter. Only light on the focal plane passes the pinhole, which leads to the detector. Adapted from reference 43.

In confocal microscopy, a lens focuses light through a beam splitter onto a sample.[43] The image spot is then directed through a pinhole in front of a detector (Fig 2.7). Only rays that come directly from the focal plane pass the pinhole. This technique allows obtaining resolution in the z-direction, hence giving rise to 3D images of the sample. Due to the fact that only light from the focal plane hits the detector, higher image contrast than in conventional optical microscopy can be achieved. The lateral resolution of the microscope can be increased by up to a factor of $\sqrt{2}$ by setting an adequate pinhole diameter. According to Ernst Abbe the resolution of optical imaging instruments is fundamentally limited by the diffraction of light. Hence, an optical microscope cannot differentiate between two objects located closer to each other than $\frac{\lambda}{2N_a}$, of which λ is the wavelength of the incident light and N_a the numerical aperture of the imaging lense.

2.3.3 Raman Spectroscopy

The Raman Effect was first reported by C. V. Raman and describes inelastic scattering of a photon by molecules, which are excited to higher vibrational energy levels.[44] In light scattering, the primary incident electric field influences the charge distribution in the target molecules and thereby induces a dipole moment. The sum of the induced dipole moments in the affected molecules act as macroscopic polarization, which in turn represents the source of a secondary electric field emitted by the molecules.[43] This secondary field represents the scattered light. The proportionality constant between the induced molecular dipole moment μ and the external electric field E is the polarizability α,

$$\mu = \alpha \mathrm{E} \tag{2.3}$$

which describes the magnitude of the incident field's capability of influencing the electronic charge distribution of the sample. In essence, the incident field disturbs the equilibrium configuration of the target molecule. The real part of the electric field E can be expressed by its vector amplitude E_0 and its oscillation frequency w_0

$$\mathrm{E} = \mathrm{E}_0 \cos(w_0 t) \tag{2.4}$$

Nevertheless, the polarizability α is not a static quantity, as it will be modified by nuclear motion in a fashion that minimizes the energy of the system. Mathematically, polarizability α can be expanded into a Taylor series around the equilibrium nuclear geometry Q_0, where Q represents all normal modes q:

$$\alpha = \alpha(\mathrm{Q}) = \alpha_0 + \sum_{q=1}^{N} [(\frac{\partial \alpha}{\partial q})_{q_0} \, q \, + \frac{1}{2} (\frac{\partial^2 \alpha}{\partial q \partial q'})_{q_0 q_0'} \, q \, q' + O(q^3)] \tag{2.5}$$

Vibrational modes with a characteristic frequency w_q along each normal coordinate q can be excited,

$$q = q_0 \, \cos(w_q t) \tag{2.6}$$

Combination of equation 2.3, 2.4 and 2.5 yields

$$\mu(t) = [\alpha_0 + (\frac{\partial \alpha}{\partial q})_{q_0} q_0 \cos(w_q t)] \, \mathrm{E}_0 \cos(w_0 t) \tag{2.7}$$

For equation 2.7 the Taylor expansion in 2.5 was truncated after the first term. Eventually, the trigonometric formula for the product of the two cosine functions in equation 2.7 can be used to derive equation 2.8:

$$\begin{aligned} \mu(t) = \overbrace{\alpha_0 \mathrm{E}_0 \cos(w_0 t)}^{Rayleigh} + \overbrace{\frac{1}{2}(\frac{\partial \alpha}{\partial q})_{q_0} q_0 \mathrm{E}_0 \cos[(w_0 - w_q)t]}^{Stokes} \\ + \overbrace{\frac{1}{2}(\frac{\partial \alpha}{\partial q})_{q_0} q_0 \mathrm{E}_0 \cos[(w_0 + w_q)t]}^{anti-Stokes} \end{aligned} \tag{2.8}$$

Equation 2.8 describes the fundamental core of Raman scattering. The time-dependent induced dipole-moment consists of three distinct frequency terms. The first term represents *Rayleigh* scattering, in which the scattered photon oscillates at the same frequency as the incident photon. This elastic scattering does not provide any valuable analytical information. The second term in equation 2.8 accounts for elastic scattering, which describes the *Stokes* shift, caused by red-shifted radiation emitted by the induced dipole moment of the molecule. *Stokes* shift depends on the vibrational frequency w_q, thus giving valuable information about the oscillating normal modes of the target molecules. Usually, Raman spectra show *Stokes* scattering, because its intensity is higher than the intensity of the identical *anti-Stokes* scattering, described in the third term of equation 2.8. Classical *Rayleigh* scattering depends on equilibrium polarizability of the molecule. Mathematically, inelastic scattering stems from the non-zero derivative of the electronic polarizability at the equilibrium geometry along the qth normal coordinate, ($(\frac{\partial \alpha}{\partial q})_{q_0} \neq 0$). This relation leads to selection rules for Raman scattering, which can be used to distinguish the technique from standard infrared spectroscopy. Symmetric normal modes do not contain a permanent dipole moment; hence the derivative of the dipole moment is zero and the vibration is IR-inactive. One can predict Raman and IR activity simply by checking, whether the electron charge distribution (i.e. electric polarizability) changes with minimal nuclear displacement in any direction.

Vice versa, if the dipole moment changes with respect to infinitesimal variation along the normal coordinate, $(\frac{\partial \mu}{\partial q})_{q_0} \neq 0$, the vibration will be IR-active. The complementarity of Raman and IR Spectroscopy is known as the mutual exclusion principle.

2.3.3.1 Raman Spectrometer Parameters – Laser wavelength, power and aperture

Due to the issue of fluorescence, Raman spectrometers are usually available with more than one excitation wavelength. The intensity of the Raman signal scales with the fourth power of the frequency υ of the incident light

$$Intensity_{Raman} \propto \upsilon^4 \tag{2.9}$$

For this reason, higher frequencies and lower wavelengths are desirable for Raman experiments. However, the lower the wavelength, the higher is the probability of encountering fluorescence, the most prominent problem of Raman spectroscopy. Raman intensity is very low, since only one photon in $10^6 - 10^9$ photons is scattered inelastically, while most of the incident photons contribute to Rayleigh scattering, which does not provide analytical information. Consequently, a photoluminescent background can easily overwhelm the whole Raman signal by generating noise in the order of the Raman signal alone. Fluorescent trace impurities suffice to make a Raman spectrum unfeasible to analyze. Choosing an appropriate wavelength for individual samples can drastically increase signal-to-noise ratio by avoiding photoluminescence. Unfortunately, this also lowers accessibility of the Raman technique, since determining the ideal wavelength requires experience.

Analogically to the wavelength, the Raman intensity is directly proportional to the power of the Raman laser exciting the sample. The higher the laser power, the higher the Raman signal. However, if laser power is too high, the sample could be burned and Raman spectroscopy loses its destruction-free characteristic. Laser power and wavelength have to be optimized according to the sample, a procedure, which cannot be generalized, but highly depends on expertise.

Instrument aperture is another key component in Raman spectroscopy, as it controls which fraction of Raman-scattered photons reaches the detector.[45] Apertures or pinholes are normally in the 10-100 μm range. An increase in pinhole size results in higher signal

strength but lower resolution. When coupled to a microscope, Raman systems often use pinhole apertures to enable confocal operations. Generally, large apertures can provide Raman spectra in cases where the overall Raman signal is very weak. Smaller apertures lead to higher spectral resolutions and should be chosen if the Raman signal is adequately high.

2.3.3.2 Raman microscopy

Raman microscopes were first developed in the 1990s, when companies combined Raman spectrometers with optical microscopes.[46] The microscope's task is to focus the excitation light to a small spot a few µm in diameter. The obvious advantage of a Raman microscope over a conventional Raman spectrometer is the fact that using a microscope objective instead of a simple lens will result in high collection efficiency for the Raman signal, due to the high numerical aperture of the objective. Raman microscopes are usually equipped with motor_driven positioning stages, which can be used to raster scan the sample surface in order to obtain respective spectra on every single sample spot in a x,y,z coordinate system. The Raman microscope used for this thesis also comes with a confocal detection setup. As explained in more detail in Section 2.3.3.1, fluorescence inherently lessens the quality of the Raman technique. In this regard, the confocal setup proves to be beneficial in Raman spectroscopy, because it limits the collection of photoluminescence to photons emitted from the focal plane. Generally speaking, amalgamation of confocal microscopy with Raman spectroscopy delivers competitive results in terms of resolution and background reduction. Figure 2.8 shows the Raman microscope used for all experiments.[47]

Figure 2.8: WITec Alpha 300R Raman microscope; a confocal microscope is coupled to a Raman spectrometer. The laser beam is directly focused through the microscope objective. The scattered light passes through the spectrometer and is detected by a cooled CCD camera.

Very high throughput of confocal Raman microscopes can be established by modern Charged Coupled Device (CCD) detectors.[43] CCD detectors are commercially available in various sizes, uncooled or cooled, with Peltier or liquid Nitrogen cooling. A CCD detector includes an array of light-sensitive Si-photodiodes, which are linked to a capacitor. Photons reaching the detector induce electron-hole pairs, which are subsequently separated by an electric field. This builds up electric charge, which is proportional to light intensity at the respective detector location. Eventually charges are converted into voltage by an array-like architecture of the CCD detector. The most important characteristic of a CCD device is Quantum Efficiency (QE), which describes the percentage of detected photons to total incoming photons. The QE determines the intensity of the detected signal in a Raman measurement, as well as the signal to noise ratio (S/N).

Basically, primary noise sources can be subdivided into shot noise, thermal noise and readout noise.[43] Ideally the noise of the spectra should be dominated by shot noise, which stems from the random arrival of photons. Thermal noise comes from thermally excited carries in the CCD detector that can be dramatically reduced by adequate cooling. Commonly, detectors in Raman microscopy are cooled to a temperature of -60°C, at which the typical background signal due to thermally excited electrons is merely 0.01

electrons/pixel/s. Thermal noise can be neglected up to integration or laser exposure times of a few minutes. Nevertheless, especially in Raman scans, integration times are usually in the range of ~1s, thus thermal noise is not a limiting factor. The most relevant source of noise in a Raman microscope is the readout noise of the detectors itself, which depends on the quality of the readout amplifier in the CCD detector. Readout noise is generated by the detector as the charge present in the pixel is transferred to the readout electronics. The S/N ratio for CCD detectors can formally be expressed as

$$\frac{S}{N} = \frac{Signal}{\sqrt{Signal + \sigma_R^2 + \sigma_T^2}} \tag{2.10}$$

where σ_R is readout noise and σ_T thermal noise. Generally speaking, the lower the integration time (i.e. in Raman scans), the more dominant readout noise becomes. In an ideal scenario, a CCD detector would have 100% QE, no thermal noise and no readout noise. In this case, shot noise would limit the quality of the spectrum. CCD detectors have the irritating property of being very sensitive to cosmic rays. Cosmic rays (CRs) are natural, random and unpredictable events, which can render spectra useless, especially during long exposure times. Fortunately, data preprocessing methods, such as filtering can sufficiently compensate for CRs.

2.4 Chemometrics

In the digital age, large datasets are increasingly common and often difficult to analyze. Chemometrics is the application of statistical and mathematical techniques to analytical data to enable maximum extraction of useful information.[48] Many techniques have been developed that help to interpret large datasets by drastically reducing their dimensionality, while preserving essential information of the data. Among the Chemometric methods applied in this thesis are Principal Component Analysis (PCA), True Component Analysis (TCA), and Partial Least Square-Discriminant Analysis (PLS-DA).

2.4.1 Principal Component Analysis

Principal Component Analysis is one of the oldest and most frequently used chemometric techniques. The idea is to reduce the dimensionality of the dataset, while maintaining as much statistical information as possible. Dimensionality reduction can be achieved by defining new "principal" components and project them into an artificial coordinate system. The new variables are then sorted according to their variance, of which the first principal component has the highest variance within the dataset. The earliest literature on PCA dates from Pearson and Hotelling, but only scientific computing made it computationally feasible on larger scales.[49]

Formally, the standard context for PCA involves a dataset X, which is then decomposed via linear combination[50]

$$X = t_1 p_1^T + t_2 p_2^T + \ldots + t_r p_r^T \tag{2.11}$$

where t are the *scores* vectors and p the *loadings* of the data matrix X. Scores and loadings are ordered by the amount of total variance they capture. Commonly, the PCA model is truncated after k components and the remaining small variance factors are consolidated into a residual matrix E:

$$X = t_1 p_1^T + t_2 p_2^T + \ldots + t_k p_k^T + E \tag{2.12}$$

In practice, equation 2.12 can be applied to various chemical systems. In spectroscopy, the spectral data, summarized in X can be decomposed in tp pairs, of which t can be replaced with intensities and p with wavelengths. Mathematically, PCA can be written as an eigenvalue problem, the process relies upon an eigenvector decomposition of the covariance matrix:

$$cov(X) = \frac{X^T X}{m - 1} \tag{2.13}$$

Equation 2.13 requires that the columns of X have been either mean-centered or auto-scaled (i.e. mean centered followed by dividing each column by its standard deviation). Mean-centering translates the axes of the coordinate system to the centroid of the data. This is essentially the point, where all principal components cross. Variance scaling or

auto-scaling is beneficial, if variables have different units, so that all variables are on equal basis. In PCA, the loadings p are eigenvectors of the covariance matrix,

$$cov(X)p_i = \lambda_i p_i \tag{2.14}$$

The *scores* t_i form an orthogonal set, $t_i^T t_j = 0 | i \neq j$, while the *loadings* p_i are orthonormal and $p_i^T p_j = 0 | i \neq j$ and $p_i^T p_j = 1 | i = j$. For any data matrix X,

$$Xp_i = t_i \tag{2.15}$$

and the *scores* are basically projections of X onto the *loadings* p. The tp pairs are then arranged according to the magnitude of λ_i, which are a measure of the total variance captured in the PCA model. The general objective of PCA modelling is to represent the data using far fewer variables than the original datasets contains, without significant loss of information. Graphically, the samples vary more along the axis, which is described by the principal component containing the highest variance. Variance can be regarded as a measure of information stored in the dataset. To assess the quality of the PCA model, a lack of fit statistic Q can be calculated by simply including the sum of squares of each row of the residual matrix E (equation 2.12):[50]

$$Q = \pm e_i e_i^T \tag{2.16}$$

In essence, the Q value is a measure of the difference between a sample and its projection into the k principal components retained in the model. High Q values for certain data points indicate that the respective sample does not fit into the calculated PCA model.

2.4.2 Partial Least Squares (PLS)

Partial Least Squares (PLS) regression is somehow related to PCA but includes several significant differences. In general, PLS attempts to find factors, which not only capture variance, but also correlation between predictor and predicted variables.[50]

PLS 1 is the most prominent PLS approach, in which two sets X and c are obtained as follows:[51]

$$X = T \cdot P + E \tag{2.17}$$

$$c = T \cdot q + f \tag{2.18}$$

The matrix q has analogies to a loadings vector, as calculated in the PCA method. In spectroscopy, the product $T \cdot P$ approximates spectral data, whereas $T \cdot q$ represents concentrations. E and f can be considered residuals. The common link between X and c is the *score* matrix T. Contrary to PCA, *loadings* in PLS are not orthonormal and *scores* and *loadings* are calculated for each compound in the dataset. *Scores* in PLS are usually labeled as Latent Variables (LV). Due to the fact that PLS is a compromise between X and c block regression, the size of successive component eigenvalues (i.e. the magnitude of the captured variance) does not decrease as each component is calculated. The aim in PLS is to develop a model, that correlates *scores* and *loadings* between X and c. By predicting the *scores* of c, the *scores* of X can be calculated iteratively. By definition, the first *scores* of the c and X block of the PLS model comprise maximal covariance. Regarding the preprocessing step, it is common to apply baseline correction prior to mean-centering and normalization of both X and c block before analysis.

Partial Least Squares-Discriminant Analysis (PLS-DA) is derived from PLS regression and involves forming a regression model between X and c.[52] However, as opposed to PLS, the c block in PLS-DA is not a set of continuous numbers, but of discrete numbers, assigned in groups or classes. Thus, PLS-DA identifies directions in the data set that discriminate classes directly. The principle is illustrated in Figure 2.9. In the case of bacteria classification, PLS-DA is the discrimination technique of choice, because the type of bacteria is already known and the goal is to find out, whether a discrimination based on spectral data is feasible. For this purpose, groups are assigned to the corresponding spectra of the respective microbe. The specification of the optimal number

of latent variables can be made based on cross-validation. In this procedure, the data set is divided into several equal size segments. A PLS-DA model is then built on all but one of the segments. This model is then used to estimate the data segment, which was previously left out. The differences between the predicted values and the actual values are presented as root-mean-square error of cross-validation (RMSECV).

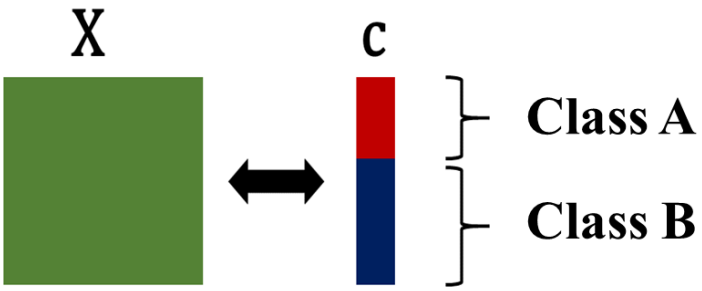

Figure 2.9: In PLS, the aim is to establish correlation between a data block X and data block c. In PLS-DA, c contains information about classes. Similar but not identical to PCA, X and c are transformed into *scores* and *loadings* vectors.

2.4.3 True Component Analysis (TCA) and "Raman TV"

Data acquisition and partly data evaluation were carried out with the software tool WITec Suite FIVE, that features the chemometrics tool True Component Analysis.[53] Unfortunately, the exact algorithm behind the software has not been made public yet by WITec and the software is still in a developing stage. Assumedly, TCA sums all spectra within a Raman image and calculates the spectral contribution of individual spots to the total spectrum. TCA proves to be very useful when it comes to the studies carried out in this thesis, because it enables to directly find distinct spectral components on the MIP surface, such as bacteria spectra and polymer spectra. The individual components can then be averaged and demixed, so that pure spectra can be extracted from confusing overlapping spectra. Nevertheless, interpretation of the TCA results can be elusive, as the exact algorithm is not known.

WITec Suite FIVE also offers a tool named Raman TV. In Raman TV, certain peaks are isolated via binomial average filtering and then recombined to produce false color pictures, containing information about intensity distributions of the respective region.

3 Experiments

All solvents and chemicals used for MIP preparation, including polymer synthesis, stamp imprinting and bacteria cultivation were purchased either from Panreac, VWR International GmBH, Merck Millipore, Thermo Scientific Fisher, or Sigma Aldrich in analytical or highest possible synthetic grade.

The Raman microscopy investigation of MIP was carried out with a WITec ALPHA300 R confocal Raman Microscope. Results were analyzed with WITec Suite FIVE and the MATLAB based chemometrics tool Solo+Mia.[53-54]

3.1 Bacteria Cultivation, Washing and Storage

E. coli was the main subject for the experiments carried out within the framework of this thesis. However, other bacteria species were cultivated and analyzed as well, so that the limits of Raman microscopy and Multivariate Data Analysis can be observed. The bacteria strains *Bacillus cereus* (*B. cereus*) ATCC® 11778™, *Staphylococcus epidermidis* (*S. epidermidis*) ATCC® 12228™ and *Escherichia coli* (*E. coli*) ATCC® 31303™ and *Lactococcus lactis* (*L. lactis*) were obtained from LGC Standards and cultivated in a common Luria-Bertani (LB) medium, which is a mixture of 1% (w/v) proteose peptone, 0.5% (w/v) sodium chloride, 0.5% (w/v) yeast extract and 0.1% (w/v) D(+)-glucose monohydrate as an energy source. In order to avoid contamination, all reagents and materials used in the cultivation process were initially autoclaved. The cells then grew overnight in Eppendorf tubes at 37°C. The next day, the cells were washed twice by centrifugation at 3500 rpm for 6.5 minutes, followed by resuspension in distilled water and further centrifugation. Upon washing, a part of the pellet was resuspended in a new LB medium for cultivation and the rest was either used for experiments or stored in the freezer at -20°C.

3.2 MIP Synthesis: Polymer Matrix

During the first months of the master's project, acrylamide-based polymers were used as MIPs. However, due to poor reproducibility of the synthesis protocol, a recipe from a previous master student, including Styrene-*co*-DVB as the polymer matrix, was used. Styrene-*co*-DVB not only turned out easier to synthesize in a reproducible manner, but also better suited for the imprinting step, due to lower surface roughness and higher homogeneity.

The Styrene-*co*-DVB was produced by mixing 0.5 mL styrene with 0.5 mL DVB and adding a catalytic amount of AIBN for initiation. The solution was left to polymerize at 70°C for ~30 minutes in a water bath. The reaction is illustrated in Figure 3.1. The mixture polymerizes into a gel-like phase, which was thereafter used for imprinting.

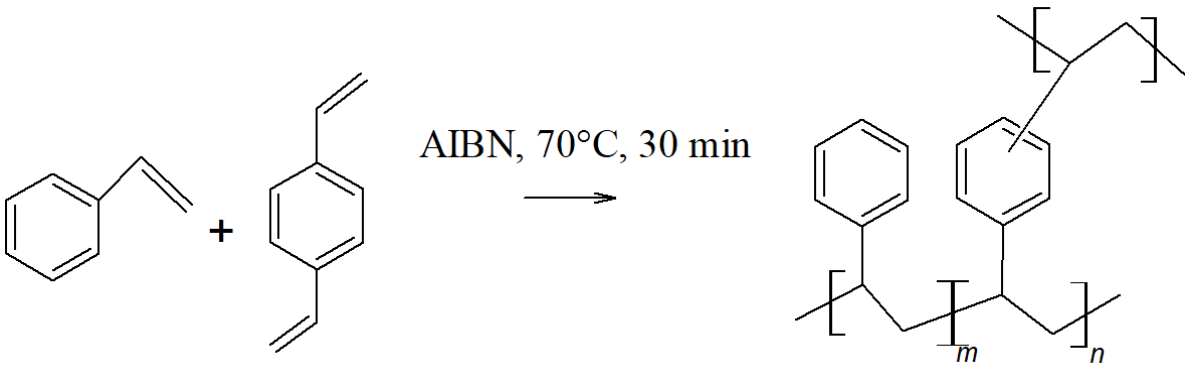

Figure 3.1: Reaction of styrene and DVB to Styrene-*co*-DVB; a catalytic amount of AIBN was used as initiator for the polymerization.

3.3 MIP Synthesis: Sensitive Layers

For the generation of sensitive layers, stamp imprinting (theoretical description in section 2.23) was used. For manufacturing the stamp, ~2x2 cm glass slides were functionalized with 3-(aminopropyl)triethoxysilane (APTES) in a first step. This could be achieved by dissolving 180 mg APTES in ~10mL acetone and pouring the resulting solution onto glass slides. After 2 minutes, the glass slides were washed with distilled water and thoroughly dried in the oven at 50°C. The dried, functionalized glass slides were then immersed into a solution of 50 mg disuccinimidyl suberate (DSS) in 10 mL DMSO and left to react for 1 hour. The slides were then washed carefully with PBS (25 mM, p.H 7.0)

and 400 µL diluted E. coli solution was spotted onto each glass slides and left to air-dry for ~1hour. Prior to imprinting, 50 µL pre-polymer was dropped onto clean glass slides. Usually, homogeneous distribution of pre-polymer on respective surfaces can be obtained via spin-coating. For stamp imprinting however, spin-coating turned out to have barely any impact on the homogeneity of the sensitive MIP layer and was skipped at all. The stamps were then lightly pressed onto the Styrene-*co*-DVB pre-polymer surface and left to polymerize in the oven at 50°C overnight. Eventually, the stamps were removed by dipping the polymer-stamp assembly into distilled water.

3.4 Atomic Force Microscopy

For surface analysis, accompanying Raman Microscopy, the AFM Nanoscope VIII from Bruker was used. Data acquisition and evaluation were carried out via the software NanoScope 9.0. The mechanically highly resistant cantilever for the AFM, SNL-10 from Bruker, is made of Silicon Nitride. All measurements were performed in contact mode at a scan rate of 2 Hz. AFM was mainly used to accompany Raman Microscopy in the task of acquiring topographical information about MIP surfaces.

3.5 Raman Microscopy and Multivariate Data Analysis

Raman microscopy investigations of MIPs was carried out with the WITec ALPHA300 R confocal Raman Microscope.[47] All samples were placed on a CaF_2 sample holder to ensure that no substrate background interfered with the sample spectra. CaF_2 shows only one characteristic peak at 321 cm^{-1} and is durable and insoluble. Before conducting the actual Raman measurements, images of the sample surface were taken to simplify navigation with the microscope and to obtain broad overview of sample topography. Four different types of diode lasers were eligible, with excitation wavelengths of 488, 532, 633 and 785 nm, respectively. Due to the unfortunate scaling of Raman intensity (equation 2.9), only the excitation laser at wavelength 532 nm was used, which proved to yield the best signal-to-noise ratio and could be controlled via the WITec software suite. The grating was kept at 600 lines/mm with a center wavelength of 596.999 and spectral center of 2046.558 rel. cm^{-1}. The grating setup achieves a spectral range from 0 rel. cm^{-1} up to 3753 rel. cm^{-1}, which covers all important anti-Stokes shifts of the analytes, however, only the range between 600 cm^{-1} and 3753 cm^{-1} was considered in data analysis, since it

contains the relevant spectral information on bacteria. The laser power was set to 8 mW for most of the experiments. Laser powers >10 mW often resulted in the degradation of the sample. Depending on the sample, both single spectra and Raman imaging scans were conducted. For single spectra, ~30 spectra were recorded on randomly chosen spots of the sample (100x magnification). The integration time for single spectra was set to 15 s, which can be regarded very long, as the WITec alpha300 R can perform ultra-fast Raman imaging with only 0.76 ms integration time per spectrum. For Raman image scans, the integration time was set in the ms range, due to the large number of data points. At first, all bacteria strains available in the lab, i.e. *Bacillus cereus* (*B. cereus*) ATCC® 11778™, *Staphylococcus epidermidis* (*S. epidermidis*) ATCC® 12228™, *Escherichia coli* (*E. coli*) ATCC® 31303™, and *L. Lactis* (*L. lactis*) were investigated separately with the aim to distinguish them by the fingerprint regions in their spectra. For this purpose, ~100 µL concentrated bacteria solution was spotted onto the CaF_2. Evaluating the bacteria concentration prior to Raman measurements was dispensable since there was no interest in obtaining quantitative information in all experiments. Following this experiment, the integration time and laser power were increased to 60 s and 10 mW, respectively, in an attempt to distinguish between the highly similar species *S. epidermidis* and *L. lactis*. Additionally, ~100 µL of the very distinct species *E. coli* and *L. lactis* were mixed and investigated to find out whether direct spectral separation within a bacteria mixture would be possible. For the Raman microscopy investigation of stamp-imprinted MIPs, the dry bacteria-imprinted Styrene-*co*-DVB sheet was directly placed onto the CaF_2 sample holder. All image scans mentioned in this thesis were carried out with 0.1 s integration time and 5 times as many pixels (points per line and lines per image) as the scan area to capitalize on the Raman microscope's lateral resolution of 200 nm. Deviant integration times are mentioned in section 4. Depending on the size of the scanned area, the Raman scans took about 30-90 minutes, whereas signal stabilization proved to have huge impact on scanning time. Signal stabilization is an invaluable setting, which diminishes thermal drift by adjustment of the focus via comparison with a reference point on the sample surface. Signal stabilization can be considered indispensable for Raman imaging. All experiments required very little sample preparation, a huge advantage and positive aspect of the Raman microscopy technique. To simulate QCM measurement conditions, the bacteria imprinted Styrene-*co*-DVB polymer sheet was immersed in concentrated *E. coli* solution prior to Raman microscopy investigation.

3.6 Data Preprocessing

Prior to PCA, PLS-DA and TCA data analysis, the data sets were preprocessed by applying baseline correction, normalization and auto-scaling. These preprocessing steps are mandatory, since they eliminate intensity discrepancies in the dataset, which can impact discriminant analysis and cause misleading results. Cosmic ray (CR) removal is another important preprocessing step, which can be performed via filter directly in the WITec Five Suite. For CR removal, the values for filter size and dynamic range were chosen to be 2 and 8, respectively.

Baseline correction was carried out via the Weighted Least Squares (WLS) technique. WLS is typically used in spectroscopic applications where the signal in some variables is caused by background only. The WLS algorithm uses theses variables as a reference to calculate how much background should be removed from nearby variables. This can be achieved by iteratively fitting a baseline to each spectrum and determining which variables are clearly above the signal and which are below. All points below the baseline are assumed to be more significant in fitting the baseline to the spectrum. Commonly, WLS uses low order polynomials to approximate the baseline. In this thesis, the polynomial used was set to be of order 2.

Normalization is another standard preprocessing technique, in which all intensities are set to the value 1. For discrimination models, such as PLS-DA, normalization is valuable if the relationship between variables (i.e. spectral regions), and not the magnitude of the response, is the most important aspect of the identification of distinct patterns. Typically, data should be normalized after baseline correction and before scaling the dataset. The normalization preprocessing method, applied in this thesis, normalizes to the sum of the absolute value of all variables and returns a vector with unit area 1. Figure 3.2 shows the effect of baseline correction, coupled with normalization, applied on a bacteria dataset.

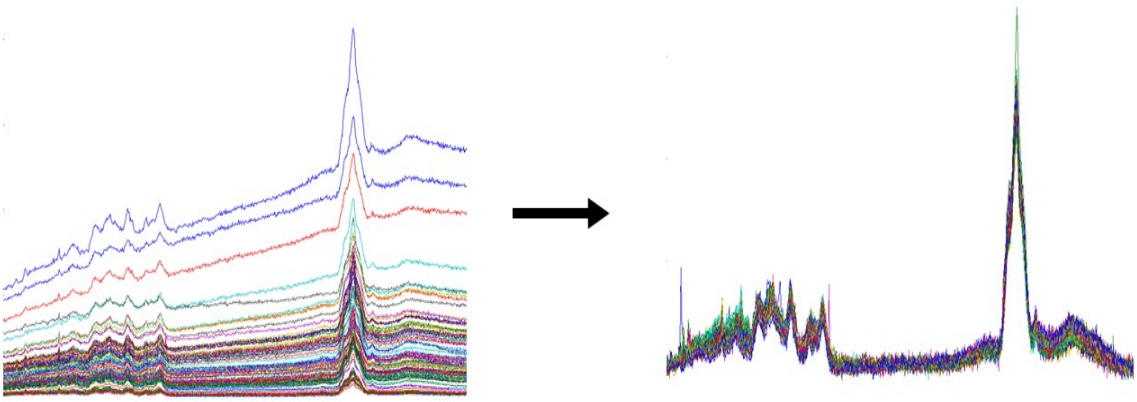

Figure 3.2: All spectra must be preprocessed prior to PCA or PLS-DA, otherwise, phenomena such as fluorescence or lack of focus of the confocal microscope will have an impact on the result of the discrimination analysis. Above, the dataset was baseline corrected and normalized.

Autoscaling is a common preprocessing method, which uses mean-centering followed by dividing each column (variable) by the standard deviation of the column. The aim of this variable scaling approach is to neutralize the dependence of variable importance from its magnitude. The scaling causes each variable's information content to be of equal magnitude. For the experiments in this thesis, autoscaling is particularly important as the fingerprint region of the bacteria spectra has a very low signal, but is equally, if not more important than the C-H region around 3000 cm^{-1}. All Multivariate Data Analysis approaches are discussed in detail in section 4.

4 Results and Discussion

4.1 Bacteria on CaF_2

The first part of the result section deals with different bacteria species on CaF_2 as a substrate. CaF_2 is widely regarded as the sample holder of choice when it comes to Raman spectroscopy, as it only exhibits 1 characteristic peak at 321 cm^{-1} with minimal fluorescence. The aim of the conducted experiments in section 4.1 is to find out, whether Raman microscopy coupled with statistical data evaluation can be used as a rapid analysis toll for the qualitative investigation of bacteria cells.

4.1.1 *E. coli*, *B. cereus*, *S. epidermidis* and *L. lactis*

The Raman spectra for the bacteria *E. coli*, *B. cereus*, *S. epidermidis* and *L. lactis* were discriminated via PLS-DA. Optically, all bacteria look very similar and it is barely possible to tell the difference between the rod-shaped *E. coli* and *B. cereus* and the spherical *S. epidermidis* and *L. lactis* microbes (Figure 4.1), respectively. As depicted in Figure 4.2, even the averaged Raman spectra look quite similar, with merely minimal variation within the fingerprint region (~600-1800 cm^{-1}) and the C-H vibrations (~3000 cm^{-1}). The average or mean spectra were calculated based on 30 individual spectra of each class. Detailed evaluation of the actual spectral difference was not subject of this thesis.

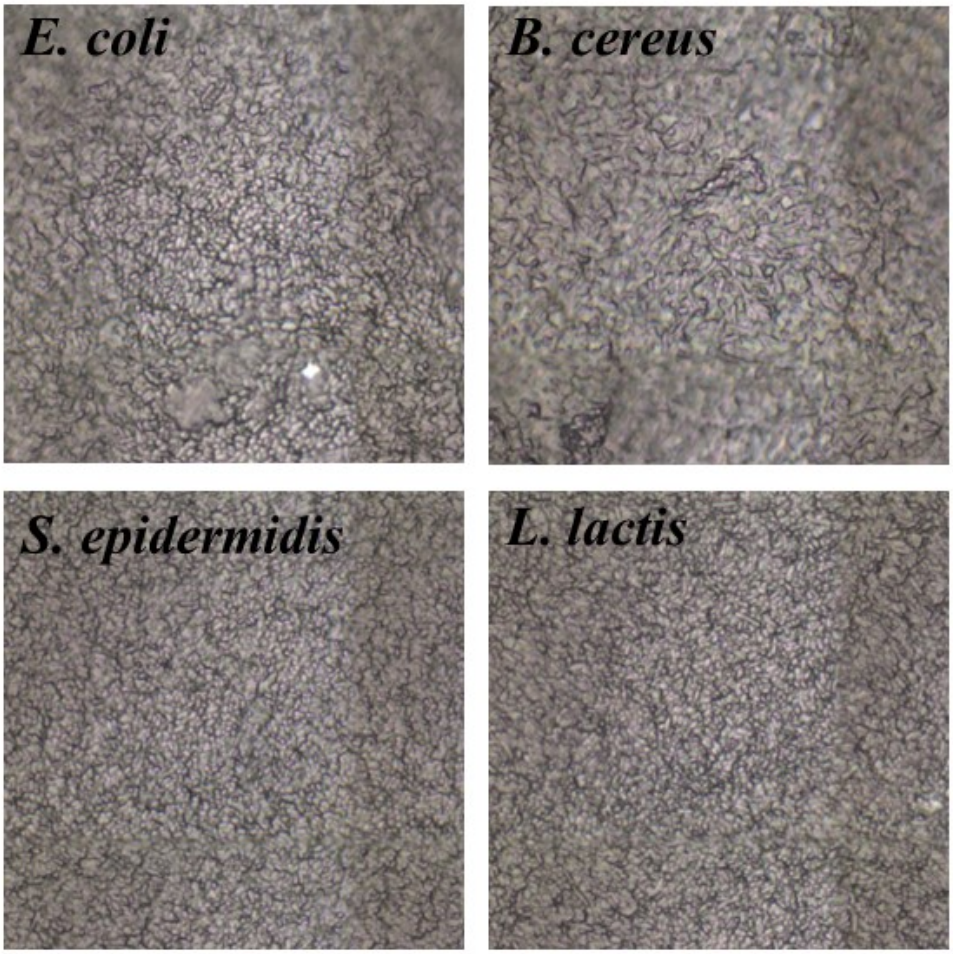

Figure 4.1: White light images of bacteria cells obtained with 100x objective. The rod-shaped *E. coli* and *B. cereus* and the spherical *S. epidermidis* and *L. lactis* can hardly be distinguished by eye.

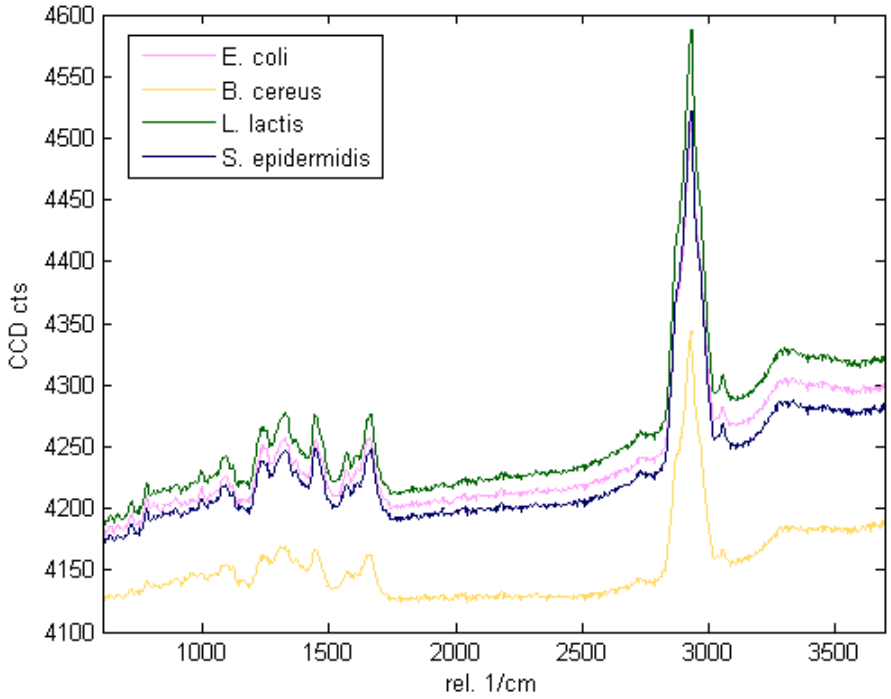

Figure 4.2: Raman Spectra of four different bacteria species; *S. epidermidis* exhibit an unusual high fluorescent background. The fingerprint region around 600-1800 cm^{-1} and the C-H vibrations around 3000 cm^{-1} are crucial for data separation with multivariate techniques.

The peaks in the fingerprint region show contributions from all major building blocks of cells, such as nucleic acids, carbohydrates, proteins and lipids. The large amount of information present in the Raman spectra is problematic when it comes to analysis. Around 3000 cm^{-1} the high frequency C-H vibrations can be observed. Especially the fingerprint region is difficult to analyze, since many peaks overlap. However, multivariate techniques allow for subtle spectral differences to be used to discriminate between different cell types.

In order to specify the optimal number of latent variables (LVs) for the PLS-DA model, the Root Mean Square Error Cross-Validation (RMSECV) was calculated and evaluated. Ideally, the LVs should level out after a certain number, but in some cases, it is not completely clear which number to choose.

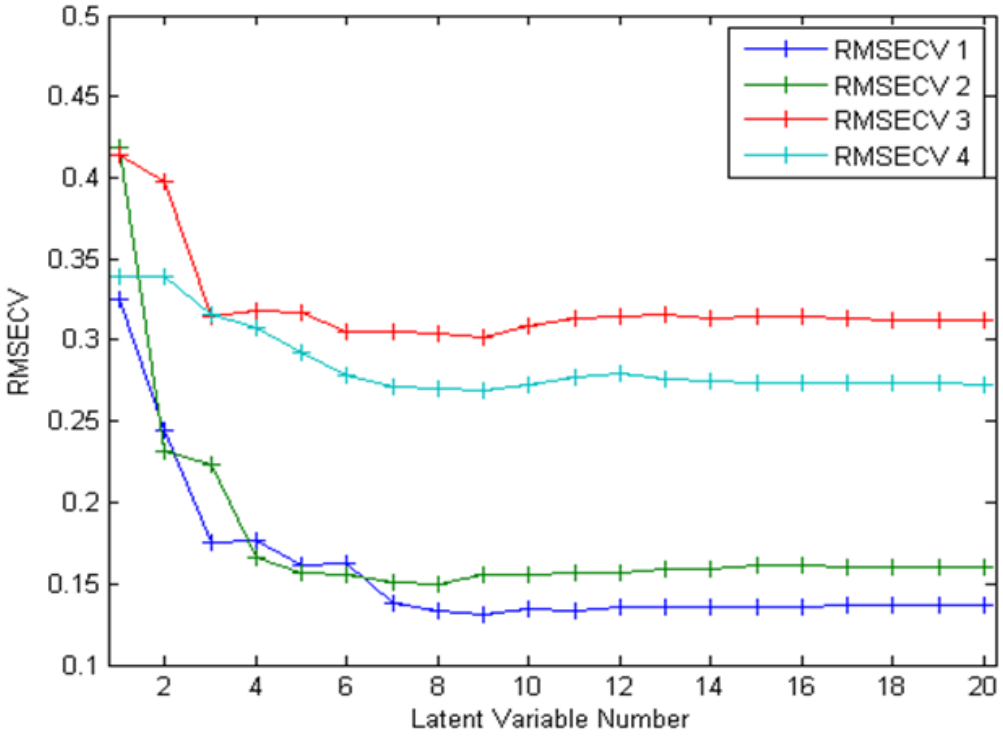

Figure 4.3: For Class 3, the RMSECV remains constant after 3 LVs. However, this is not the case for the remaining classes. Class 4 varies only slightly, considering its stable RMSECV, whereas class 1 and 2 take up to 7 and 4 LVs, respectively, until there is no significant change anymore.

As illustrated in Figure 4.3, all bacteria classes take a different number of LVs until their RMSECVs level out. For this purpose, 4 LVs were chosen as a compromise, because in average it takes 4 LVs until the RMSECV remains considerably constant. There are different opportunities to analyze the final PLS-DA model. A powerful tool is Y-Prediction, in which a threshold is calculated based on Bayes Theorem, which assists in estimating class affiliation of the samples. The threshold is predicted for every class and can be seen in Figure 4.5. All bacteria clusters correspond to their respective classes. The distance of the centroids of the clusters is a measure of the quality of separation. It can easily be noted, that *E. coli* and *B. cereus* are more distinct than *S. epidermidis* and *L. lactis,* because they do not overlap with other classes and their centroids are further away from the threshold line at 0.39.

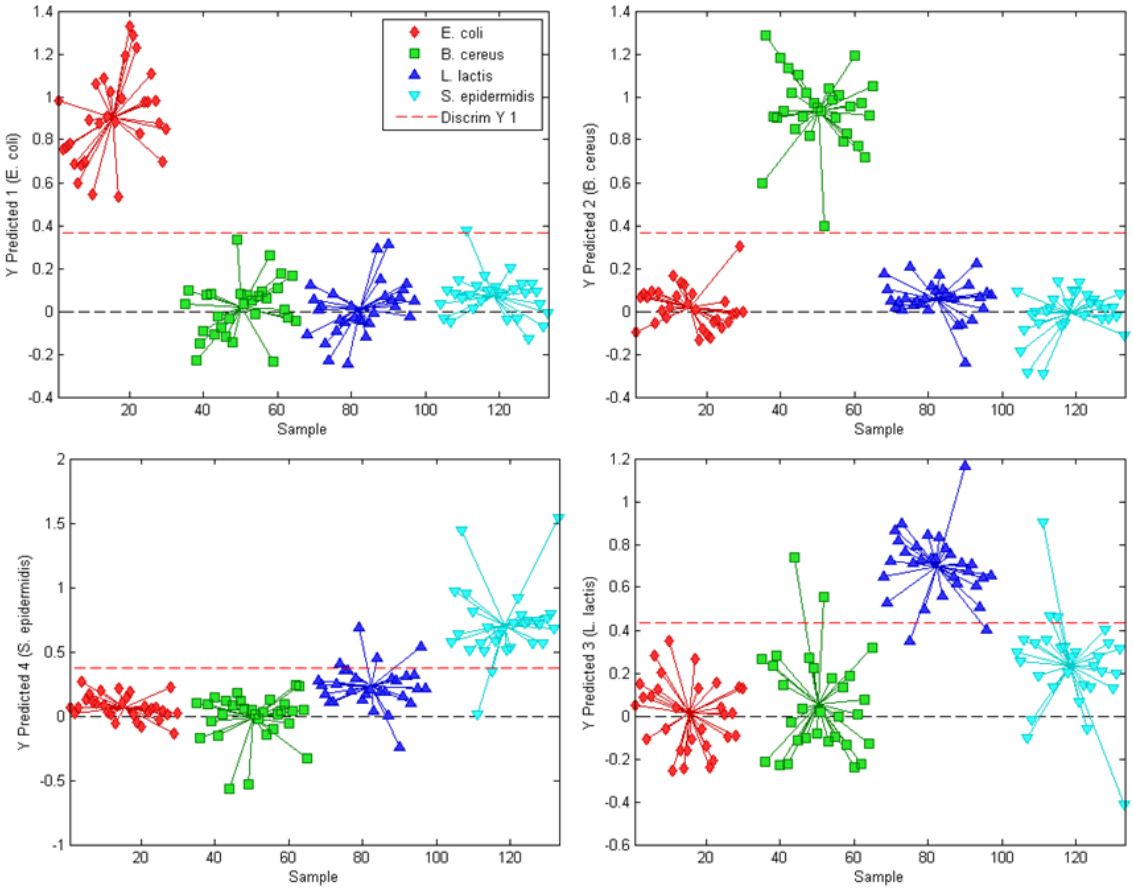

Figure 4.4: The Bayes Theorem is the fundament of Y-Prediction; the red dashed line represents a threshold, which separates the respective group from the remaining samples. *E. coli* and *B. cereus* are more distinct as *S. epidermidis* and *L. lactis*, since in average, the centroids of their clusters are further away from the threshold.

Another means to analyze PLS-DA results is to plot the sample numbers versus the 1. LV, which captures most variance of the system. Figure 4.5 illustrates, that *E. coli and B. cereus* cannot be separated considering only 1 LV. However, they differ from *L. lactis and S. epidermidis*, respectively, as their centroids lead to positive, instead of negative values on LV 1.

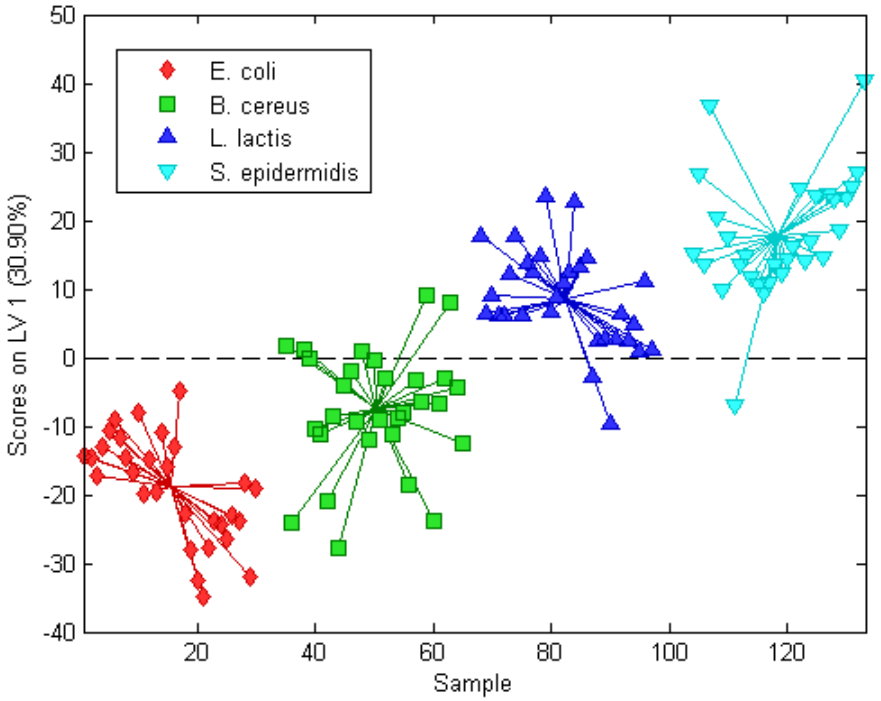

Figure 4.5: *E. coli and B. cereus* lead to negative values on LV 1, whereas *L. lactis and S. epidermidis* can be found in the positive range. 1 LV is not sufficient to separate all specimens from each other, since data of all bacteria overlap significantly.

Most of the time, considering only 1 dimension is not sufficient for decent separation of classes. In these cases, it can prove to be valuable to plot different LVs versus one another to obtain a distinct perspective on the system modelled. Figure 4.6 and Figure 4.7 show LV1 plotted against LV2 and a 3-dimensional LV1 versus LV2 versus LV3 diagram, respectively. Plotting LV 1 vs LV 2 in a 2-dimensional manner reveals, that *E. coli* and *B. cereus* can easily be separated from each other, whereas *L. lactis* and *S. epidermidis* still overlap dramatically. A 3-dimensional plot solves this issue, since including LV 3 sufficiently depicts all species of interest as belonging to individual classes.

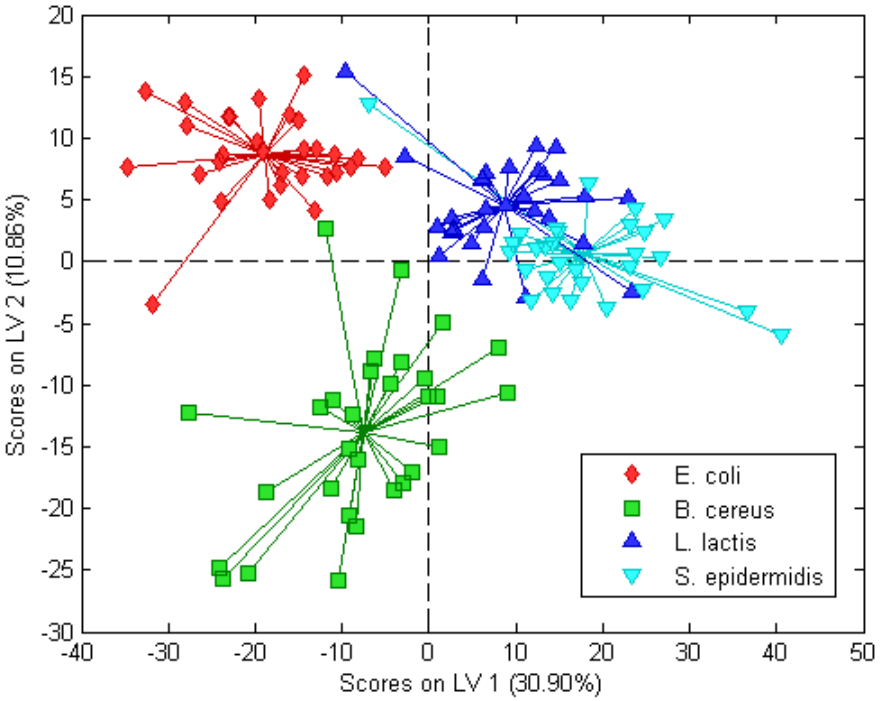

Figure 4.6: The inclusion of a 2nd LV leads to a decent separation of *E. coli and B. cereus* from *L. lactis* and S. epidermidis. However, the latter species can barely be distinguished from each other, since their centroids are very close to each other on the positive scale of the 2D LV diagram.

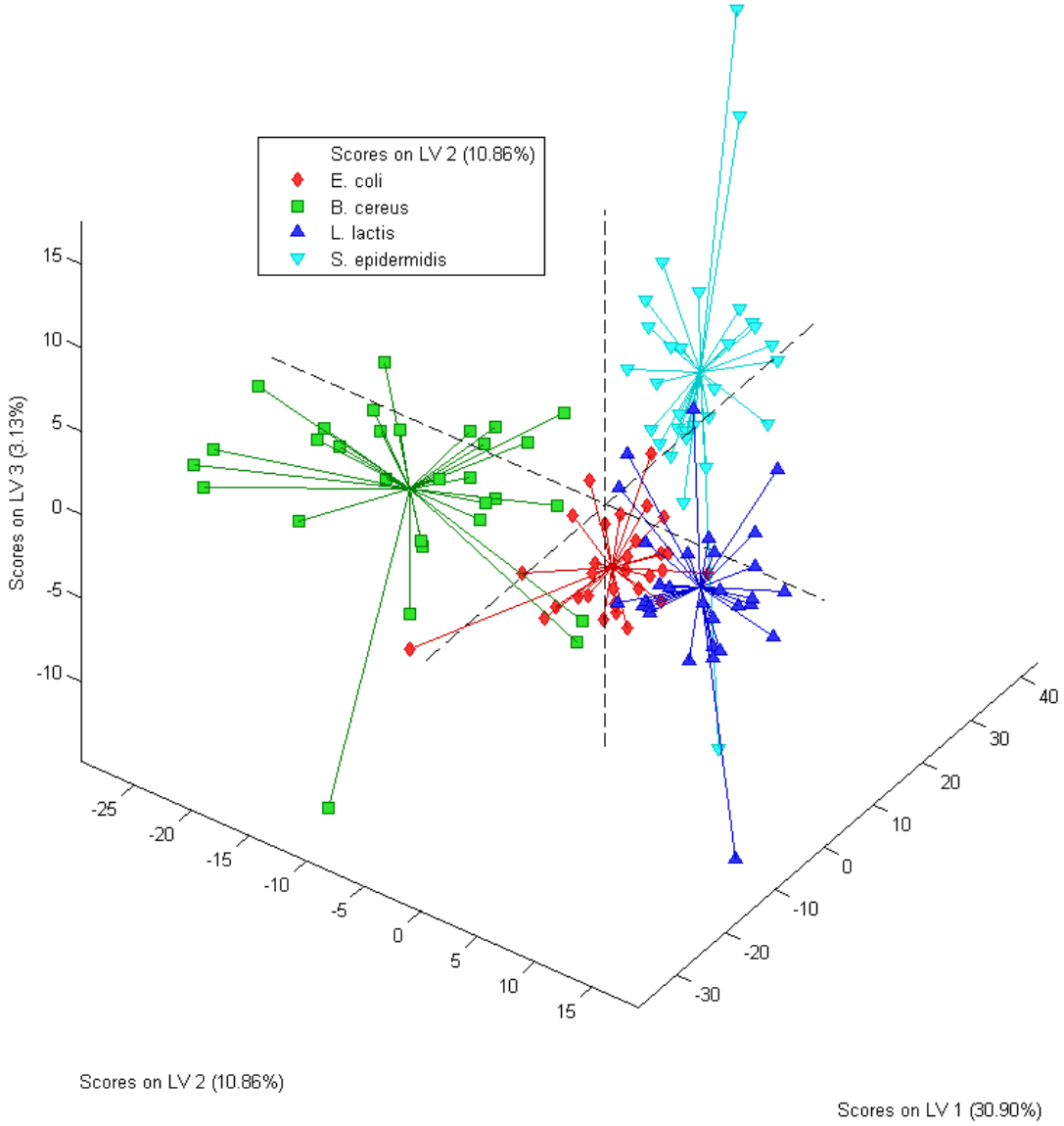

Figure 4.7: The 3D LV diagram is an efficient means to separate all investigate species. Nevertheless, the more LV are considered, the more complicated a PLS-DA model gets.

Another attempt to separate the very similar *L. lactis* and *S. epidermidis* species from each other was made by increasing the integration time of the single spectra to 60 s and the laser power to 10 mW. This setting might lead to sample degradation and cannot be applied in a reproducible manner. The mean spectra of the bacteria (Figure 4.2) reveal that *S. epidermidis* samples exhibit unusually high background. This reduces to normal levels after sufficiently long laser exposure times due to photobleaching. Photobleaching is caused by impurities within the sample, which are degraded by photolysis. The cause for photobleaching within the *S. epidermidis* sample is unknown. For this experiment, baseline correction was skipped to find out whether the spherical bacteria can be separated based on the background of the spectra. Figure 4.8 shows that both species can be

separated if integration time and laser power are high enough. The removal of baseline correction doubles the magnitude of variance captured in LV 1, due to inherently high fluorescence of *S. epidermidis*. This leads to better separation of *L. lactis* and *S. epidermidis* along the first LV and demonstrates the high impact of preprocessing on Multivariate Data Analysis.

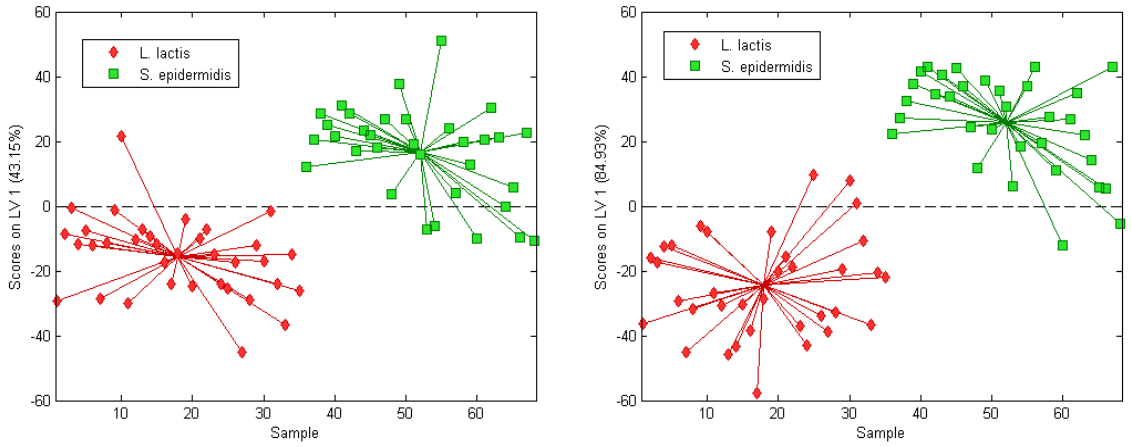

Figure 4.8: The plot of sample number vs LV1 demonstrates that the exclusion of background subtraction (right image) leads to better separation than its inclusion (left image). Preprocessing must be modified according to the sample type.

4.1.2 *E. coli* / *L. lactis* mixture

4.1.2.1 Single Spectra

In order to find out whether bacteria could be separated, while being present in the same sample, a sample containing the most distinct *E. coli* and *L. lactis* bacteria was investigated. The supervised classification method PLS-DA is not feasible in this case, because classes cannot be assigned to individual spectra in mixtures. Thus, PCA was applied for this experiment, with the aim to observe clustering within the mixture. Like LVs in PLS-DA, the choice of the number of principal components is not a trivial task. One prudent means to find the decent number is by plotting the eigenvalues and looking for a sudden jump or "knee" in the plot.[50] Figure 4.9 suggests a good model can be built by using 4 PCs, as a sharp decrease in the magnitude of captured variance can be observed between eigenvalue 4 and 5. Choosing more than >4 PCs makes no sense, because most of the model variance is already captured in the first 4 PCs. Figure 4.10 shows a plot of the first PC versus the sample number, which captures merely 15.29% of total variance.

The plot demonstrates high variance within the sample, which affirms that the sample is a mixture of different bacteria species.

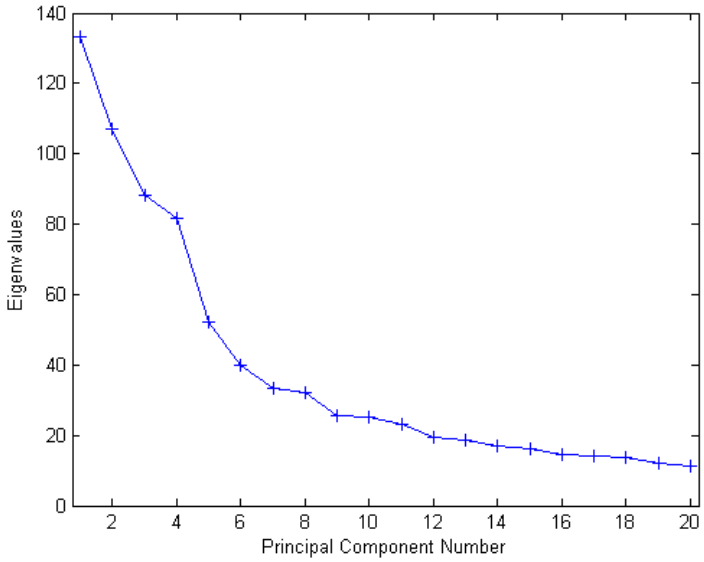

Figure 4.9: The eigenvalue plot shows a sharp decrease from 4 to 5, which suggests that the choice of 4 PCs is ideal, since they capture the greatest amount of total variance.

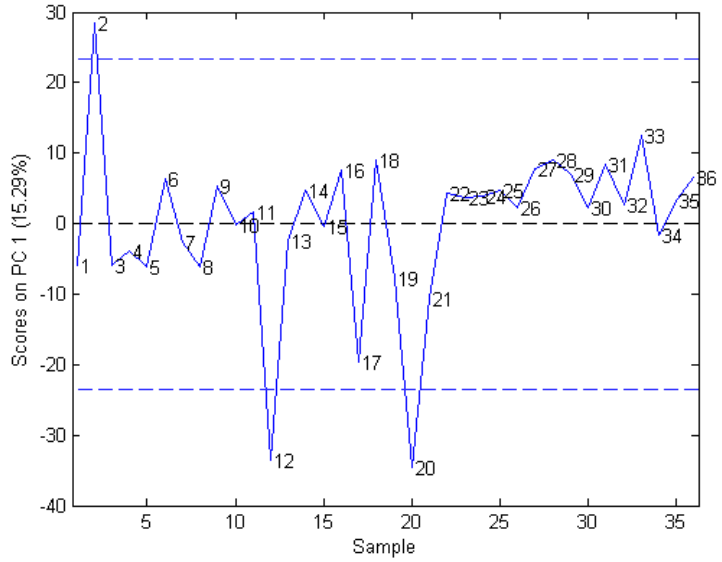

Figure 4.10: The PC1 plot shows that sample number 2, 12, 17 and 20 have unusually high scores. This observation suggests that the spectra belong to different types of bacteria.

Spectra 2, 12, 17 and 20 appear to have unusual scores that are substantially different from the remaining samples. Still, they appear to stem from bacteria cells. This indicates

that these spectra belong to a different bacteria species. Figure 4.11 shows a white light image of the investigated bacteria sample next to the respective spectra.

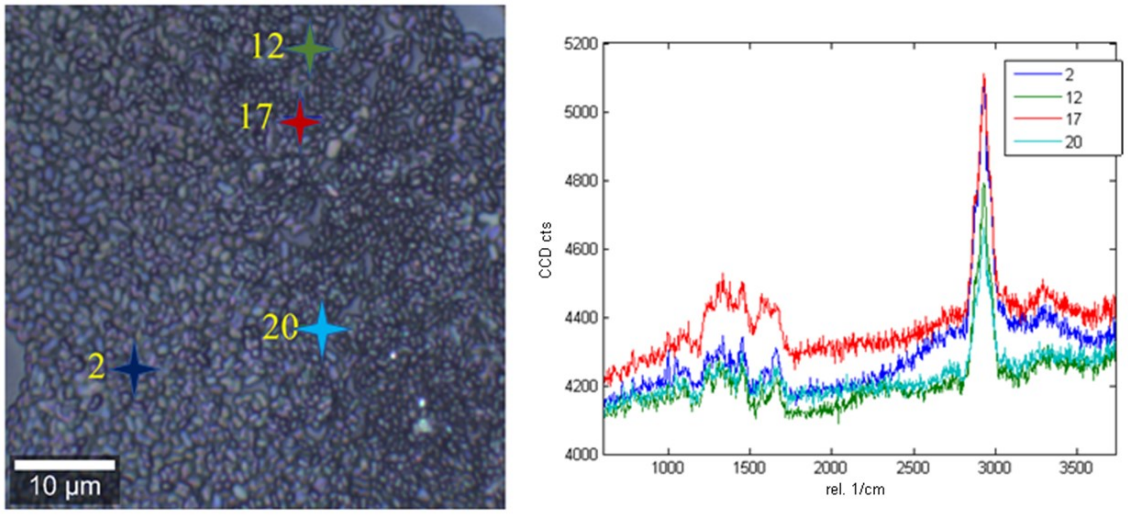

Figure 4.11: Left; the positions of sample number 2, 12, 17 and 20 are marked in the white light picture of the *E. coli*/*L. lactis* mixture. Within the bacteria cluster, it is difficult to separate distinct specimen by eye. Right; The spectra for sample number 2, 12, 17 and 20 appear to stem from bacteria samples. It is important to note, that the background signal has no influence on the PCA, due to the preprocessing step.

This assumption can be confirmed by the simple kNN technique, which separates a sample from another one by calculating the distance to the class of its nearest k neighbors (Figure 4.12).

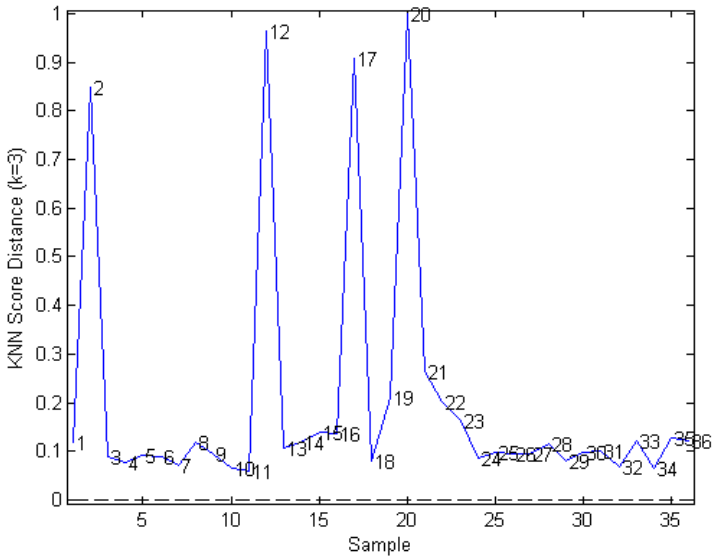

Figure 4.12: The kNN plot confirms the results of the PC1 plot, as shown in Figure 4.10. For the calculation of the k Score distance, k was set to 3.

However, the scores plot as well as the kNN distance provides no information concerning what variables (i.e. spectral regions) are useful in separating classes from each other. The loadings plot can give valuable information about the importance of the individual spectral regions. As expected, the fingerprint region, as well as the C-H vibrations, load heavily on the negative and positive PC1 score, respectively (Figure 4.13).

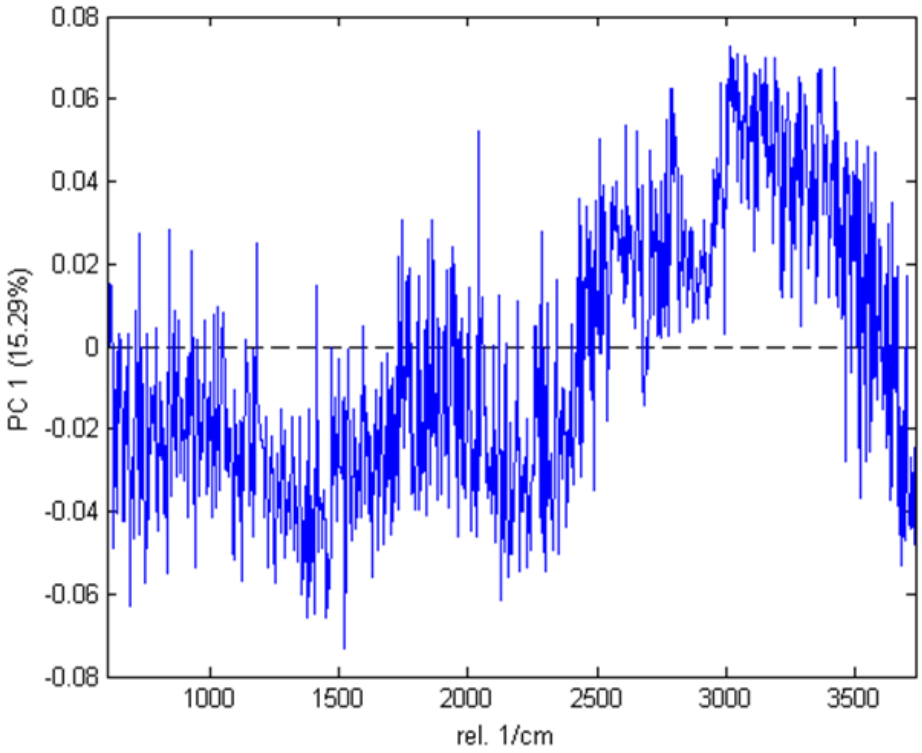

Figure 4.13: The loadings plot provides information on what spectral regions contribute the most to the separation of the modelled organisms. Apparently, the fingerprint region around 1400 cm^{-1} and the C-H vibrations around 3000 cm^{-1} contribute the most to the eigenvalues of the PCA model.

4.1.2.2 Image Scans

In further experiments, image scans of the *E. coli*/*L. lactis* mixture on CaF_2 were recorded. For the evaluation of these scans, PCA as well as TCA were applied. For every scan in this thesis, the number of pixels was set according to the size of the region of interest, so that a lateral resolution of 200 nm could be achieved. In the µm range this equates to 5 times as many points per line and lines per image as sampled μm^2. The aim of these experiments was to obtain spectral separation of different bacteria cells directly via scans

instead of single spectra. Figure 4.14 shows a white light image (100x magnification) of a region of 10 10 µm which was scanned with comparatively long integration time of 1 s, and the respective PCA false color picture, which was calculated via Solo + Mia.

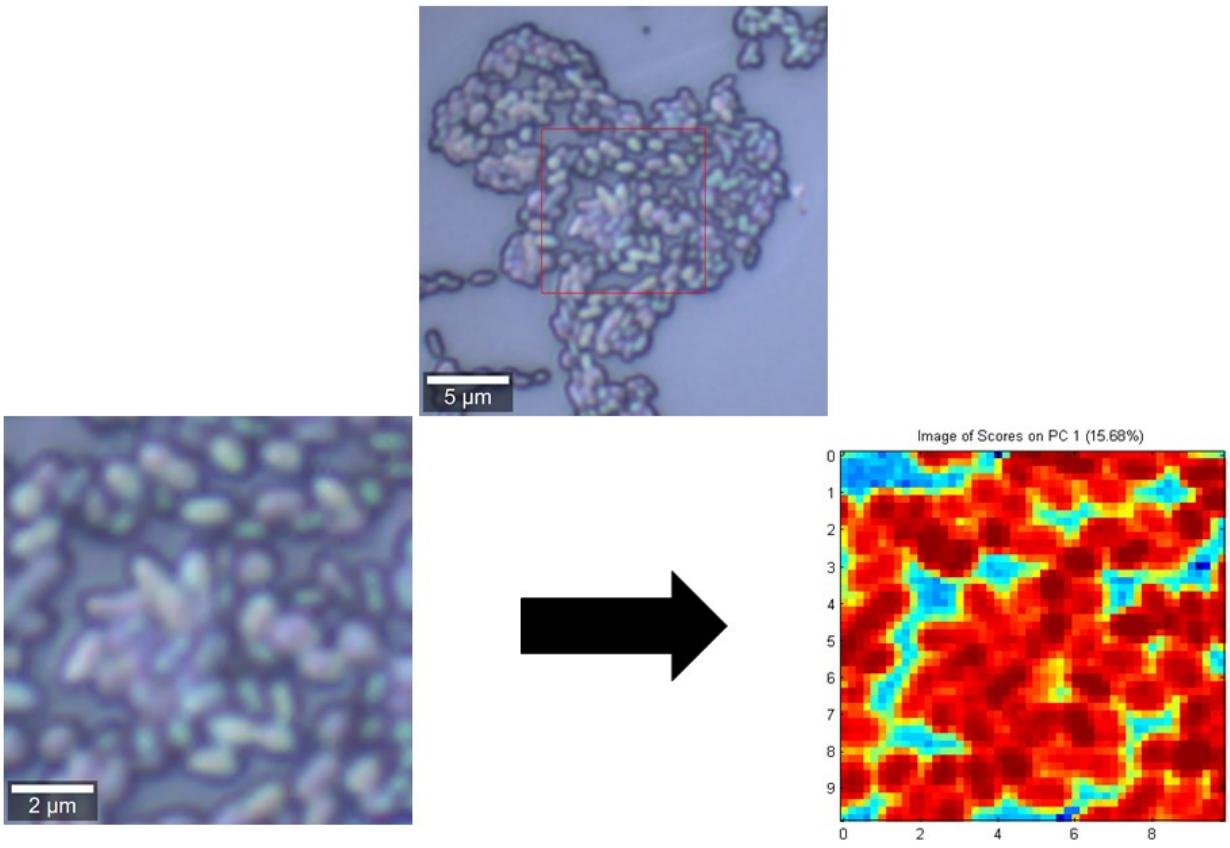

Figure 4.14: A 10 10 µm area of *E. coli* and *L. lactis* bacteria was scanned with an integration time of 1 s for a total of 131072 pixels. The PCA false color image illustrates the Bacteria and the CaF_2 substrate as red- and turquoise-colored, respectively. The first principal component in the *E. coli/L. lactis* model captures only 15.68% total variance. This suggests a high magnitude of residuals, which might be caused by fluorescence or remaining cosmic rays.

The PCA image of the scores depicts merely 2 major components, namely the CaF_2 substrate and the bacteria species. However, the false color picture does not distinguish between the bacteria species. Plotting of PC 1 versus PC 2 shows that the bacteria cluster, but there is no separation within the samples (Figure 4.15).

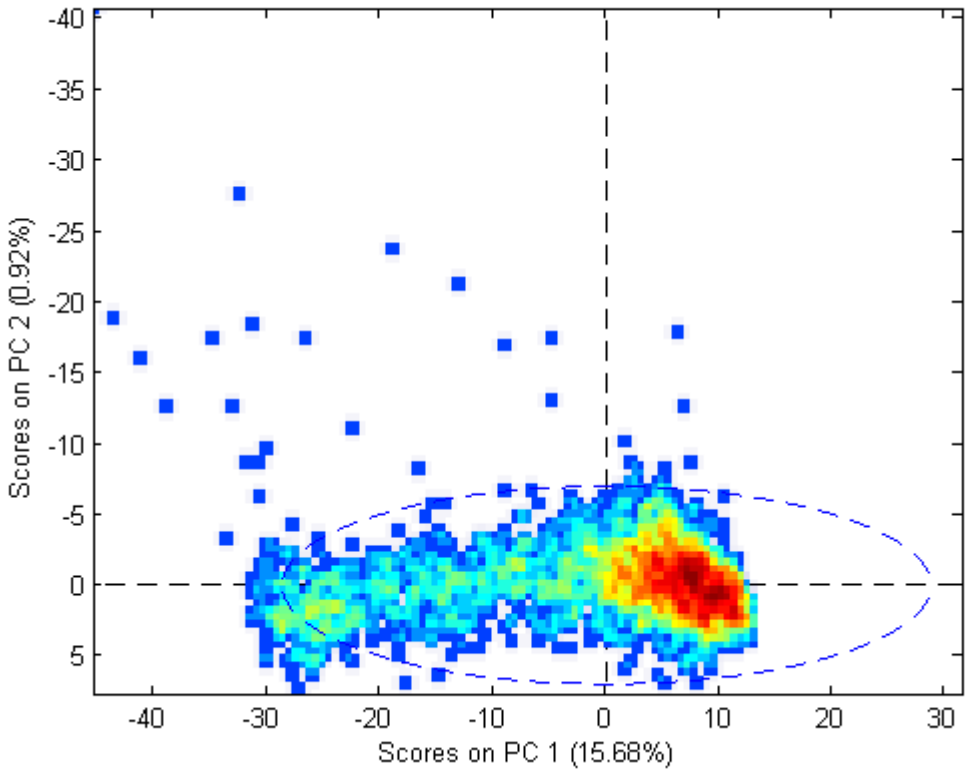

Figure 4.15: Plot of PC1 versus PC2 of the *E. coli*/*L. lactis* mixture. The clustering on the positive score of PC1 (red area) represents the bacteria spectra. However, PC2 captures below 1% variance of the system and should not be included in the final model.

Another attempt to separate *E. coli* from *L. lactis* was based on True Component Analysis (TCA). In TCA, the software tries to automatically detect individual spectral components, calculates the average for each component and combines them to obtain a false color picture. The average spectra and the combined Raman picture of the scanned region can be seen in Figure 4.16 and 4.17, respectively.

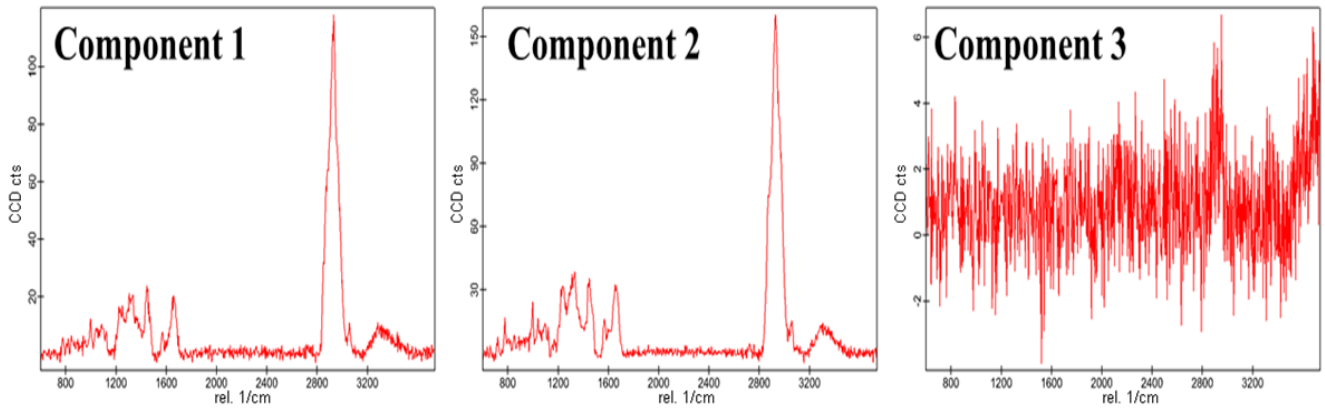

Figure 4.16: TCA automatically calculates 3 different components for the *E. coli*/*L. lactis* system. Component 1 and Component 2 apparently belong to different bacteria spectra. Component 3 depicts only noise and is spectral information provided by the CaF_2 substrate.

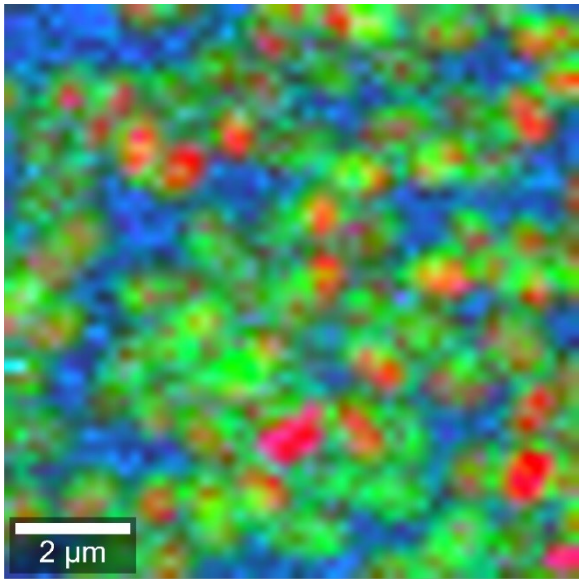

Figure 4.17: False color Raman image; all previously averaged components can be combined to a false color Raman image. The blue arears represent the CaF_2 spectra and the green and red areas stem from assumedly different bacteria cells.

Considering the data shown in Figure 4.16, no apparent difference in Component 1 and 2 can be spotted by eye. Component 3 clearly represents CaF_2. By assembling a false color picture from the different components, certain spots appear in a different color because they are built from a different component. In essence, the red spots on the false color picture represent component 2, whereas the green and blue hues stand for Component 1 and 3, respectively. Ideally, different components from bacteria spectra should prove the presence of different species. A huge drawback of this method is the

fact that the algorithm of the TCA method is only known to its author. For this thesis, it was assumed that the algorithm finds spectral components that take up a certain percentage of the sum of all spectra recorded for the scan. However, it remains speculation which criteria define a component and separate components from each other. Furthermore, the integration time might be too low for optimal segregation. Increasing the integration times further results in time-consuming scans, where the sample could even be altered by heavy laser exposure.

Concluding Section 4.1, different bacteria cells can be separated according to their spectra via PLS-DA. Nevertheless, *S. epidermidis* and *L. lactis* can barely be separated using standard preprocessing procedures. Excluding baseline correction leads to clear distinction between the spherical specimens. The most distinct *E. coli* and *L. lactis* can be separated by extensive single spectrum measurements with long integration times, but when it comes to image scans, shorter integration times lead to lower resolution. Even by scanning with 1 s integration time per pixel, Solo + Mia is unable to differentiate between *E. coli* and *L. lactis.* Even though TCA finds different components of bacteria spectra within the sample, the obscure algorithm requires careful interpretation of the result

4.2 Bacteria-Imprinted Polymer

Following the investigation of bacteria on CaF_2 as a substrate, the goal was to analyze *E. coli*-imprinted DVB/styrene polymer networks. The main topic was to find out whether a MIP cavity was occupied by *E. coli* cells or empty. For stamp imprinted polymer, the surface is expected to be devoid of bacteria cells, since the analyte was previously covalently attached to the stamp. Preparation of the bacteria cells, polymer synthesis and generation of the sensitive layers are explained in detail in section 3.1, 3.2 and 3.3, respectively. Successful production of stamps (section 3.3) can be reviewed by labelling the bacteria with fluorescein and checking the stamp surface under the microscope under UV light. However, the presence of fluorescein would be counterproductive for Raman investigations, since fluorescence can be regarded the nemesis of this spectroscopic method. Thus, the stamps were analyzed by white light only, prior to imprinting. Apparently, *E. coli* aggregates on the glass surface (Figure 4.18). The extent of aggregation depends on the bacteria concentration in solution.

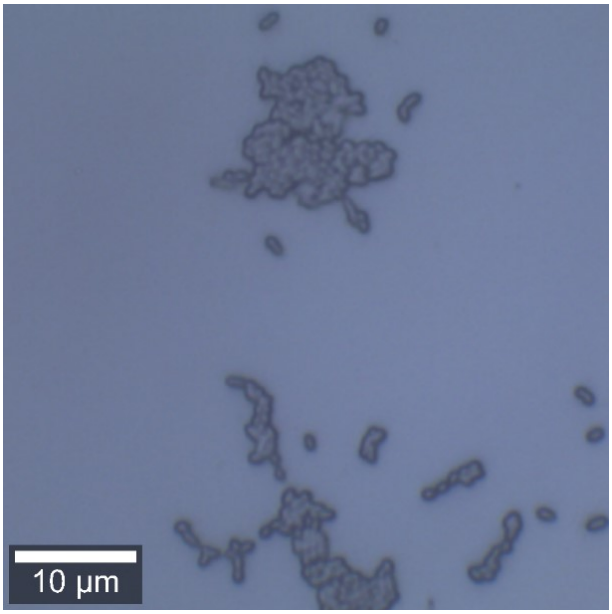

Figure 4.18: *E. coli* cells aggregate on the surface of a glass stamp. The microbes are covalently attached to the surface through DSS and APTES.

After stamp imprinting, the polymer sheet was expected to contain no bacteria at all. To test this hypothesis, Raman scans of the MIP surface were investigated via TCA and Raman TV. Figure 4.19 shows a white light image of the stamp-imprinted MIP surface.

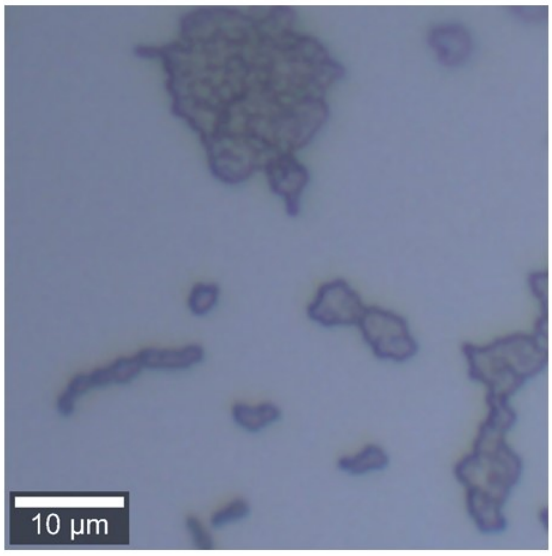

Figure 4.19: *E. coli* cells pressed in a DVB/Styrene matrix; after the imprinting process, white light pictures reveal apparent *E. coli* cavities. The spectroscopy part of the Raman microscopy technique aids in analyzing these spots qualitatively.

The aim of this Multivariate Data Analysis approach is to answer the question whether imprints are imprints at all, and if yes, whether they are occupied by bacteria cells or empty. Via WITec Suite Five's, two different major components were found: Component 1 as a part of the *E. coli* spots and component 2 represents the polymer matrix (Figure 4.20).

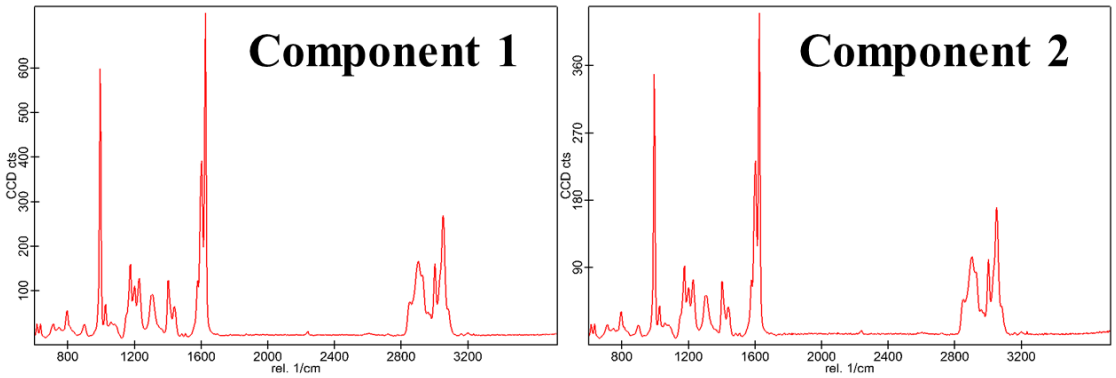

Figure 4.20: TCA components of *E. coli* stamp imprinted DVB/styrene. Apparently, there is no difference between the cavity component 1 and the matrix component 2. This suggests that, either there are no bacteria present at all, or the polymer spectra overlap all bacteria spectra.

However, both components seem to only differ in intensity. Qualitatively, no difference between component 1 and component 2 can be spotted by eye. Either, there are no bacteria present in the sample, or the spectra of the polymer system significantly overlap with the spectra of the microbes. The lack of distinction between component 1 and component 2 can be considered a general limitation of the Raman technique, when it comes to the investigation of bacteria embedded in polymer matrices. Figuratively speaking, for this system the bacteria are needles in a haystack. This might be due to the generally higher intensities of the polymer system, which overlap the bacteria spectrum. All experiments with bacteria spectra demonstrated that even slight fluctuations in the z-direction cause a loss of focus. This issue can partly be resolved by applying signal stabilization to the measurements. Nevertheless, the ruggedness of the Styrene-*co*-DVB assembly can provoke focus instabilities. Future experiments including z-scans might give more insight into this problem. Furthermore, even when focusing directly on the bacteria spots, the laser light will partly penetrate through the microbes and the detector collects information about the polymer matrix too. For the experiments in section 4.1, the weak bacteria signals were no issue at all, since the spectra of the CaF_2 substrate exhibit only 1 characteristic peak, thus there is no interference between bacteria and matrix. All in all, isolated Bacteria

spectra could not be recorded within the Styrene-*co*-DVB matrix. There is, however, the possibility to demix spectra. Basically, it can be assumed that the dark spots in Figure 4.19 are imprinted cavities. Subtracting component 1 from component 2 by demixing should result in pure bacteria spectra, if *E. coli* cells are present in the imprints. Subtracting component 2 from component 1 results in a demixed component 1 spectrum (Figure 4.21). Component subtraction was halted after the characteristic polymer peaks around 1000 cm^{-1} and 1600 cm^{-1} vanished.

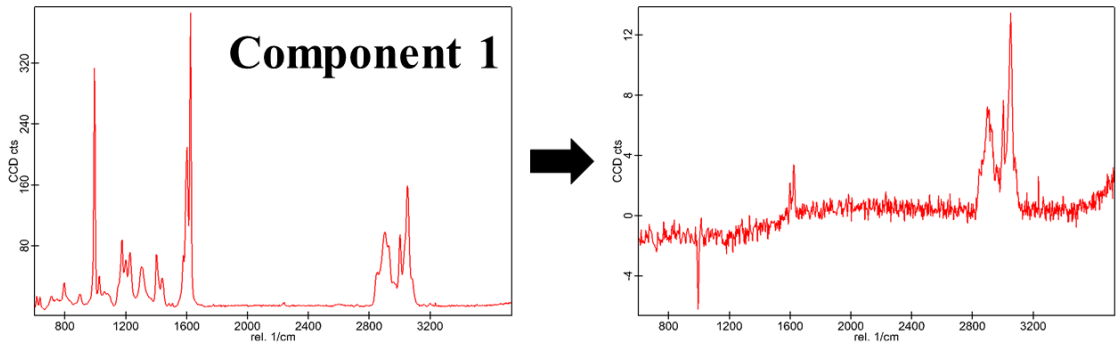

Figure 4.21: Demixing component 1 leaves only traces of the peaks around 3000 cm^{-1}. However, these traces can be considered negligible, due to their very low intensities of about 1/20 of the original spectrum.

The remaining component reveals that there is indeed no characteristic bacteria spectrum left. TCA helps in the identification of different components within the data of a Raman scan. However, the components alone do not yield topographical information, i.e., they don't reveal information about the depth of an imprint. For Raman TV, a certain area of the spectrum is isolated by a moving average filter, so that 2D and 3D images can be constructed. These false color images include information about the intensity distribution along the surface. For all Raman TV investigations, the spectral filter size was set to 10 and the peak around 1600 cm^{-1} was considered for the filter. Figure 4.22 shows a 2D Raman TV picture next to a 3D picture. The 3D Raman TV picture suggests that the spots on the Styrene/DVB surface are indeed cavities. Accompanying the Raman microscopy investigation and to confirm the results, AFM measurements were carried out on the same sample. Figure 4.23 shows a topographical image of the sample's surface, which also proves our initial assumption that no intact bacteria cells remain on the surface after stamp imprinting.

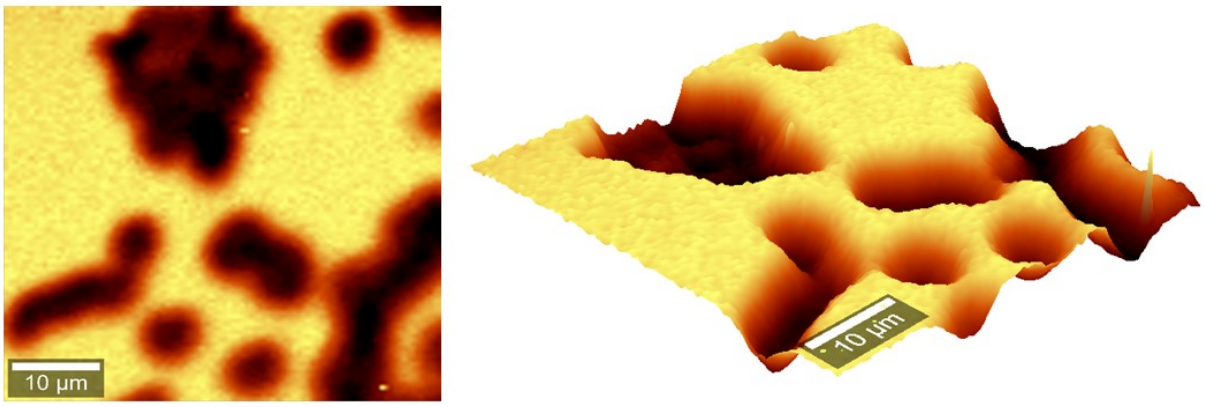

Figure 4.22: Raman TV images of *E. coli* imprinted DVB/Styrene; the bright spots are areas of high intensity and the dark spots areas of low intensity. For all Raman TV measurements, the predominant peak around 1600 cm^{-1} was considered. The 3D picture suggests that the dark spots are indeed cavities, due to the lower Raman signal in this area.

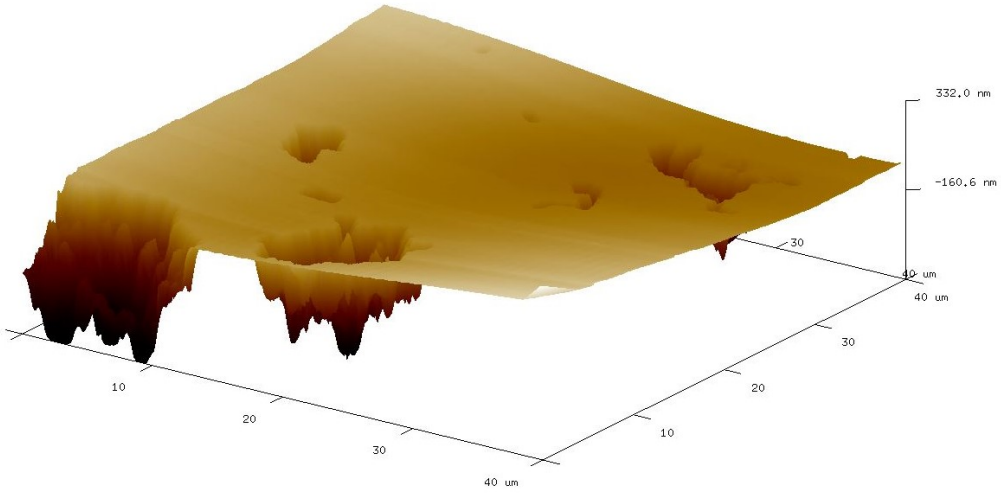

Figure 4.23: AFM image of *E. coli* imprinted DVB/Styrene; AFM is a more reliable method than TCA or Raman TV when it comes to answer topographical questions. The 50 μm^2 area was measured with a scan rate of 2Hz in contact mode.

However, it is important to note that the 3D Raman TV picture only gives information about the intensity distribution of a single peak within the scanned region. The dark spots represent areas with lower intensities. During the scan, the distance between the surface of the sample and the objective is kept constant via signal stabilization (Section 3.5), thus

lower intensities suggest larger distance between the objective and the sample. This leads to the conclusion that the increase in distance in certain areas is due to the presence of cavities. Anyhow, this observation purely depends on the focus of the laser light and reproducibility can be an issue. Considering this limitation of the Raman TV method, it is more prudent to apply AFM or complete 3D Raman scans, which were not subject of this thesis, for topographical investigation.

For the following experiment, the *E. coli* imprinted polymer sheet was immersed in concentrated *E. coli* solution to measure bacteria on the DVB/styrene surface. All measurements were carried out in the same way as the preceding investigation of *E. coli* imprinted DVB/styrene. The aim was to find out whether the dark spots are bacteria cells or imprints. Figure 4.24 shows a white light image of an apparent E. coli conglomeration on the polymer surface. TCA reveals 2 distinct components, of which component 1 belongs to the bacteria spots and component 2 to the polymer matrix (Figure 4.25). In contrast to the previous experiment, in which the cavities turned out to be vacant, component 2 looks different than component 1 in this case.

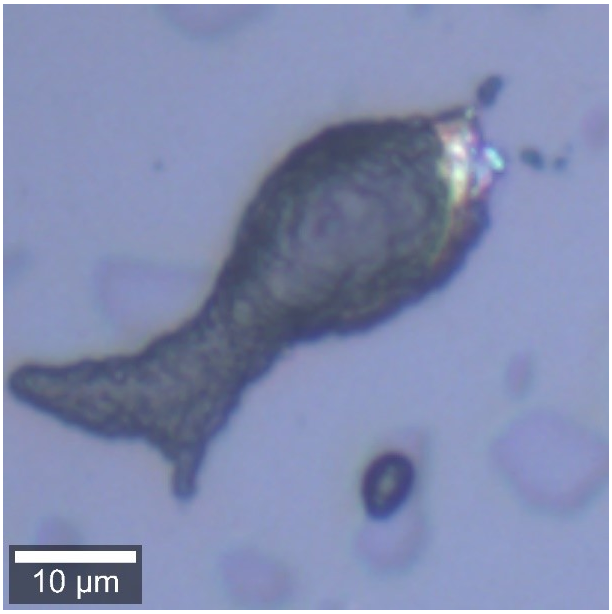

Figure 4.24: *E. coli* cells embedded in a DVB/Styrene matrix; after the immersion of the *E. coli* imprinted polymer sheet in *E. coli* solution, different spectral components were expected.

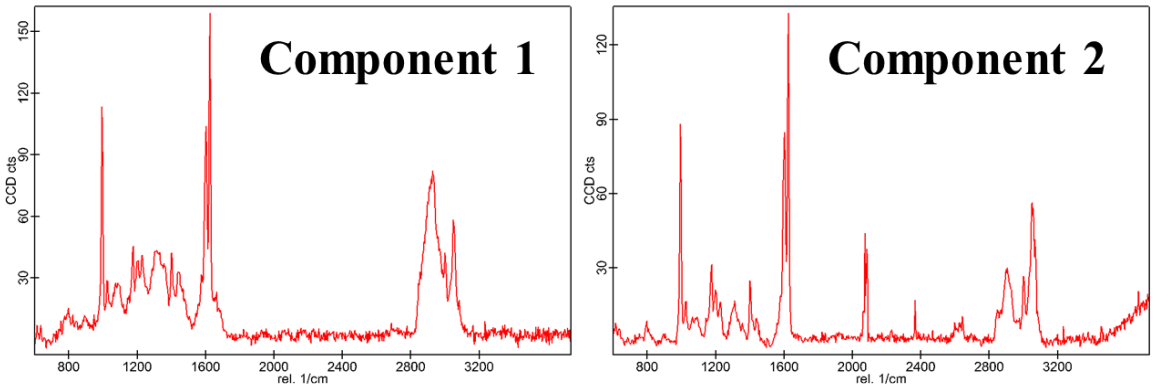

Figure 4.25: TCA components of *E. coli* stamp imprinted DVB/styrene immersed in E. coli solution; component 1 differs from component 2, however the averaged bacteria spectrum still exhibits characteristic polymer peaks.

Major discrepancies can be observed in the region around 1200 cm^{-1} and 3000 cm^{-1}. Nevertheless, there is still significant overlap between the bacteria and the polymer spectrum. Subtraction of component 2 from component 1 yields the demixed component 1 (Figure 4.26). Demixed component 1 reveals a spectrum which is quite similar to a pure E. coli spectrum (Figure 4.2). Not only the typical fingerprint region, ranging from 600 cm^{-1} to 1600 cm^{-1}, but also the broad single peak in the 3000 cm^{-1} region can be observed.

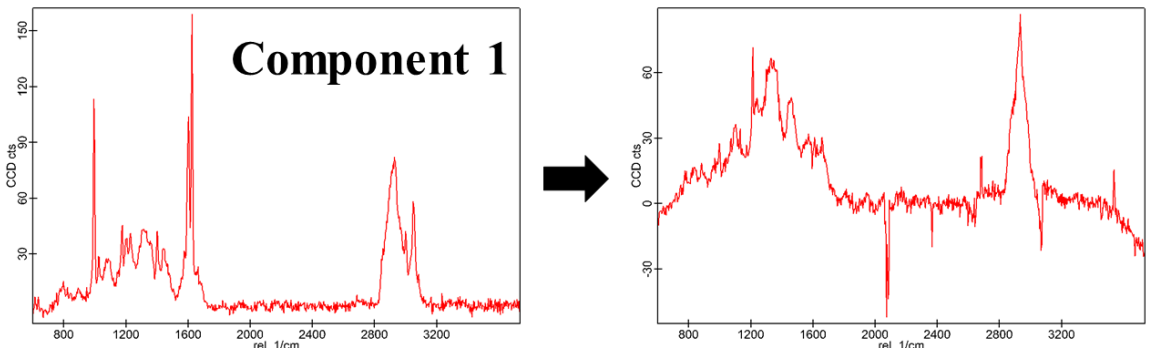

Figure 4.26: Demixing component 1 reveals an *E. coli* spectrum.

Figure 4.27 shows a 2D Raman TV image next to a 3D image. Like in the previous experiment the Raman TV investigation was complemented by AFM measurements, which reveal hills on the polymer surface, which stem from bacteria cells (component 1). However, the methods employed in this thesis are unable to tell whether the bacteria cells occupy a cavity or not. For this purpose, further experiments in liquid system with water immersion objectives could be carried out in the future.

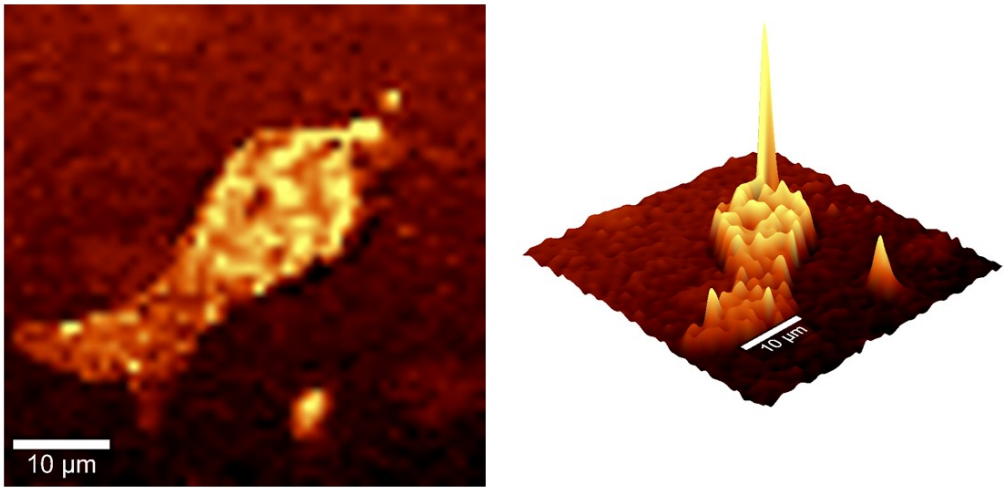

Figure 4.27: Raman TV images of *E. coli* imprinted DVB/Styrene after immersion in concentrated E. coli solution; The peak around 1600 cm^{-1} was used as a probe to create intensity dependent Raman TV images. Bright spots can be attributed to higher intensities. The 3D picture suggests that the distance between the objective and the sample's surface decreases over the polymer matrix.

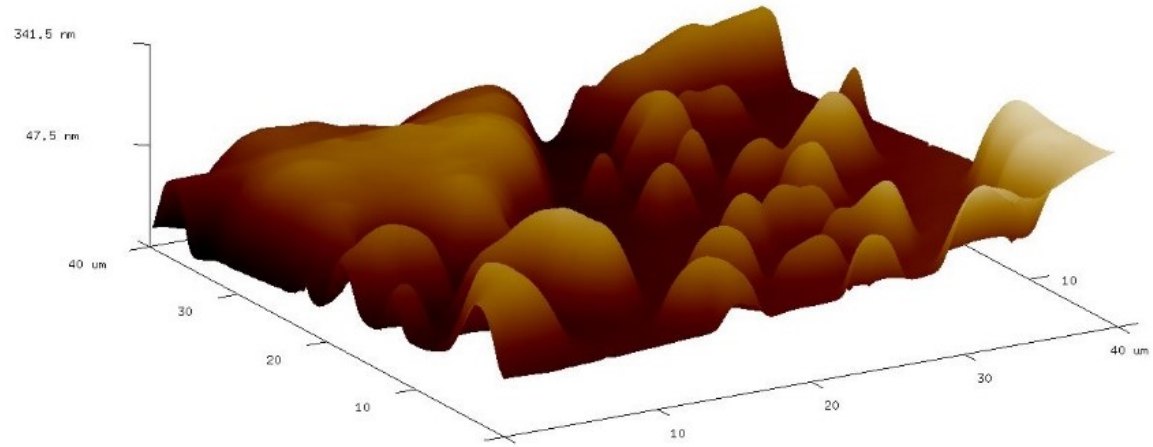

Figure 4.28: AFM image of *E. coli* imprinted DVB/Styrene after immersion in concentrated *E. coli* solution; the AFM image reveals reliable topographical information about the sample surface. In contrast to the AFM image depicted in Figure 4.23, this AFM image suggests that bacteria occupy the polymer surface.

Figure 4.27, as well as Figure 4.28 can be regarded as the negatives of Figure 4.22 and Figure 4.23, respectively. However, only the inclusion of AFM experiments can guarantee reliable topographical information, as Raman TV investigations depend on the focus of the objective.

In order to classify the calculated component spectra, a PLS-DA model including pure DVB/Styrene and *E. coli* spectra was set up. All components calculated in this section were included in the PLS-DA model in order to confirm which class they belong to (Figure 4.29). As expected, component 1 of the bacteria sample (Figure 4.26) scores closely to the E. coli cluster on LV 1. Component 2 of each experiment, as well as component 1 of the imprinted sample all score near the DVB-*co*-Styrene cluster.

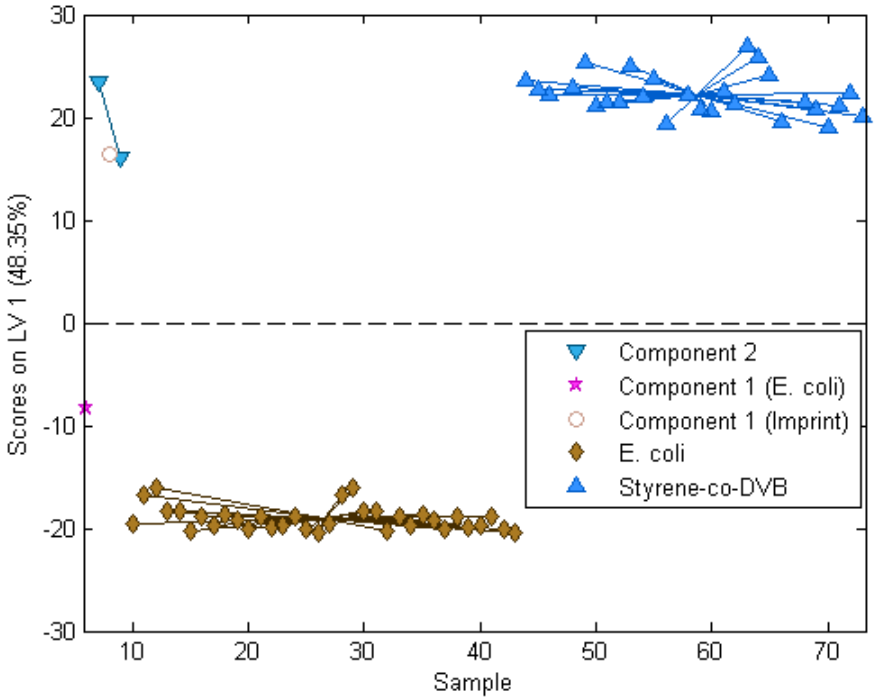

Figure 4.29: The PLS-DA model confirms the assumption that component 1 of the E. coli imprinted polymer sample, which was immersed in concentrated E. coli solution, scores near the E. coli cluster. All remaining components score close to the Styrene-*co*-DVB cluster, as expected.

5 Challenges and Optimization

All results presented in previous sections can be regarded reproducible in terms of experimental approach. However, many other attempts to describe MIP systems yielded less satisfying results and are clearly optimizable. This final section introduces challenges encountered and possible optimization steps for Raman microscopy investigation of MIP experiments.

5.1 QCM measurements

MIP results, as presented in Section 4 were investigated on CaF_2 sample holders to avoid disruptive background signal. Nevertheless, direct analysis of MIPs coated on QCM electrodes might give valuable insight into the physico-chemical properties of the system too. Figure 5.1 shows a white light image of a gold coated QCM (10 x magnification). 100 x magnification reveals that the gold coating includes several holes and overall insufficient homogeneity of the coating. Holes expose the quartz surface, which does not contribute to QCM signals in measuring cells. For this reason, homogeneous coating of QCM quartzes is of great importance for larger signals. Regarding Raman spectra, quartz as well as gold exhibit various peaks and high background signal, which might interfere with bacteria imprinted polymer signals, respectively (Figure 5.2, Figure 5.3). The desired polymer layer on QCM is ideally a few hundred nanometers thick, so that damping effects can be kept at a minimum. However, laser light will penetrate thin polymer layers and record background signal too. Generally, MIP measurements on QCM substrates can be challenging due to unwanted background signals.

Figure 5.1: White light picture of a QCM electrode. The black area surrounding the gold electrode is quartz substrate.

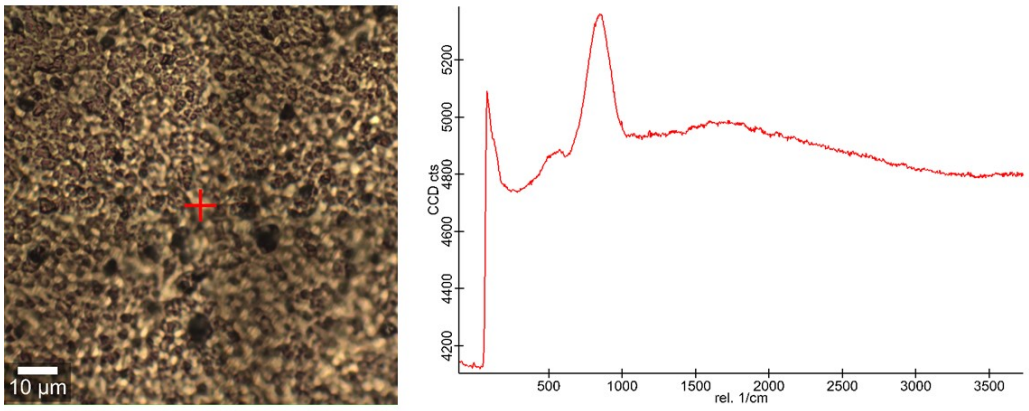

Figure 5.2: Spectrum of pure gold electrodes exhibits high fluorescent background signal, which can be disruptive when investigating MIPs. The spectrum was recorded with 532 nm laser, laser power of 8mW and 100 s integration time.

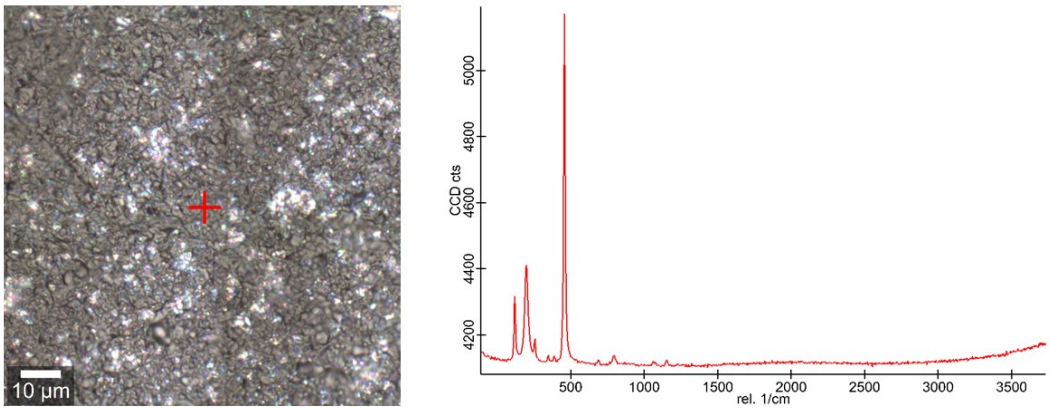

Figure 5.3: Spectrum of pure quartz regions on QCMs exhibit exhibits various signals <600 cm^{-1}, which can be disruptive when investigating MIPs. The spectrum was recorded with 532 nm laser, laser power of 8mW and 100 s integration time.

For future experiments higher wavelength lasers can be used to investigate MIP systems directly on QCM substrates. The laser wavelength was kept at 532 nm for all experiments mentioned in this thesis, because bacteria inherently exhibit weak Raman signals and higher wavelengths will lessen the signal strength even further.

5.2 Signal Stabilization

Signal stabilization is a prerequisite for Raman scans conducted with WITec ALPHA 300R confocal Raman microscope. Raman scans without signal stabilization will result in a loss of ~1 µm focus in z-direction within 30 minutes. The lateral resolution of the Raman microscope is ~200 nm and bacteria cells usually exhibit Raman signals of feasible strength in a range of ~500 nm. Figure 5.4 shows a white light image of E. coli imprinted acrylamide-based polymer next to Raman image, obtained via Raman TV (Section 4). The bright hues represent higher intensities and the dark hues regions of lower intensity. Apparently, intensity decreases with scan time, due to thermal expansion. Irradiation with high intensity laser heats the air between the sample surface and the Raman microscope objective, which subsequently results in a drift in z-direction. Signal stabilization checks the focus after each line and compensates for the respective loss of focus. However, there is a major drawback with signal stabilization: Stabilization time depends on the set integration time of the Raman scan. Comparatively long integration times of about 1 s / point will result in exorbitantly high total scan times, as the signal stabilization will take up a major fraction of the total scan time. Furthermore, 3D scans are more or less impossible to be conducted within a reasonable time scale, as signal

stabilization will cover the largest fraction of the scan time. Additionally, exposing the sample to high laser power for extended time periods might result in sample degradation. Future software updates might resolve this problem, but as of right now, mandatory signal stabilization proves to be inefficient for longer integration times and 3D scans. Fast high-intensity Raman scans without signal stabilization <30 minutes pose a meaningful middle ground.

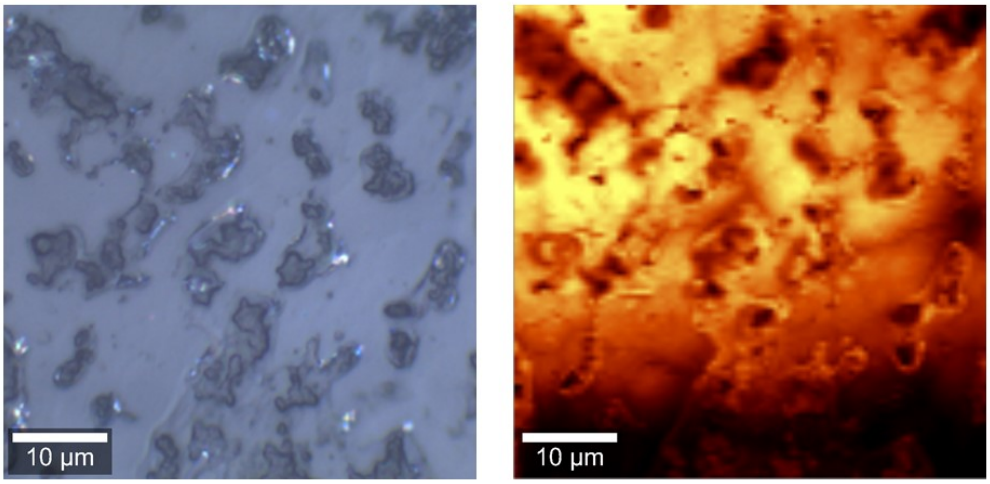

Figure 5.4: Conducting Raman scans without signal stabilization results in extensive loss of focus in z-direction, caused by thermal drift. The bright to dark hue gradient in the Raman TV image on the right side demonstrates the loss of focus. Scan time for this sample was set to 83 minutes. The laser power for the 532 nm laser was set to 8 mW.

6 Conclusion

For this master's thesis, Multivariate Data Analysis models for spectra of bacteria imprinted polymers could successfully be established. The required Styrene-*co*-DVB MIPs were synthesized using standard polymerization protocols coupled with the novel stamp imprinting technique. Prior to the application of PLS-DA, PCA and TCA, Raman images and scans of various bacteria samples were recorded. CaF_2 proved to be the ideal substrate for bacteria investigation, since any fluorescent source impedes the investigation of the fundamentally weak Raman signal. PLS-DA reliably distinguished between most of the bacteria species except for *S. epidermidis* and *L. lactis*, which exhibit very similar Raman spectra. Unsupervised PCA detected high variation within a sample consisting of 2 different bacteria species, which indicates that direct separation of different species is possible with sufficiently long integration times. For bacteria MIP, the results often appeared less clear at first sight. Pure E. coli spectra could not be recorded by focusing bacteria spots on the polymer substrate, because the weak microbe Raman bands significantly overlapped with the polymer spectra. For future work, this investigation can be extended by conducting scans in z-direction, so that the focus of the objective varies within a measurement. TCA and Raman TV delivered promising results, but especially TCA's obscure algorithm prevents representative interpretation of the components. Raman TV must not be confused with topographical information, since it only provides 3D pictures of intensity distributions. However, if the composition of the sample is already well known, spectral filters can be a powerful tool accompanying AFM measurements. All in all, WITec Project FIVE and WITec Suite FIVE are still in their infancies and as of right now, no manual even exists. Most of the work carried out within the framework of this master's thesis can be regarded pioneering. For future Raman microscopy investigation of MIPs, water immersion objectives can aid in getting a deeper understanding of the binding and rebinding process of analytes during QCM measurements, as many interesting sensors respond in liquid phase.

7 Bibliography

1. WHO *Millennium Development Goals: progress towards the health-related Millennium Development Goals.* ; 2011.

2. WHO *Water Sanitation and Health.* ; 2010.

3. Craun, M. F.; Craun, G. F.; Calderon, R. L.; Beach, M. J., Waterborne outbreaks reported in the United States. *Journal of water and health* **2006,** *4 Suppl 2*, 19-30.

4. National Primary Drinking Water Regulations. https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations.

5. (a) Malakoff, D., Water quality. Microbiologists on the trail of polluting bacteria. *Science (New York, N.Y.)* **2002,** *295* (5564), 2352-3; (b) Edberg, S. C.; Rice, E. W.; Karlin, R. J.; Allen, M. J., Escherichia coli: the best biological drinking water indicator for public health protection. *Symposium series (Society for Applied Microbiology)* **2000,** (29), 106s-116s.

6. (a) Davey, H. M.; Kell, D. B., Flow cytometry and cell sorting of heterogeneous microbial populations: the importance of single-cell analyses. *Microbiological Reviews* **1996,** *60* (4), 641-696; (b) Yamamoto, Y., PCR in Diagnosis of Infection: Detection of Bacteria in Cerebrospinal Fluids. *Clinical and Diagnostic Laboratory Immunology* **2002,** *9* (3), 508-514.

7. Samardzic, R.; Sussitz, H. F.; Jongkon, N.; Lieberzeit, P. A., Quartz Crystal Microbalance In-Line Sensing of Escherichia Coli in a Bioreactor Using Molecularly Imprinted Polymers. *Sensor Letters* **2014,** *12* (6-7), 1152-1155.

8. Lieberzeit, P. A.; Samardzic, R.; Kotova, K.; Hussain, M., MIP Sensors on the Way to Biotech Application: Selectivity and Ruggedness. *Procedia Engineering* **2012,** *47* (Supplement C), 534-537.

9. Figueiredo, L.; Erny, G. L.; Santos, L.; Alves, A., Applications of molecularly imprinted polymers to the analysis and removal of personal care products: A review. *Talanta* **2016,** *146* (Supplement C), 754-765.

10. Schie, I. W.; Huser, T., Methods and applications of Raman microspectroscopy to single-cell analysis. *Applied spectroscopy* **2013,** *67* (8), 813-28.

11. Huang, W. E.; Griffiths, R. I.; Thompson, I. P.; Bailey, M. J.; Whiteley, A. S., Raman Microscopic Analysis of Single Microbial Cells. *Analytical Chemistry* **2004,** *76* (15), 4452-4458.

12. Schuster, K. C.; Reese, I.; Urlaub, E.; Gapes, J. R.; Lendl, B., Multidimensional Information on the Chemical Composition of Single Bacterial Cells by Confocal Raman Microspectroscopy. *Analytical Chemistry* **2000,** *72* (22), 5529-5534.

13. Lieberzeit, P. A.; Dickert, F. L., Sensor technology and its application in environmental analysis. *Analytical and bioanalytical chemistry* **2007,** *387* (1), 237-47.

14. Hulanicki, A.; Glab, S.; Ingman, F., Chemical sensors: definitions and classification. In *Pure and Applied Chemistry*, 1991; Vol. 63, p 1247.

15. (a) Janshoff, A.; Galla, H.-J.; Steinem, C., Piezoelectric Mass-Sensing Devices as Biosensors—An Alternative to Optical Biosensors? *Angewandte Chemie International Edition* **2000,** *39* (22), 4004-4032; (b) Denzler., M. P., Lebensdauer und Zuverlässigkeit dynamisch betriebener piezokeramischer
Biegewandler. . *Cuvillier* **2004**.

16. Dickert, F. L.; Lieberzeit, P. A., Imprinted Polymers in Chemical Recognition for Mass-Sensitive Devices. In *Piezoelectric Sensors*, Janshoff, A.; Steinem, C., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 2007; pp 173-210.

17. Vashist, S. K.; Vashist, P., Recent Advances in Quartz Crystal Microbalance-Based Sensors. *Journal of Sensors* **2011,** *2011*, 13.

18. Surfaces and interfaces for biomaterials A2 - Vadgama, Pankaj. In *Surfaces and Interfaces for Biomaterials*, Woodhead Publishing: 2005; p ii.

19. Latif, U.; Qian, J.; Can, S.; Dickert, F., Biomimetic Receptors for Bioanalyte Detection by Quartz Crystal Microbalances — From Molecules to Cells. *Sensors* **2014,** *14* (12), 23419.

20. Ruigrok, Vincent J. B.; Levisson, M.; Eppink, Michel H. M.; Smidt, H.; van der Oost, J., Alternative affinity tools: more attractive than antibodies? *Biochemical Journal* **2011,** *436* (1), 1-13.

21. (a) Haupt, K., Molecular Imprinting. *Springer* **2012,** *325*; (b) Haupt, K.; Mosbach, K., Molecularly Imprinted Polymers and Their Use in Biomimetic Sensors. *Chemical Reviews* **2000,** *100* (7), 2495-2504.

22. Lieberzeit, S. Z. B. a. P. A., Recognition principle of Cu2+ imprinted polymers
- Assessing interactions by combined spectroscopic and mass-sensitive measurements. *Sens. Actuator B Chem* **2015**.

23. Chunta, S.; Suedee, R.; Lieberzeit, P. A., Low-Density Lipoprotein Sensor Based on Molecularly Imprinted Polymer. *Anal Chem* **2016,** *88* (2), 1419-25.

24. Spieker, E.; Lieberzeit, P. A., Molecular Imprinting Studies for Developing QCM-sensors for Bacillus Cereus. *Procedia Engineering* **2016,** *168* (Supplement C), 561-564.

25. Uzun, L.; Turner, A. P. F., Molecularly-imprinted polymer sensors: realising their potential. *Biosensors and Bioelectronics* **2016,** *76* (Supplement C), 131-144.

26. Vasapollo, G.; Sole, R. D.; Mergola, L.; Lazzoi, M. R.; Scardino, A.; Scorrano, S.; Mele, G., Molecularly Imprinted Polymers: Present and Future Prospective. *International Journal of Molecular Sciences* **2011,** *12* (9), 5908-5945.

27. Ertürk, G.; Mattiasson, B., Molecular Imprinting Techniques Used for the Preparation of Biosensors. *Sensors (Basel, Switzerland)* **2017,** *17* (2), 288.

28. Nantasenamat, C.; Isarankura-Na-Ayudhya, C.; Naenna, T.; Prachayasittikul, V., Quantitative structure-imprinting factor relationship of molecularly imprinted polymers. *Biosensors and Bioelectronics* **2007,** *22* (12), 3309-3317.

29. Muhammad, T.; Nur, Z.; Piletska, E. V.; Yimit, O.; Piletsky, S. A., Rational design of molecularly imprinted polymer: the choice of cross-linker. *The Analyst* **2012,** *137* (11), 2623-8.

30. Mujahid, A.; Iqbal, N.; Afzal, A., Bioimprinting strategies: From soft lithography to biomimetic sensors and beyond. *Biotechnology Advances* **2013,** *31* (8), 1435-1447.

31. Whitcombe, M. J.; Chianella, I.; Larcombe, L.; Piletsky, S. A.; Noble, J.; Porter, R.; Horgan, A., The rational development of molecularly imprinted polymer-based sensors for protein detection. *Chemical Society Reviews* **2011,** *40* (3), 1547-1571.

32. Hillberg, A. L.; Tabrizian, M., Biomolecule imprinting: Developments in mimicking dynamic natural recognition systems. *IRBM* **2008,** *29* (2), 89-104.

33. Zahedi, P.; Ziaee, M.; Abdouss, M.; Farazin, A.; Mizaikoff, B., Biomacromolecule template-based molecularly imprinted polymers with an emphasis on their synthesis strategies: a review. *Polymers for Advanced Technologies* **2016,** *27* (9), 1124-1142.

34. Lin, T.-Y.; Hu, C.-H.; Chou, T.-C., Determination of albumin concentration by MIP-QCM sensor. *Biosensors and Bioelectronics* **2004,** *20* (1), 75-81.

35. Dickert, F. L.; Lieberzeit, P.; Gazda-Miarecka, S.; Halikias, K.; Mann, K.-J., Modifying polymers by self-organisation for the mass-sensitive detection of environmental and biogeneous analytes. *Sensors and Actuators B: Chemical* **2004,** *100* (1), 112-116.

36. Hacker, J.; Blum-Oehler, G., *In appreciation of Theodor Escherich*. 2007; Vol. 5.

37. Blount, Z. D., The unexhausted potential of E. coli. *eLife* **2015,** *4*, e05826.

38. Wang, Y.; Ping, J.; Ye, Z.; Wu, J.; Ying, Y., Impedimetric immunosensor based on gold nanoparticles modified graphene paper for label-free detection of Escherichia coli O157:H7. *Biosensors & bioelectronics* **2013,** *49*, 492-8.

39. De Middeleer, G.; Dubruel, P.; De Saeger, S., Characterization of MIP and MIP functionalized surfaces: Current state-of-the-art. *TrAC Trends in Analytical Chemistry* **2016,** *76* (Supplement C), 71-85.

40. Binnig, G.; Quate, C. F.; Gerber, C., Atomic Force Microscope. *Physical Review Letters* **1986,** *56* (9), 930-933.

41. Chang, K.-C.; Chiang, Y.-W.; Yang, C.-H.; Liou, J.-W., Atomic force microscopy in biology and biomedicine. *Tzu Chi Medical Journal* **2012,** *24* (4), 162-169.

42. (a) Israelachvili, J., *Intermolecular and Surface Forces*. 2011; (b) A.De Stefanis, A. A. G. T., Scanning Probe Microscopies: From Surface Structure to Nano-scale Engineering (Materials Science Foundation). **2001**.

43. Thomas Dieing, O. H., Jan Toporski, Confocal Raman Microscopy. **2011**.

44. (a) Schmitt, M.; Popp, J., Raman spectroscopy at the beginning of the twenty-first century. *Journal of Raman Spectroscopy* **2006,** *37* (1-3), 20-28; (b) Raman, C. V., A new radiation [Reproduced from Indian J. Phys., 1928, 2, 387–398]. *Current Science* **1998,** *74* (4), 382-386.

45. Wieboldt, D. Understanding Raman Spectrometer Parameters. http://www.spectroscopyonline.com/understanding-raman-spectrometer-parameters.

46. I. R. Lewis, H. G. M. E., *Handbook of Raman Spectroscopy*. Dekker, New York,

Basel, 2001.

47. WITec http://www.witec.de/de/produkte/alpha300-reihe/alpha300-r/.

48. El-Gindy, A.; Hadad, G. M., Chemometrics in pharmaceutical analysis: an introduction, review, and future perspectives. *Journal of AOAC International* **2012,** *95* (3), 609-23.

49. Jolliffe, I. T.; Cadima, J., Principal component analysis: a review and recent developments. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* **2016,** *374* (2065), 20150202.

50. Barry M. Wise, J. M. S.; Neal B. Gallaghe, W. W.; Rasmus Bro, R. S. K., *Chemometrics Tutorial for PLS_Toolbox and Solo*.

51. Lorber, A.; Wangen, L. E.; Kowalski, B. R., A theoretical foundation for the PLS algorithm. *Journal of Chemometrics* **1987,** *1* (1), 19-31.

52. G. Brereton, R.; Lloyd, G., *Partial least squares discriminant analysis: Taking the magic away*. 2014; Vol. 28.

53. SOFTWARE WITEC SUITE. http://www.witec.de/products/accessories/software-witec-suite/.

54. Incorporated, E. R. http://www.eigenvector.com/software/solo+mia.htm.

Summary / Zusammenfassung

In this work, Raman Microscopy followed by Multivariate Data Analysis were applied to study different bacteria systems. Vibrational Raman spectra were obtained from different bacteria species and Principal Component Analysis (PCA) and Partial Least Square-Discriminant Analysis (PLS-DA) were employed for classification and data evaluation. Furthermore, bacteria-imprinted Styrene-*co*-DVB assemblies were synthesized and topographically examined. For this purpose, Atomic Force Microscopy (AFM) accompanied Raman Microscopy for the detailed description of the Molecularly Imprinted Polymer (MIP) surfaces. The analyte of interest was *E. coli*, a gram-negative bacterium, which is universally acknowledged as a bio-indicator for water contamination. PCA and PLS-DA models demonstrated that different bacteria species can be separated by their spectral fingerprints, even though differences within their spectra cannot be registered by the naked eye. Coupling of Raman experiments with atomic force microscopy enabled topographical and qualitative insight into the surface of bacteria-imprinted polymers.

In dieser Arbeit wurde Raman Mikroskopie in Kombination mit Multivariater Datenanalyse angewandt um verschiedene Bakterien-Systeme zu untersuchen. Schwingungsspektren wurden von unterschiedlichen Bakterien-Spezies aufgenommen und mittels Hauptkomponentenanalyse und Partial Least Square-Discriminant Analysis (PLS-DA) unterschieden. Außerdem wurden mit Bakterien geprägte Styrene-*co*-DVB Systeme synthetisiert und topografisch untersucht. Als Analyt wurde mit E. Coli gearbeitet, einem gram-negativen Bakterium, das universell als Bio-Indikator für Wasserverschmutzung anerkannt ist. Hauptkomponentenanalyse und PLS-DA Modelle demonstrierten, dass unterschiedliche Bakterien anhand ihrer individuellen spektralen Fingerabdrücke unterschieden werden können, obwohl die Spektren kaum qualitative Unterschiede aufweisen. Die Verbindung von Raman Mikroskopie Experimenten mit Rasterkraftmikroskopie erlaubte das Erstellen topografischer Modelle der Oberfläche der geprägten Polymere.