Title

Understanding the State of Bacterial Systems with Machine Learning-Enabled Interpretation of Surface Enhanced Raman Scattering Spectra Produced by Novel Nanomanufacturing

Author
Thrift, William John

Publication Date
2019

UNIVERSITY OF CALIFORNIA,
IRVINE

Understanding the State of Bacterial Systems with Machine Learning-Enabled Interpretation of Surface Enhanced Raman Scattering Spectra Produced by Novel Nanomanufacturing

Dissertation

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Materials Science and Engineering

by

William John Thrift

Dissertation Committee:
Professor Regina Ragan, Chair
Professor Filippo Capolino
Associate Professor Allon Hochbaum

2019

DEDICATION

To the people who made this dissertation possible: my family, Regina, and Hurik.

Also Roger Penrose, whose book, “The Road to Reality” inspired me to go to get a PhD.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

It takes a village to write a PhD thesis and I am deeply indebted — and forever grateful — to the people who have helped me along the way.

First and foremost I would like to thank **Regina Ragan**. Regina has always been supportive of me and has enabled me to pursue all of my ideas, some fruitful, and some crazy. She pushed me to explore science outside of my comfort zone, and that has shaped who I am today. For many, the PhD experience is a difficult, yet rewarding undertaking. Because of Regina, it was for me as fun as it was rewarding. For these reasons, I hold you in the highest of esteem and am grateful for everything that you've done for me.

Along the way I have also been fortunate to have worked with some great collaborators. My committee members, **Filippo Capolino** and **Allon Hochbaum**, have given my work depth and clarity that could never have been achieved without them. Just read our mutual works and you will surely see their impact. I want to also thank **Mahsa Darvishzadeh-Varcheie**. I'm proud of the work that we did together, you've taught me a lot over the years.

I'd like to thank my research group. **Nicholas Sharac**, thank you for mentoring me when I first joined the group, without you I would have been lost. **Robert Sanderson, Chen Wang, Adrian Garcia, Hong Wei,** and **Chloe Groome**: thanks for talking to me all these years. I had fun.

I've taken great joy in mentoring several students during my PhD, and I'd like to thank all of them: **Megan Quan**, **Antony Cabuslay**, and **Andrew Laird**. Thanks for sticking it out and working so hard. Especially you Antony, take some Vitamin D, you need if after all that time in the pitch black laser room.

And of course I'd like to thank **Cuong Nguyen.** You were both a collaborator and, more importantly, a friend. I learned so much from our discussions, and our arguments sharpened my mind, and for that I am grateful. Our brainstorming sessions remain the pinnacle of my PhD. Thank you.

I'd like to thank my family. My parents **John** and **Patti Thrift**, your unconditional love and support has enabled everything good in my life. My sisters **Emma** and **Jessica Thrift**, you shaped my childhood and continue to shape me to this day. My dog **Daisy**, who still can't read.

And finally, **Hurik Muradyan**. I don't have the words that could do you justice. You've made me a better person (and incidentally, a better scientist). My years with you are the best in my life. It's worth writing this entire dissertation just to acknowledge you. I love you.

This work was supported by National Science Foundation (NSF) EECS-1449397, NSF IGERT Fellowship, UCI Public Impact Fellowship, and UCI Graduate Dean's Dissertation Fellowship. I thank and acknowledge UCI's Laser Spectroscopy Facility and the Irvine Materials Research Institute for the use of their facilities.

CURRICULUM VITAE

WILL THRIFT

Education

PhD	University of California, Irvine, Materials Science and Engineering Dissertation: "Understanding the State of Bacterial Systems with Machine Learning-Enabled Interpretation of Surface Enhanced Raman Scattering Spectra Produced by Novel Nanomanufacturing" Committee: Regina Ragan, Filippo Capolino, Allon Hochbaum GPA: 3.96 (3.91 through masters)	September 2019
BS	University of California, San Diego, Nanoengineering GPA: 3.31 (3.74 major)	June 2013

Research and Work Experience

Development and Implementation of Surface Enhanced Raman Scattering Biosensors, **UC Irvine** **2014-2019**

Advisor: Regina Ragan

- Invented new chemical self-assembly method for manufacturing optical structures composed of gold nanoparticles. The work proved that chemical crosslinking can be activated with electrokinetic colloidal driving forces
- Developed a deep convolutional neural network model for determining analyte concentration at ultralow concentrations from spectroscopic data. The work demonstrated that concentration can be inferred from single sensing events by determining the distribution of these events.
- Developed a convolutional neural network method for identifying odorant source direction from spatial spectroscopic data.
- Implemented a templated chemical assembly method using electron beam lithography.
- Mentored 6 undergraduate students and 2 masters students from chemical engineering, chemistry, materials science, computer science, and data science.
  - Helped them develop:
    - an in-line biosensing platform for longitudinal monitoring of biofilm growth
    - an AC electrohydrodynamic flow chemical self-assembly method
    - a generative adversarial network method for data augmentation of Raman spectra.
- Collaborated in a multidisciplinary team of optical engineers, chemists, biologists, and computer scientists to translate simulation into experiment and experiment into biosensing platform
- Authored and coauthored three peer reviewed journal articles and five conference proceedings, and an additional three journal articles in press.

- Lab safety manager 2017-2019
- Communicate science to laypeople as a writer for an NPR syndicated podcast, The Loh Down on Science
- Technical skills
    - **Machine Learning:** Python programming, classification/regression, dimensional reduction, data visualization, data augmentation, anomaly detection, feature learning, transfer learning, few shot learning, semi-supervised learning
    - **Nanomanufacturing:** Scanning electron microscopy, electron beam lithography, energy dispersive x-ray spectroscopy, e-beam evaporation, sputtering, nanoparticle synthesis, nanoimprint lithography
    - **Optics:** Raman spectroscopy, dark-field scattering spectroscopy, light microscopy
    - **Analytical Chemistry:** IR spectroscopy, UV-Vis, HPLC, TLC, NMR
- Soft skills: Team leadership, mentorship, public speaking, scientific writing

Bayer Healthcare, Berkeley California **2014**

Associate Engineer

- Manufacturing support in the manufacturing sciences department
- Maintained small scale (15L) bioreactors
- Insured regulatory compliance in manufacturing by standardizing SOPs and interviewing manufacturing staff
- Technical skills: bioreactor maintenance, regulatory compliance

Moore's Cancer Center **2011-2013**

Undergraduate researcher under Dr. Sadik Esener

- Assisted in the development and characterization of ultrasound based drug delivery systems for cancer cells
- Cultured HUVEC cells
- Coauthored one peer journal reviewed article
- Technical skills: cell culture, optical microscopy, fluorescence microscopy

Bayer Healthcare, Berkeley California **2012**

Summer Intern

- Developed a small scale cell separation device (100mL scale) which received an internal patent.
- Technical skills: device engineering

Bayer Healthcare, Berkeley California **2011**

Summer Intern

- Characterized mass transfer characteristics of small scale bioreactors (1L)
- Technical skills: control engineering

Teaching Experience

University of California, Irvine **2014-2018**

- Cbems 241 – graduate level nanomaterials course
    - Ran biweekly office hours and wrote and delivered lectures on nano-optics while professor was at conferences
    - 3.85/4 median teaching evaluation
- Cbems 65a – undergraduate level materials thermodynamics class
    - Wrote and delivered weekly discussion and ran biweekly office hours
    - 3.7/4 and 4/4 median teaching evaluations

Honors and Awards

UCI Graduate Dean’s Dissertation Fellowship	2019
UCI Grad Slam Finalist	2019
UCI Distinguished Public Impact Fellowship	2019
Loh Down on Science Hive Fellowship	2018
Judges Choice TED-style Talk at the UCI AGS Symposium	2018
Research Highlight from the Department of Energy	2018
Research Highlight from UC Irvine	2018
Research Highlight from Nature Nanotechnology	2017
Visiting Scholar at Karlsruhe Institute of Technology	2017
NSF GRFP Honorary Mention	2016
NSF IGERT Fellowship	2015-2017
Henry Samueli Fellowship	2014

Publications

Journal Articles

(1) **Thrift, W. J.**; Nguyen, C. Q.; Darvishzadeh-Varcheie, M.; Zare, S.; Sharac, N.; Sanderson, R. N.; Dupper, T. J.; Hochbaum, A. I.; Capolino, F.; Abdolhosseini Qomi, M. J.; et al. Driving Chemical Reactions in Plasmonic Nanogaps with Electrohydrodynamic Flow. *ACS Nano* **2017**, *11* (11), 11317–11329.

(2) **Thrift, W. J.**; Cabuslay A.; Laird, A. B.; Ranjbar, S.; Hochbaum, A.; Ragan, R. SERS-Based Odor Compass: Locating Multiple Chemical Sources and Pathogens. *ACS Sensors,* **2019**.

(3) Nguyen, C. Q.; **Thrift, W. J.**; Bhattacharjee, A.; Ranjbar, S.; Gallagher, T.; Darvishzadeh-Varcheie, M.; Sanderson, R. N.; Capolino, F.; Whiteson, K.; Baldi, P.; et al. Longitudinal Monitoring of Biofilm Formation via Robust Surface-Enhanced Raman Scattering Quantification of Pseudomonas Aeruginosa-Produced Metabolites. *ACS Appl. Mater. Interfaces* **2018**, *10* (15), 12364–12373.

(4) Darvishzadeh-Varcheie, M.; **Thrift, W. J.**; Kamandi, M.; Ragan, R.; Capolino, F. Two-scale structure for giant field enhancement: Combination of Rayleigh anomaly and colloidal plasmonic resonance. *Physical Review Applied,* **2019**, 11, 054057

(5) Neal, J. A.; Oldenhuis, N. J.; Novitsky, A. L.; Samson, E. M.; **Thrift, W. J.**; Ragan, R.; Guan, Z. Large Continuous Mechanical Gradient Formation via Metal–Ligand Interactions. *Angewandte Chemie International Edition* **2017**, *56* (49), 15575–15579.

(6) Schutt, C. E.; Ibsen, S. D.; **Thrift, W. J.**; Esener, S. C. The Influence of Distance between Microbubbles on the Fluid Flow Produced during Ultrasound Exposure. *The Journal of the Acoustical Society of America* **2014**, *136* (6), 3422–3430.

Conference Proceedings

(7) **Thrift, W. J.**; Bhattacharjee, A.; Darvishzadeh-Varcheie, M.; Lu, Y.; Hochbaum, A.; Capolino, F.; Whiteson, K.; Ragan, R. Surface Enhanced Raman Scattering for Detection of

Pseudomonas Aeruginosa Quorum Sensing Compounds; International Society for Optics and Photonics, 2015; Vol. 9550, p 95500B.

(8) **Thrift, W. J.**; Darvishzadeh-Varcheie, M.; Capolino, F.; Ragan, R. Templated Electrokinetic Directed Chemical Assembly for the Fabrication of Close-Packed Plasmonic Metamolecules. In *Plasmonics: Design, Materials, Fabrication, Characterization, and Applications XV*; International Society for Optics and Photonics, 2017; Vol. 10346, p 103461M.

(9) Nguyen, C.; **Thrift, W. J.**; Bhattacharjee, A.; Whiteson, K.; Hochbaum, A.; Ragan, R. Robust SERS Spectral Analysis for Quantitative Detection of Pycocyanin in Biological Fluids. In *Biosensing and Nanomedicine X*; International Society for Optics and Photonics, 2017; Vol. 10352, p 1035205.

(10) Nguyen, C.; **Thrift, W. J.**; Bhattacharjee, A.; Darvishzadeh-Varcheie, M.; Capolino, F.; Hochbaum, A.; Ragan, R. Self-Assembled Plasmonic Nanogaps: Enabling Early Detection of Biofilm Formation. In *2017 IEEE SENSORS*; 2017; pp 1–3.

(11) Darvishzadeh-Varcheie, M.; **Thrift, W. J.**; Kamandi, M.; Ragan, R.; Capolino, F. Electric Field Enhancement by Two-Scale Structure. In *Conference on Lasers and Electro-Optics (2018), paper JTh2A.68*; Optical Society of America, 2018; p JTh2A.68.

In Press Journal Articles

(12) **Thrift, W. J.**; Ragan, R.; Quantitative Single Molecule SERS Sensing with Convolutional Neural Networks. *Under Review*.

(13) **Thrift, W. J.** Nguyen, D.; Ronaghi, S.; Hochbaum, A.; Ragan, R. Generative Deep Machine Learning for Rapid Antimicrobial Susceptibility Testing with SERS. *In Preparation*

Patents

(1) Regina Ragan, Fillipo Capolino, **William Thrift**, Nano Biosensing System, U.S. Provisional Patent Application No. 62/549,903, 2018

(2) Russell Wong, **William Thrift**, Small Scale Cell Separation Device, Internal Bayer Patent, 2012

Presentations and Invited Lectures

Oral Presentations

"Quantitative label free surface enhanced Raman scattering spectroscopy in the single molecule regime enabled by deep convolutional neural networks", American Chemical Society National meeting, Revolutionizing Chemistry with Artificial Intelligence 2018

"Electrohydrodynamic flow as a driving force for the directed chemical assembly of plasmonic meta-molecules", Invited speaker Karlsruhe Institute of Technology 2017

"Electrohydrodynamic flow as a driving force for the directed chemical assembly of plasmonic meta-molecules", SPIE Nanoscience + Engineering, Plasmonics: Fundamentals and Applications 2017

"Electrohydrodynamic flow as a driving force for the directed chemical assembly of plasmonic metasurfaces " Materials Research Society, Molecular and Colloidal Plasmonics 2017

Poster Presentations

“Illuminating bacterial communities with plasmonic nanoantennas” 2019
Micro- and Nanotechnologies for Medicine

“Illuminating bacterial communities with plasmonic nanoantennas” 2019
Materials Research Society, Materials for Biomedical Applications

“Single molecule quantitative biosensing with self-assembled surface enhanced Raman scattering surface enabled by deep learning”, Gordon Research Conference, Bioanalytical Sensors 2018

“Large Area, Self-Assembled Surface Enhanced Raman Scattering Biosensors Reaching Single Molecule Detection Limits”, Materials Research Society, Colloidal Nanoparticles 2016

ABSTRACT OF THE DISSERTATION

Understanding the State of Bacterial Systems with Machine Learning-Enabled Interpretation of Surface Enhanced Raman Scattering Spectra Produced by Novel Nanomanufacturing

By

William John Thrift

Doctor of Philosophy in Materials Science and Engineering

University of California, Irvine, 2019

Professor Regina Ragan, Chair

Surface enhanced Raman scattering (SERS) spectroscopy is a powerful tool for identifying and quantifying complex mixtures of small molecules, such as the metabolites that indicate the state of a bacterial system. SERS frequently takes advantage of metal nanoparticles to confine light onto their surface, increasing the local electric field. This field enhancement massively increases Raman scattering, and can enable single molecule detection. In this thesis, I introduce a new nanomanufacturing technique called 2-dimensional physically activated chemical self-assembly (2PAC). The focus of 2PAC is to manufacture extremely powerful and uniform SERS sensors, with a SERS enhancement factor of 10^9 that has a relative standard deviation of 10% over a 1 mm x 1 mm area. This technique is characterized by electron microscopy and Raman spectroscopy to elucidate its physical origin. I show that SERS sensor's field enhancement can be further increased by 3-fold by taking advantage of Rayleigh's anomaly using electron beam manufactured optical gratings.

While SERS enables fantastic chemical sensing, it also increases the complexity of analyzing spectra. This is because the ligands involved in 2PAC produce their own spectral signature, and the nanostructures produce a field enhancement that decays rapidly from the hotspot in withhin gaps between nanoparticles. In order to address this complexity, I have developed machine learning techniques that greatly improve analyte concentration regressions. Using a convolutional neural network, I demonstrate quantitative sensing down to 10 fM, well into the single molecule detection regime. I use these methods to address another complex problem, identifying the state of a bacterial system through its metabolome. First, I show that by tracking pyocyanin, a metabolite of *Psuedomonas aeruginosa* (PA), PA can be monitored as it forms a biofilm. PA is detected in just 6 hours, well before it becomes resistant to antibiotics. Then, I demonstrate a semi-supervised method of antimicrobial susceptibility testing (AST) using SERS. This method identifies which antibiotics a bacterium is susceptible to with minimal 24-hour cell culture. AST of PA is performed in just 30 minutes with over 99% accuracy. In sum, this thesis points a way forward to better healthcare through nanotechnology and machine learning.

Part One: Introduction

“In a scientific pursuit there is continual food for discovery and wonder.”

-Mary Shelley, Frankenstein, or the Modern Prometheus

Chapter 1

Introduction

The goal of this chapter is to place this dissertation within the broader context of chemical sensing. Each chapter will include an introduction section that will provide the literature background, unique challenges, and impact of the discussed work. Here, we will see that chemical sensing is a subfield of analytical chemistry, a field with a rich history. We will see that surface enhanced Raman scattering (SERS) is one of many competing strategies for chemical sensing with unique benefits and challenges associated with it. I will show that biological systems have interesting chemical sensing strategies, and that I have used many of these as inspiration for this work. I will show that advances in analytical methods have led to advances in the field of metabolomics, which has had a meaningful impact in diagnosing a variety of health problems. We will see that SERS is a promising technique for metabolomics, and that advances in self-assembly based nanomanufacturing have enabled much of this work. Then, I will discuss the work of others within this field of chemical sensing using SERS. Finally, I will give an overview of this dissertation.

1.1 The History of Analytical Chemistry and Chemical Sensing

Humans have been interested in identifying materials since at least antiquity, and the first recorded chemical identification test (assay) dates back to 1385-1361 BC in ancient egypt. The assay was for the identification of gold. The fire assay, as it is now called, involves putting the suspect material into a furnace, oxidizing the non-gold elements (except for silver). The material would be weighed before and after, and thus the amount of gold was identified.[1] Later, Archimedes would solve the problem of detecting silver in gold by measuring the specific weight of the material. Pliny the Elder made the first record of using a chemical reagent for chemical detection, which detected the adulteration of copper sulphate with iron sulphate. In this method, a strip of papyrus was soaked in an extract of oak apples and then dipped into the suspicious solution. If iron sulphate was present, the papyrus would turn black.[2] In the 4th-5th century the Greek alchemist Synesios introduced the hydrometer to measure a liquid's density.[2]

The first true sensor measurements were performed by Robert Boyle, who contends for the title of first chemist with Jabir ibn Hayyan. In 1674 Robert Boyle performed an experiment to identify the limit of detection of chloride ions through precipitation of silver nitrate. His experiment involved preparing serial dilutions of salted water and adding silver nitrate, taking note of precipitate.[3] Boyle was also a prolific introducer of reagent methods for identifying substances, developing natural extracts for testing acidity and alkalinity. Predating an important contemporary use of chemical sensors – detecting poisons and chemical weapons – Boyle developed a method for detecting poisonous arsenic through the use of $HgCl_2$ as an indicator.[2]

Quantitative chemical analysis began in earnest in the 18th century. First, the blowpipe was popularized as an important quantitative tool for determining the composition of solid materials. The blowpipe was a controlled method of heating samples by blowing a controlled amount of air through a flame. Using a known amount of material, one could quantitatively identify the composition of solid mixtures by melting or burning.[4] Then, the work of Antoine Lavoisier in the 18th century had a marked impact on analysis.[5] Lavoisier popularized the careful measurement of the weights of reagents and constructing chemical balances for reactions. Through this work, he helped develop stoichiometry and the principle of the conservation of mass. As most chemical detection at the time was performed through the use of chemical reagents, this approach enabled many existing chemical detecting methods to become quantitative. Lavoisier also developed early devices for analysis of the composition of organic substances, these devices measured the amount of gas produced by combustion of the material.

The first part of the 19th century marked the beginning of the field analytical chemistry. Jöns Jacob Berzelius identified the atomic weights of several elements, and developed an electrochemical cell which he used to separate and identify ionic compounds.[6] The latter, along with Humphry Davy's work in electrolysis ignited the field of electrochemistry which was used to identify various metals.[2] During this time chemists began using the microscope to examine crystals formed during chemical reactions to identify the product.[2] Titration was also popularized at by Joseph Gay-Lussac, starting the field of volumetric analysis to identify chemicals using indicators.[7] The first comprehensive analytical chemistry textbook was published by Pfaff in 1821.[8] Later, Carl Fresenius founded the first analytical chemistry journal *Zeitschrift für analytische Chemie* in 1862, which was shortly followed by *The Analyst* in 1875, both of which are still publishing today.

1.2 The Development of Modern Chemical Sensors

1.2.1 Light Spectroscopy

For most of the history of analytical chemistry one parameter was observed, or at best quantified in order to identify a chemical. Except for niche applications like the fire assay of gold, this necessitated the use of indicators, which would only indicate upon exposure to the chemical of interest. This all changed in 1860 when Gustav Kirchhoff and Robert Bunsen systematized flame emission spectroscopy and began assigning unique spectral patterns to specific elements.[9] With spectroscopy, many parameters in a system could be measured simultaneously, for example, flame emission spectroscopy yields all of the energy levels of the electronic states in a gas in the visible spectrum. For the first time, chemical assignments could be made with high confidence, without the need for developing a completely new indicator. August Beer would turn optical spectroscopy into a quantitative technique through his observation of the relationship between light absorption and analyte concentration.[10] Beer also developed the color comparator which enabled researchers to quantitatively compare optical spectra. The color comparator would evolve into the UV-visible spectrophotometer (UV-vis), which will be used in this work in Chapter 3 for characterizing the extinction spectral peaks of optical antenna. The UV-vis is a powerful tool for measuring electronic transitions from bonding to antibonding orbitals, which are often in this energy regime. This technique is excellent for chemical sensing for three reasons: 1) many of these transitions exist for any given molecule 2) the transition energy is determined extremely sensitively by the configuration of atoms in a molecule 3) electronic transitions have a characteristic absorption cross-section, enabling precise quantification.

Infrared (IR) radiation had been discovered as early as 1800 by John Herschel, who used a prism and blackened thermometers to measure adsorption beyond the visible spectrum.[11] Research into the IR spectrum began in earnest after the bolometer was invented in 1878, which measured temperature changes very sensitively through an electrical resistance.[12] Using this device Willem Henry Julius published the IR absorption spectra of 20 organic molecules.[13] IR spectroscopy is a powerful tool for chemical sensing because it interrogates the vibrational models available in a molecule. Molecules are extremely complex, and have potentially hundreds of measurable normal modes, overtones, and superpositions. Compared to electronic transitions interrogated with UV-vis, vibrational modes have extremely narrow linewidths which allow many more modes to be resolved for IR spectroscopy than for UV-vis. One region of particular interest is the so-called “fingerprint” region of the IR spectrum. This region is densely occupied with vibrational modes, and nearly any molecule can be identified by observing its modes in this region.[14]

Raman scattering spectroscopy also interrogates vibrational modes in materials. Raman scattering was discovered in 1927 by Sir Chandrasekhara Venkata Raman, and was important evidence for quantum mechanics, which as a nascent field at the time.[15] Raman scattering spectroscopy is extremely important to this work, and it gets a full theoretical treatment in Chapter 2, as it will be used in nearly every chapter of this thesis. Raman scattering spectroscopy probes molecular vibrations through the inelastic collision of light (usually in the visible or near IR) with vibrational modes. Energy is transferred from the high energy light into the molecule's vibrational modes, and the energy shift, called the Raman shift, is measured as a spectrum in energies near the incident wavelength. The Raman spectrum probes the same energies as IR spectroscopy, but a key difference between the two makes them complementary: selection rules. IR spectroscopy can only detect asymmetrical vibrations because a dipole moment must be formed, while Raman scattering spectroscopy can only detect symmetrical vibrations because the polarizability of the molecule must change.[14]

For a long time after Raman's discovery his spectroscopy was not used in practical applications. This is because Raman scattering is a nonlinear optical phenomenon (it relies on light's interaction with electrons), and light-matter interactions are extremely weak, resulting in very low Raman scattering cross-sections.[16] This made the invention of the laser in 1960 by Theodor Maiman[17] particularly important for Raman scattering spectroscopy. The laser enabled huge optical power densities to be achieved, greatly increasing the amount of Raman scattering that could be observed. When operated with modern lasers, microscopes, optical gratings, and detectors, Raman scattering spectroscopy has several key advantages over IR spectroscopy. First, the environment has much less of an effect on the measured spectra for Raman scattering, this is because both CO2 and water are broad and strong IR absorbers, while they have extremely small

Raman scattering cross-sections. Second, visible detectors and gratings currently have significantly greater resolution than IR detectors and gratings, so much more detailed spectra can be acquired. Finally, visible light has a much smaller wavelength than IR light, so Raman scattering spectroscopy can be used for microanalysis, where this is impossible except in sophisticated (usually near-field scanning optical microscopy) academic microscopes.

1.2.2 Mass Spectrometry

In principle, mass spectrometry (MS) uses an electric field to push an ionized atom or molecule via Lorentz force. The displacement of the ionized analyte is measured by a detector and the mass is calculated. In practice, MS instruments are among the most complicated, and most powerful chemical sensors available with sophisticated and diverse separation tools, ion sources, mass analyzers, and detectors.

MS began with the discovery of cathode rays in 1858 by Julius Plucker, who found that fast moving, negative charges were emitted when an electric current was passed through a vacuum.[18] These charges would later be identified as electrons. Then, in 1886 Eugen Goldstein and Wilhelm Wein discovered that if a perforated cathode was used, new, positively charged rays would be produced.[11] These positive charges were ions formed when the electrons ionized some of the gas molecules in the vacuum tube. The study of cathode rays became a method of analysis by 1912 when J. J. Thomson constructed a parabola mass spectrograph, a forerunner of the mass spectrometer, to analyze the deflection of positive ions and thus identify their mass-to-charge ratio. Thomson used this information to produce the first mass spectrum and to identify the signature of the air components in the vacuum tube (water vapor, nitrogen, oxygen, etc).[19] Even as fresh as MS was at the time, Thomson was able to use the tool to prove the existence of isotopes.[11] Much later, in the 1950s, Roland Gohlke developed the use of MS as the detector in gas chromatography, enabling the components of complex mixtures to be identified and quantified.[20] After the popularization of high performance liquid chromatography (HPLC) in the 1960s, MS would often be used as HPLC's detector as well.

Initially, electron ionization (EI) was the only ionization method used in MS, but this method is extremely harsh and can break molecules down into fragments that are too small to give useful information for inferring their identify. Today many soft ionization methods now exist, including chemical ionization (CI), electrospray (ES), and matrix-assisted laser desorption/ionization (MALDI), among others. CI was the first developed (in the 1960s) and consists of a reagent gas, usually methane, which is ionized by EI and is used as a proton donor to the analyte molecule.[21] The resulting positively charged analyte ions are more stable and the mass of the entire molecule can be more easily inferred. ES was an early method (developed by Malcolm Dole in 1968[22]) for analyzing polar and ionic biomolecules, and uses a jet of analyte dissolved in solvent (called a matrix), which aerosolized and eventually produces ionic analyte. MALDI is a more recent method, being developed in 1985 by Franz Hillenkamp, Michael Karas.[23] MALDI uses a laser to vaporize and ionize analyte on a surface. This enables microscopic imaging based on MS, and has made MALDI one of the most popular ionization methods in use today. Another important development has been the tandem MS, where a soft ionization technique like ES or MALDI are used, the masses are detected, and then the ion beam is fragmented with hard ionization and the masses are measured again. This enables structural information to be obtained from the fragments, from which the molecule can be identified.[24] Tandem MS is an important reason why MS is considered by many to be the best chemical sensing method available today.

1.2.3 Nuclear Magnetic Resonance Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is based on the resonant frequency of the spin switching of (typically hydrogen) nuclei. The key to using NMR for chemical identification is that this resonant frequency shifts slightly based on the local atoms and functional groups around a hydrogen. NMR produces a spectra of resonant frequencies of all the hydrogen nuclei in a molecule and this can be used to infer the identity of the molecule. By using adding a known amount of a standard molecule, the concentrations of the various molecules within a mixture can also be identified, although the number of components must be small as NMR is not usually used with separation techniques. Throughout the second half of the 20th century, NMR was considered the premier chemical sensor and remains one of the most important tools in chemistry for chemical sensing. Yet, NMR, like IR spectroscopy, requires high concentrations of analyte molecules, and for this reason is slowly losing ground to MS systems.

The nuclear magnetic moment was first postulated by Wolfgang Pauli in 1924,[25] and then later experimentally verified by Isidor Rabi in 1938.[26] Felix Bloch and Edward Purcell then independently developed a method to measure nuclear magnetic moments in liquids and solids, respectively.[27,28] The phenomenon of chemical shift, the change in resonant frequency depending on the nuclei's local environment, was then discovered by Yu and Proctor in 1950.[29] Quickly after this discovery, NMR was commercialized with instruments being produced by the Varian Corporation for oil companies to identify the relative amounts of organic molecules in gasoline.[11] Modern advances in NMR include 2D NMR, such as correlation NMR, which uses spin correlation between nearby nuclei to improve the differentiability of similar molecules and molecules within mixtures. Another important area of active research is solid state NMR, which uses magic angle spinning to identify the structure of solids. This technique has promise in combination with X-ray crystallography for determining tertiary and quaternary protein structure, an extremely challenging problem.

1.2.4 Electrochemical Sensors and Immunoassays

After reading the prior three sections, one would be forgiven for thinking that reagent methods in analytical chemistry were dead. In fact, nothing could be further from reality. I will typically refer to reagent methods as either 1) labeled chemical sensing (or more likely I will refer to label-free chemical sensing as the opposite of this method) if one observes the transformation of some reagent, such as the (un)quenching of a fluorophore or redox of a small molecule, as the measurement readout, or 2) capture-agent chemical sensing (again more likely I will discuss label-free, capture-agent-free methods that are the focus of this work) where some agent, typically an antibody (aptamers and chemical crosslinking are also popular) captures the analyte molecule and this event is the measurement readout. Labeled and capture-agent based reagent methods are alive and well. Glucose sensors are the most important biosensor today, accounting for 71% of the total biosensing market, and accounting for 15.6 billion dollars in value. Glucose sensors are labeled electrochemical sensors, using an oxygen cosubstrate enzyme to generate hydrogen peroxide. The enzyme, FAD, oxidizes upon reaction with glucose to form reduced FAD (FADH), which then reacts with oxygen and water to form hydrogen peroxide. The hydrogen peroxide is then measured electrochemically with an amperometric measurement (discussed below). In addition to electrochemical sensors, other labeled or capture agent based sensors are extremely popular, such as enzyme-linked immunosorbent assays (ELISA), and surface plasmon resonance (SPR) based immunoassays.

Electrochemical sensing began in 1909 with a groundbreaking experiment by Franz Haber. 3 years prior, it was discovered that glass lamellas were pH-sensitive membranes; Haber characterized electrical the potential difference between the fluid on both sides of the membrane, producing the first potentiometric electroanalytical measurement.[30] The next development came in 1922 when Jaroslav Heyrovský developed what he called polarography (later renamed voltammetry), where drops of mercury were dropped into an analyte solution and the electrolytic current was measured as a function of voltage applied.[31] This technique was neither quantitative nor time efficient, so electrochemical sensors remained relatively obscure until the 1960s when differential pulse voltammetry[32] and electrochemical stripping analysis[33] enabled the detection of trace metals in the high parts per trillion concentration regime, an excellent limit of detection, even by today's standards. Then in 1973, the modern enzyme-based glucose electrochemical sensor was developed by Guilbault and Lubrano.[34] This progress continued through the 1980s with the development of different ion selective membranes, alternative electrode materials, and surface functionalizations of electrode materials.[30] Electrochemical sensor research continues to flourish, driven mainly by new developments in nanotechnology based electrode materials, and continued development of enzymes, antibodies, and aptamers for detecting new analyte molecules.

Immunoassays are also extremely popular, especially in the pharmaceutical industry. ELISA was developed in 1972 by Engvall and Perlmann[35] and uses antibodies as capture-agents and enzymes to generate labels which enable the detection of analyte molecules. In the technique, analyte is placed on a surface and a recognition antibody is added to conjugate with the analyte, the antibody is conjugated with an enzyme which produces a color change by reacting with some reagent. This color change is measured, and produces a quantitative signal. Because the signal can be allowed to grow (the enzyme will continue to produce the color change as long as there is reagent), ELISA has a huge gain, and can be used for single-molecule detection.[36] In the section below, we shall see that this strategy resembles chemical sensing strategies performed by nature.

Another important immunoassay based method is that of SPR, which was developed by Liedberg, Nylander, and Lundström in 1983.[37] A theoretical treatment of the phenomenon in the context of localized surface plasmon resonance will be given in section 2.1. The important feature of SPR in chemical sensing is that the phase-matching condition for exciting a surface plasmon is extremely narrow, and extremely dependent on the local (~10 nm) region around the gold film in which the surface plasmon propagates. This enables researchers to measure changes in the local dielectric function near the gold film extremely precisely. In SPR sensing, analyte is placed on the film and a capture-agent is introduced onto the film. When the capture-agent recognizes the analyte molecule, the dielectric function changes, and this is measured by the absorption into the SPR.[38] SPR was commercialized in 1990 by Biacore.

1.3 Chemical Sensing by Bacteria

Long before humans, single-celled organisms pioneered chemical sensing. The perception and processing of chemical signals is likely the first sense developed by life.[39] Chemosensing nutrients and toxins provides an enormous advantage for survival, and bacteria have developed chemical sensing methods that are still superior in many respects to artificial methods. Thus, scientists have a lot to learn about chemical sensing from bacteria.

Life transduces chemical signals through chemoreceptors, proteins that undergo a structural change when a small molecule docks in the receptor. The structural change results in signal through one of two mechanisms. 1) ionotropic reception where the structural change enables an ion flux that either results in a change in the cell membrane potential or results in an increase of intracellular ion concentration which leads to signalling. 2) metabotropic reception where the structural change increases enzymatic activity of the receptor and results in the production of secondary signalling molecules. The chemoreceptor strategy is effective because the signal is amplified, e.g. for metabotropic reception many secondary signalling molecules are produced for each molecule recognized. The process is also long lived, which makes the process robust to noise, in spite of the large gain, similar to ELISA.

Biological chemoreception is extraordinarily powerful. *Escherichia coli* (EC) can detect amino acids at nanomolar concentrations.[40] EC and *Salmonella typhimurium* are capable of detecting a chemical signal in 0.1-0.5 seconds, and *Dictyostelium discoideum* can detect concentration differences of cAMP that are just 5% over its 10 μm diameter.[41] Bacteria's ability to detect small concentration differences for chemotaxis serves as the inspiration for Chapter 6.

Besides identifying nutrients and avoiding toxins, chemosensing serves another important purpose: communication. Bacteria use chemical signals to trigger gene regulation in their communities in a process called quorum sensing. Quorum sensing enables decentralized decision making to be performed in bacterial communities which makes them much more robust. For example, EC uses the molecule AI-2 to regulate cell division and thus population density.[42] Quorum sensing also enables intraspecies warfare, *Escherichia carotovora* uses quorum sensing to trigger the production of antibiotics to fend off other bacteria.[43] Also consider *Psuedomonas aeruginosa* (PA) which uses the small molecule pyocyanin as a virulence factor, eg. it uses pyocyanin as a signal to colonize its host.[44]

Quorum sensing in bacteria points to an important application of chemosensing: detection of bacteria. Chapter 5 is devoted to detecting pyocyanin in order to identify PA before it forms a biofilm (a robust community of bacteria on a surface), but this strategy is not new. Bacteria have been using quorum sensing to detect bacteria for millennia, for example *Salmonella enterica* uses quorum sensing to identify pathogenic bacteria within the human gut microbiome.[45] Viruses too use this strategy. The bacteriophage VP882 hides in *Vibrio cholerae* (VC) DNA until it senses a quorum sensing molecule that VA produces when it reaches a high cell density. VP882 then emerges from its latent state and begins reproducing.[46]

1.4 Olfaction as a Medical Device

Multicellular organisms have evolved an even more sophisticated chemical sensing method: the sense of smell (olfaction). Olfaction works as follows: First, you inhale air into your nose. Within the air, there may be odor molecules which can be smelled. These molecules are small, volatile, and hydrophobic. Incidentally, the nose is shaped to concentrate these odor molecules to improve limits of detection, a phenomenon that has been mimicked using a 3D printed nose and a mass spectrometer.[47] Once inside your nose, odor molecules eventually reach the olfactory epithelium, mucous covered patches of olfactory neuron dendrites. These dendrites are covered in olfactory receptors, of which humans have approximately 20 million. Signals from the olfactory neuron dendrites are then propagated to the olfactory bulb where glomeruli and mitral cells produce a signal that is interpreted by the brain as a smell.

Olfaction has several properties worthy of biomimicry, but in this thesis two stand out as the most important. 1) Unknown odor molecules may be recognized by olfaction. 2) Complex mixtures of small molecules can be understood and differentiated from individual types of small molecules and other complex mixtures. These two properties are very different – and much more powerful – than typical assays that are developed where a receptors (eg an antibody or an aptamer) is developed to recognize an individual type of small molecule. To achieve this, olfaction uses hundreds of different receptors.

Olfactory receptors (ORs) are a family of G protein-coupled receptors expressed on cell membranes. In 2004 Linda Buck and Richard Axel won the Nobel Prize in Physiology or Medicine for their discovery of ORs.[48] The OR system of recognizing molecules is powerful because individual ORs are not made to identify specific small molecules. Rather, they tend to identify functional groups on molecules, and the relative size of these molecules. When one type of molecule enters the olfactory epithelium many different ORs will be activated (via a metabotropic process). The olfactory bulb then uses the pattern of activated ORs to uniquely identify the odor.

The pattern recognition element of olfaction is the inspiration for this work's use of machine learning, which is developed in Part 3 of this text, and is why new molecules can be recognized: they will produce a novel pattern of OR activation. This system also produces some interesting unintended consequences such as smells changing as the concentration increases, for example, grapefruit mercaptan smells of grapefruit at low concentrations, but skunky at high concentrations.[49] The pattern recognition system of olfaction is also the cause of the second property described above. Complex mixtures can be identified and differentiated because odors are perceived by the pattern of OR activation. In this sense olfaction is much like the sense of sight: we perceive odors as objects. In fact, even trained perfumers are unable to recognize more than 3 different molecules simultaneously, yet they can remember tens of thousands of different perfume mixtures.[50]

Human olfaction is capable of some remarkable feats. Humans have about 300 unique ORs, and typical limits of detection of odors in the parts per million range, with some molecules that can be smelled in parts per trillion concentrations.[51] Humans are capable of smelling and identifying fear in sweat.[52] One controversial study calculated that humans can differentiate 1 trillion different odor objects.[53] Humans can smell the age of other humans.[54] Women have been shown to identify immune system health and differences in immune system genetic composition of men by smell.[55]

While humans have used their sense of smell to detect diseases like *Mycobacterium tuberculosis* (TB) since 2000 BC,[56] when it comes to the healthcare applications of olfaction dogs reign supreme. Scientists have trained dogs to detect bacteria, such as *Clostridium difficile,*[57] TB,[58] and others.[59] Dogs can detect colorectal cancer,[60] prostate cancer,[61] bladder cancer,[62] and others.[63] They can also be used to monitor their diabetic owner's blood sugar levels.[64,65] In fact, dogs were recently used to demonstrate that epileptics release a small molecule before having a seizure.[66] Dogs and humans are capable of these impressive healthcare diagnoses because they sense the diverse set of small molecules that are produced in these disease states. Put another way, they use metabolomics.

1.5 Metabolomics

At the core of biology is the central dogma of molecular biology.[67] The dogma concerns how information flows, loosely put, the central dogma is that DNA information is copied into RNA transcripts, RNA is used to synthesize proteins, and proteins produce and use small molecules called metabolites. The scientific community is very interested in knowing all of the information as it is passed along, and the complete set of each of these elements, DNA, RNA, proteins, and metabolites, all have an -omic associated with them: genomics, transcriptomics, proteomics, and metabolomics, respectively. The grand plan of what to do with all of this information is called systems biology, the study of all of the interactions between all of the constituents of a biological system.[68] The idea is that if we know all of the elements in a certain biological system, say the glucose-insulin system, we could gain a mechanistic understanding of how to identify new drugs for diabetes. Even as of 2019 when this work was written, this level of mechanistic understanding is far away. Yet just gathering an individual's -ome has enormous potential for personalized medicine, with genomics analysis giving immediate diagnosis of dangerous diseases like phenylketonuria,[69] and can be used to identify cancer risk.[70,71]

Of these -omic's metabolomics is the most recent to be formalized as a scientific discipline. This is primarily because of challenges in chemical sensing, metabolomic profiling requires scientists to analyze a huge diversity of molecules, often in extremely small concentrations, and it requires reasonably good quantification of these metabolites to draw useful conclusions. For this reason, the history of metabolomics and the history of analytical chemistry are inextricably linked (discounting the development of humoralism by Galen). The concept of a metabolic profile that could be used to identify disease states was introduced in the 1940s by Roger Williams,[72] who used the then new tool of paper chromatography[73] to try to show that schizophrenics had a different metabolic profile than non-schizophrenics. Shortly after the development of the GC-MS discussed in section 1.2.2, the concept of the metabolic profile was cemented in the 1970s by rigorous, quantitative measurements of steroids, drugs, and acids in urine.[74] NMR was also rapidly advancing in the 1970s, and soon became the quintessential tool in metabolomics. For example, in 1974 metabolites in in-tact tissues were measured with NMR, demonstrating that 90% of ATP in the body is complexed with magnesium.[75] Then, the development of electrospray MS enabled major developments in metabolomics, like the discovery that oleamide has sleep inducing properties.[76] Throughout this time, metabolomics was not a field of research in its own right; that changed with the development of tandem MS, which enabled quantitative analysis and identification of complex mixtures of diverse and low concentration metabolites. The first metabolomics database, METLIN (characterized with tandem MS), was published in 2005 by Gary Siuzdak.[77] METLIN now contains over 450,000 metabolite spectra.

The field of metabolomics has been able to recreate many of the olfaction-based diagnoses described in section 1.4. In cancer, for example, metabolomics has been used extensively to diagnose colorectal,[78] lung,[79] oral,[80] head and neck,[81] prostate,[82] bladder,[83] pancreatic,[84] breast,[85] liver,[86] and thyroid[87] cancers. Autoimmune diseases like diabetes,[88] arthritis,[89] asthma,[90,91] autoimmune hepatitis,[92] Crohn's disease,[93] and coeliac disease[94] have all been diagnosed with metabolomics. A variety of infections have been diagnosed with metabolomics such as *Clostridium difficile*,[95] TB,[96] PA,[97] *Helicobacter pylori*,[98] and EC.[99] The vast majority of these studies were carried out using some form of MS, which has come to dominate the field.

MS has been an extremely useful and successful tool in metabolomics. Yet, MS has not become the ubiquitous diagnosing tool that a general metabolome profiling method has the potential of being. This is for a number of reasons: 1) Cost. A modern tandem MS instrument is extremely expensive, with instrument costs in the 100,000 dollar range. 2) Sample preparation and chromatography. Sample preparation is extremely important in MS, slowing the process, and leading to irreproducible results. Chromatography must also be involved to get meaningful results from MS, which is relatively time consuming if one wants a diagnosis during a doctor's visit. 3) Limits of detection in the 10s of parts per billion[100] (ppb). Many metabolites are in the low ppb to parts per trillion concentration range,[101] which is currently unattainable with MS without preconcentration of analyte.[102] We have seen that advances in metabolomics naturally follow advances in chemical sensing. For this reason, this thesis seeks to advance the chemical sensing technique: surface enhanced Raman scattering (SERS) spectroscopy to tackle challenging problems in metabolomics.

1.6 Surface Enhanced Raman Scattering (SERS)

SERS has a tortuous, fitful history that has been driven as much by serendipity as by advances in fundamental knowledge. In the 1970s, electrochemists were interested in using spectroscopic methods to measure electrochemical reactions in real time. Importantly, these reactions occurred at electrode surfaces, so only a monolayer of the analyte molecules are available for measurement; the reactions occur in aqueous media, so IR spectroscopy couldn’t be used effectively due to the intense IR absorption of water. Given these considerations, the electrochemists turned to Raman spectroscopy.[103] In retrospect was likely against their better judgement as the Raman scattering cross-section of most molecules is on the order of 10^{-30} steradians,[14] and as a result is generally not a useful tool for measuring monolayers, let alone in real time. Yet, owing to the unusually large Raman cross-section of mercury ions, redox reactions involving Hg_2Cl_2 and Hg_2Br_2 were measured by Martin Flieschman in 1973.[104] Then, in 1974, Flieschman inadvertently discovered SERS after he roughened silver electrode, with the goal of increasing the number of molecules in the laser spot size.[105] Unbeknownst to him, localized surface plasmon resonance (LSPR) of the jagged silver electrode had dramatically increased the local field enhancement of the pyridine monolayer he had been investigating, enabling his measurement. The unusually large Raman scattering intensity was noted by the Van Duyne group in 1977,[106] and then the LSPR enhancement mechanism was proposed by Martin Moskovits in 1978.[107] The technique was then expanded for use with colloidal silver nanoparticles in 1980.[108]

The 1980s and early 90s saw the first wave of excitement for SERS as a tool in chemical sensing, but ultimately the technique wouldn’t take off then due to challenges in nanomanufacturing. Owing to the fact that primary and secondary amines chemisorb onto silver surfaces and the simultaneous genomic revolution, nucleic acids were among the first biomolecules characterized with SERS.[109–111] Similarly, peptides and amino acids were also subject of intense interest.[112–114] While the concept of the SERS "hotspot" that would push the field into nanogap based surfaces and into measuring small molecules had not yet been developed, the groundwork for contemporary SERS sensing was also laid out, with groups using the technique for explosives,[115] drugs,[116] pesticides,[117] and neurotransmitters.[118] Yet, despite these works, it became increasingly clear that SERS (as it was then) was not useful for quantitative chemical sensing due to the incredible variation of the SERS enhancement from SERS substrates. For example, the roughened silver electrode that was in vogue at the time generated 24% of the SERS signal from just .0063% of the surface,[119] leading to wild variations in signal that prevented quantification of analyte concentration. Similarly, SERS signal from colloid which arises due to salt-induced aggregation of the nanoparticles is highly nonuniform due to nonuniform nanoparticle assemblies and gap spacings that are produced.[120]

SERS was revitalized in the late 90s when two groups independently demonstrated single molecule detection.[121,122] This was a momentous achievement for SERS, making it one of the few chemical sensing methods capable of the ultimate limit of detection, and with rich spectral information that enables chemical identification to boot. Single molecule SERS will be discussed in full in Chapter 7. SERS also benefited from related interest in surface plasmon polaritons (SPPs) that started in 1998 with Thomas Ebbesen's discovery of extraordinary optical transmittance through perforated metal films,[123] and grew to a fever pitch with the development of plasmonics, devices based on propagating SPPs and LSPR.[124,125] The increased attention these advancements lent to optical nanomaterials led to a boom in nanomanufacturing techniques, particularly self-assembly, which would drive much of the development of SERS.

Inspired by contemporaneous advances in self-assembled monolayers (SAM),[126] the first self-assembled SERS surface was developed in 1995 by the Natan group.[127] The group used a straightforward strategy: apply an amine or thiol terminated SAM onto a surface, then add a colloidal nanoparticle solution. The nanoparticles would chemisorb onto the SAM and be immobilized. The resultant surfaces produced extremely uniform SERS surfaces by the standards of the time, and the technique was widely used and expanded upon.[103] For the next 10 years, a large number of works using SAM immobilization of nanoparticles and salt-induced aggregation of colloidal nanoparticles would be published. Unfortunately, even as late as 2005 researchers in the field were not sufficiently aware of the critical role that small gap spacings play in achieving good SERS surfaces. This can be observed, for example, in a well-regarded review by Baker and Moore where they state, "several authors have claimed the importance of junctions between neighboring nanoparticles as being key to single-molecule sensitivity in SE(R)RS… The exact mechanism remains obscure, however."[128] The importance of small gap spacings was gradually realized in the late 2000s with the development of nanoparticle on mirror geometries with tunable gap spacings,[129] and the development of the plasmon hybridization model.[130] A thorough background of self-assembly methods since 2005 will be given in section 3.1.

The Ragan lab entered the nanoparticle assembly space in 2009 with J H Choi's work, "Design of a versatile chemical assembly method for patterning colloidal nanoparticles".[131] This work was motivated by the plasmonics community, which had been interested in nanostructure arrays for novel optical phenomena.[132–134] The authors sought to expand upon preliminary work that demonstrated nanoparticle assembly on block copolymer templates,[135,136] with the goal of transferring the complex pattern provided by the block copolymer to the nanoparticle assemblies. Importantly, instead of using the SAM assembly techniques used by Natan and others, this work used the bioconjugation technique of carbodiimide coupling to link the nanoparticles onto a SAM on the copolymer surface. Later, the processing conditions were improved by Sarah Adams who used the technique for SERS, and studied the importance of discrete nanoparticle assemblies and gap spacings for SERS enhancements.[137] Then, motivated by contemporaneous research into the use of colloidal driving forces for self-assembly nanomanufacturing,[136,138–140] Adams incorporated electrophoretic deposition into the assembly procedure that she developed earlier.[141] Interestingly, Adams observed dramatically increased assembly of Au nanospheres into discrete clusters with small gap spacings, which are ideal for SERS surfaces. While the assembly mechanism was not understood at the time, in Chapter 3 I will demonstrate that Adams had in fact serendipitously discovered a method for driving the carbodiimide crosslinking between Au nanoparticles using electrohydrodynamic flow.

1.7 Chemical Sensing with SERS

We saw in the previous section that as early as the 1980s and 90s SERS was actively pursued for chemical sensing applications, particularly for sensing warfare agents and poisons. SERS has several features that make is a promising platform for chemical sensing: 1) Like MS and NMR, it is a label free method, so the constituents of complex fluids can be measured without the need for developing capture agents. 2) SERS can achieve single molecule sensing, even for non-resonant molecules,[142] which enables sensing at ultralow concentrations, with a limit of detection of 1 part per trillion being demonstrated for label-free single molecule sensing[143], much lower than either MS or NMR. 3) Vibrational spectroscopy is an excellent method for molecular identification or for the discrimination of complex mixtures of molecules. The energy states typically probed from within the “fingerprint” region of vibrational spectroscopy contains a huge amount of information, and it parallels natural olfaction in important ways. Both use about hundreds of overlapping but unique parameters that capture details about molecular structure (energies for SERS and olfactory receptors for natural olfaction). While neither use chromatography like MS (except for some notable lab-on-a-chip implementations for SERS[144]), we have seen in section 1.4 that this has not limited the metabolomics applications of natural olfaction.

A large amount of work has been done using SERS for bacteria detection. Due to interest in single cell sensing over the last two decades, much of this work has been performed by simply adding nanoparticle colloids to bacterial samples.[145–150] While discrimination of bacteria can be achieved with this method, the large variations of signal enhancement for colloidal SERS prevents this from being a quantitative method. Therefore, manufactured SERS surfaces are also commonly used for bacteria detection.[151–154] Antimicrobial resistance has also been a subject of interest, with groups using recognition elements to discriminate between resistant and susceptible bacteria strains.[155,156] Analyzing urinary tract infections (UTIs) is particularly interesting for SERS, in part because UTIs are amongst the most common hospital acquired infection,[157] but also because urine can be directly analyzed with SERS surfaces with no sample preparation. Several groups have reported UTI bacteria identification and discrimination with SERS.[158–160] Metabolomics applications have also been recently explored for detection of bacteria,[161,162] and for cancer diagnosis.[163–165]

Much of the recent progress in chemical sensing with SERS has been driven by machine learning tools. Principal component analysis, an important dimensional reduction tool, has been with SERS since the early days for spectral differentiation and has been used to demonstrate single molecule SERS,[166] detect nasopharyngeal cancer,[167] colorectal cancer,[168] gastric cancer,[169] and cervical cancer.[170] Linear[171–173] and multilinear[174] regressions have also been popular for concentration quantification using the langmuir adsorption curve,[175] where multilinear regressions are far preferred because they take advantage of the rich spectral information acquired with SERS. The ongoing machine learning revolution has begun to impact SERS, with several groups using modern techniques to analyze SERS spectra. Support vector machines, which provide excellent spectral discrimination and concentration regressions even with small datasets, have been used for diagnosis of prostate cancer[176] and detection of drugs in urine.[177] Decision trees have been used to differentiate different metabolites in blood,[178] differentiation of bacteria,[179,180] detection of gastric cancer,[181] and detection of pesticides,[182] while hiearchical cluster analysis has provided an unsupervised method of finding spectral similarities in applications like bacteria[183] and virus[184] differentiation, detection of chemical warfare agents,[185] and detection of food contaminants.[186] Fully connected artificial neural networks have been used for multiplexed concentration quantifications,[187] and convolutional neural networks have been used to differentiate metabolites produced by cells *in vivo*.[188] While these techniques have been implemented, the machine learning community has been fast at work, developing deep machine learning, generative methods, and sophisticated data visualization tools like non-negative matrix factorization and t-distributed stochastic neighbor embeddings. In this thesis I will show how these new tools will enable SERS sensing for advanced metabolomics applications.

1.8 Dissertation Outline

This dissertation is divided into three parts: I) Introduction (Chapters 1 and 2), II) Nanomanufacturing Large-Area, High Performance Optical Devices (Chapters 3 and 4), and III) Machine Learning Enabled Biosensing with Surface Enhanced Raman Scattering Spectroscopy (Chapters 5-9), with a brief conclusion given in Chapter 10. Chapter 1 concerned itself with giving an extensive historical and technical background as well as the motivation for the dissertation.

Chapter 2, Background, gives an introductory, theoretical background for important and complicated tools used in this work, namely LSPR, SERS, electrohydrodynamic flow, and machine learning.

Chapter 3, Driving Chemical Reactions in Plasmonic Nanogaps with Electrohydrodynamic Flow, is the successor to Sarah Adams' "Directing Cluster Formation of Au Nanoparticles from Colloidal Solution". In it, I demonstrate that the clustering behavior she observed is actually carbodiimide crosslinking between the ligands on Au nanospheres. This is shown through SERS spectral analysis, transmission electron microscopy, and molecular dynamics simulations. In order to resolve how nanospheres are selectively crosslinked at the electrode surface yet not in bulk, I propose an electrohydrodynamic (EHD) flow mechanism. This mechanism is elucidated by varying the sphere size, temperature, and voltage. I leverage this understanding to improve the SERS performance of the manufactured sensors. With electromagnetic simulations compared to real data, I show that the close-packed assemblies produced with this method have a remarkably uniform optical response. Finally, I show that the SERS enhancements are among the largest, most uniform, and over the biggest area of any SERS surface ever made. This new and improved assembly method is dubbed a 2-dimensional physically activated chemical (2PAC) self-assembly, and is used for all of the following chapters.

I conclude Part 2 with Chapter 4, Templated Electrokinetic Directed Chemical Assembly for the Fabrication of Close-Packed Plasmonic Metamolecules. The goal of this chapter is to show that the 2PAC method developed in Chapter 3 is broadly applicable to nanophotonic applications. In it, I show that 2PAC can be reformulated as a directed self-assembly method. Here, templates composed of Au nanopillar arrays manufactured with electron beam lithography are used to drive EHD flow. I show that coating these pillars with SAMs for carbodiimide crosslinking, and the surface with an omniphobic SAM, assembly of circular, close-packed Au nanosphere assemblies can be produced. Electromagnetic simulations of the optical response of these as-manufactured assemblies demonstrate that these assemblies provide magnetic field enhancements which are highly unusual for nanophotonic structures.

Part 3 begins with Chapter 5, Longitudinal Monitoring of Biofilm Formation via Robust Surface-Enhanced Raman Scattering Quantification of *Pseudomonas aeruginosa* - Produced Metabolites. This chapter provides the basis for the machine learning-driven SERS-based metabolomics done for the rest of the thesis. In it, pyocyanin, a metabolite of PA is measured as a proxy of PA growth. In section 1.3 we saw that pyocyanin is a virulence factor and is thus well suited for this application. This is also a particularly desirable system for fundamental studies such as these due to the fact that pyocyanin is a fluorescent molecule with an electronic transition near 785 nm, the laser used for SERS. This simplifies the SERS analysis by ensuring that all spectra will be dominated by pyocyanin. The chapter begins by acquiring SERS spectra from concentration gradients of pyocyanin in water and cell media and training a partial least squares regression model to map these SERS spectra to concentration values. With the trained model, PA is monitored as it grows on a biofilm, a planar community of bacteria. The model identifies the PA as early as 3 hours, with good concentration results at 9 hours. Importantly, this is shown to be less time than that which PA needs to become resistant to an antibiotic treatment.

Chapter 6 continues SERS-based sensing of bacterial biofilms, and is titled, SERS-based Odor Compass: Locating Multiple Chemical Sources and Pathogens. In it, I propose the SERS-based odor compass, a device that detects the direction from which an odor source (a volatile small molecule) emerges. The device is composed of a grid of 2PAC manufactured SERS sensors, which are placed nearby the odor source. The chapter seeks to solve the multiple odor source problem for the first time, and show the first use of an odor compass for identifying odor sources of biological origin. I show that non-negative matrix factorization (NMF) is an ideal tool for decomposing SERS spectra into parts that correspond to different odor sources. Then, I show that various machine learning classifiers achieve good accuracy in solving the 2 odor source problem. Finally, I demonstrate that the SERS-based odor compass can identify the location of an EC biofilm growing nearby.

Chapter 7, Improved Concentration Regressions with Convolutional Neural Networks for Surface Enhanced Raman Scattering Sensing, seeks to explore the convolutional neural network (CNN) as a tool for analyte concentration regressions with 2PAC manufactured SERS sensors. The chapter shows that the CNN is a superior tool for concentration regressions compared to the fully-connected artificial neural network or the multilinear regression. I explore the robustness of the CNN model to interpolation, and show very positive results. Finally, I show that the CNN model can be interpreted through the use of gradient assisted class activation mapping (GradCAM). While CNNs are considered black-box models that are prone to making choices based on spurious correlations between inputs, GradCAM demonstrates that the CNN model uses pyocyanin spectral information in its inferences.

With Chapters 6 and 7 both successfully utilizing the CNN model for SERS spectra, Chapter 8, Quantification of Analyte Concentrations in the Single Molecule Regime Using Convolutional Neural Networks, pushes the model further, using it for single molecule concentration quantification. I recast the problem as an image classification problem, (with SERS spectra as pixels) for which the CNN model is particularly well suited. I show that even NMF is poorly suited for this task, due to the large variation in NMF loading observed at trace concentrations. Yet, the CNN model is shown to produce excellent results, with a limit of quantification down to 10 fM. This tool is shown to be particularly powerful due to its compatibility with transfer learning, a method of training neural networks with small amounts of data. With transfer learning, I show that CNNs can be used for single molecule concentration quantification of a new analyte with only 50 examples per concentration.

Finally, I conclude Part 3 with Chapter 9, Semi-Supervised and Unsupervised Rapid Antimicrobial Susceptibility Testing (AST) with SERS. This chapter demonstrates AST in unprecedentedly short times. Again, a CNN model is used to interpret SERS spectra, but here in the form of a variational autoencoder (VAE). The unsupervised VAE builds a latent space that the spectra occupy which can be easily interpreted and used for semi-supervised and unsupervised classification. PA is either treated with antibiotics that it is resistant to, susceptible to, or not treated and SERS spectra are acquired from the PA lysate. These spectra are used to build the VAE latent space, and a clear trend emerges that is associated with cell health, and AST classification using machine learning tools is greatly improved. Finally, I show that this VAE visualization creates a virtuous cycle, where the computer informs the user of important information, enabling the user to acquire more useful spectra for the VAE to use. This results in ultra-accurate VAE models beating 99% AST classification accuracy.

Chapter 2

Background

2.1 Localized Surface Plasmon Resonance (LSPR)

From a classical standpoint, LSPR is simply exploiting a pole in the polarizability of a metal/dielectric interface to excite a guided wave in the interface. To make sense of this, let us first consider the polarizability of a metal/dielectric interface. For conceptual and mathematical simplicity we may consider the polarizability of a dipole embedded in a bulk material depicted in Figure 2.1, which yields the well-known Clausius-Mossotti relation.[189]

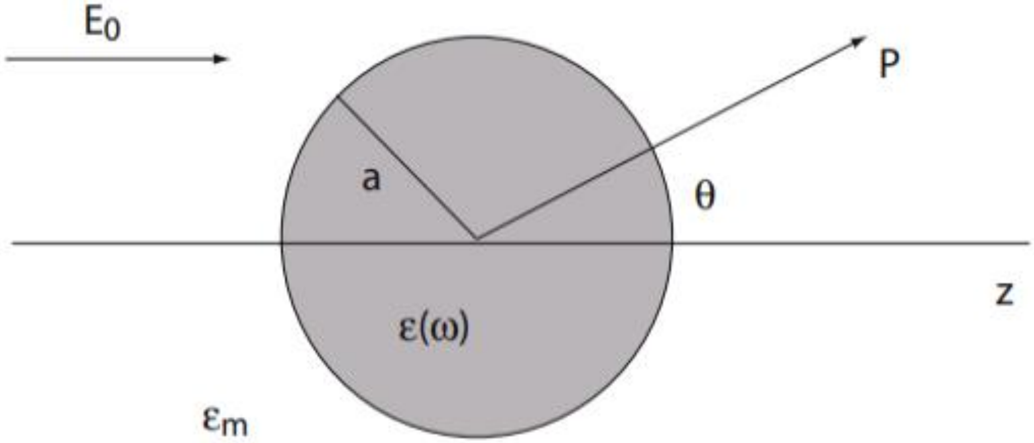

Figure 2.1: Sketch of the physical system considered in the derivation of LSPR. Adapted from Maeir with permission from Springer.[38] Copyright 2007.

$$\alpha = 4\pi a^3 \frac{\varepsilon(\lambda)-\varepsilon_m(\lambda)}{\varepsilon(\lambda)+2\varepsilon_m(\lambda)} \tag{2.1}$$

Where α is the polarizability of the nanosphere, a is the radius of the nanosphere, ε is the permittivity of the dielectric environment, ε_m is the permittivity of the metal, and λ is the wavelength. This system is isomorphic to the case of a spherical nanoparticle in a dielectric environment after applying the quasi-static approximation. From equation 1 we immediately observe that a pole exists where $\varepsilon(\lambda) = -2\varepsilon_m(\lambda)$, a condition (the Fröhlich condition) easily satisfied for a metal nanoparticle embedded in a dielectric medium. This condition yields an infinite polarizability, in realistic situations the polarizability is bounded by the loss/storage term that comes from the imaginary component of a material's permittivity. It is worth mention that the quasi-static approximation which requires that the radius of the particle is much smaller than the wavelength of light places the upper bound on how large a particle may be to exhibit this unbounded polarization. At some point LSPR cannot be excited by a plane wave and surface plasmon polariton behavior is observed, which is significantly more challenging to excite due to momentum matching conditions.

This simple dipole model can be extended further to reveal the usefulness of metal nanoparticles in surface enhanced spectroscopy techniques. Under plane-wave illumination, the electric near field of an oscillating electric dipole (our simplified model for a metal nanoparticle) is described by the following formula:[189]

$$E = \frac{3n(n \cdot p) - p}{4\pi\varepsilon_0\varepsilon_m}\frac{1}{r^3} \tag{2.2}$$

Where n is the unit normal vector, p is the dipole moment, ε_0 is vacuum permittivity, and r is the distance from the dipole. Given the dipole moment is a function of the polarizability times the incident electric field, we immediately see that, having met the Fröhlich condition, the electric field near the nanoparticle is unbounded. We shall see below that the intensity of light (the electric field squared) observed in many spectroscopy techniques is proportional to the incident light intensity and is thus modulated by the particle's polarizability.

2.2 Surface Enhanced Raman Scattering (SERS) Spectroscopy

Raman scattering is a second order nonlinear optical phenomenon wherein light inelastically scatters off of a molecule and interacts with molecular vibrations, resulting in the light energy being reduced by the vibrational energy. While a full quantum mechanical description of Raman scattering is out of the scope of this work, the classical description of this phenomenon describes the salient features of Raman scattering. First, consider a molecule that undergoes a vibration. For vibrations that are Raman active (that is they have large Raman scattering cross sections) the vibration will result in a shift of its electron cloud relative to the nuclei in the system, resulting in a dipole moment. Coupling of this dipole moment to the dipole moment caused by an incident (significantly higher frequency) plane wave is the cause of Raman scattering. Explicitly, consider the dipole moment caused by incident light.[189]

$$p_0 = \alpha_0 E_0(\omega_0 t) \tag{2.3}$$

Where p_0 is the dipole moment of the molecule caused by the light, α_0 is its polarizability, E_0 is the electric field of the light, ω_0 is the frequency of the light, and t is the time coordinate. If a molecular vibration induces a dipole moment that is sufficiently small compared to p_0 we can determine the total polarizability via Taylor expansion[189]

$$\alpha = \alpha_0 + \frac{\partial \alpha_v}{\partial Q} Q_0(\omega_0 t) \tag{2.4}$$

Where α_v is the polarizability from the vibration, Q is the total electron cloud displacement relative to nuclei, and Q_0 is the displacement from just incident light. Using this polarizability to determine the total dipole moment and assuming harmonic oscillation of the electric field and displacement we find (after a simple trigonometry identity)[189] $P = \alpha_0 E_0 \; cos \; cos \; (\omega_0 t) \; + (\frac{\partial \alpha_v}{\partial Q} \frac{Q_0 E_0}{2}) \{cos[(\omega_0 - \omega_v)t] + cos[(\omega_0 + \omega_v)t]\}$ (2.5)

This is the most important result for Raman scattering. Inspection of the above formula shows that induced dipole moments are created at three frequencies: ω_0, ω_0 - ω_v, and ω_0 + ω_v. Light scattered from a molecule will be observed at these energies, which correspond to Rayleigh, Stokes, and anti-Stokes scattering, respectively, as depicted in Figure 2.2. Most electrons are in the ground state so Raman scattering (inelastic scattering) is measured at ω_0 - ω_v. We see now that Raman scattering may be used to measure the vibrational energies of molecules.

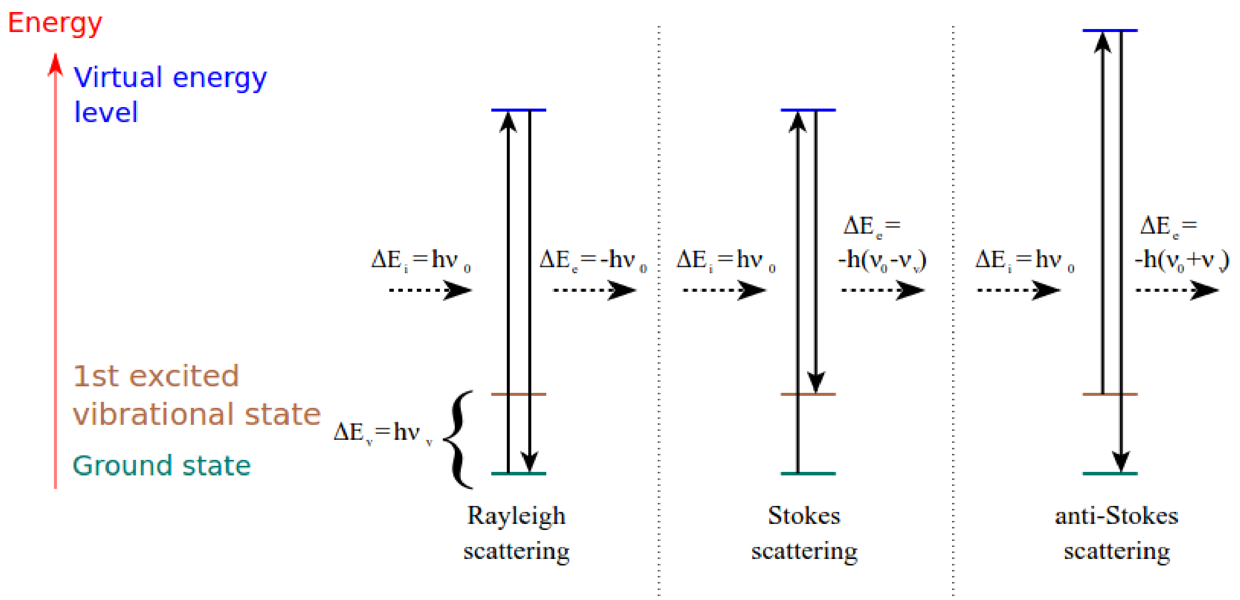

Figure 2.2: Diagram of various light scattering phenomenon. ν_0 and ν_v correspond to the normal frequency of the incident light and vibration, respectively. Adapted from Wikipedia.[190]

From here the benefit of using LSPR is simple; we replace the E_0 in 5 with the electric field produced in metal nanoparticles near field from 2, which is bounded only by causality and losses. Further, note that we measure Raman intensity, not electric field, so the measured field will be E^2. Further still, the intensity is increased both by the electric field brought to the molecule *and* the electric field brought from the scattered light emitted from the molecule back to the far field. SERS enhances the electric field in both cases, giving a total enhancement of E^4, relative to the incident electric field. This E^4 dependence makes SERS very powerful, doubling the electric field increases Raman scattering by a factor of 16. Plasmonic nanogaps – which use mode hybridization to improve field confinement (analogously to how dihydrogen increases electron density between hydrogen atoms relative to a single hydrogen) – give field enhancements over 100 fold, resulting in greater than 10^8 Raman scattering intensity.

2.3 Electrohydrodynamic (EHD) Flow

EHD flow arises in colloidal, electrolytic, solutions at an electrode surface. An analytical description of EHD flow is sufficiently complex to make it outside the scope of this work so a qualitative description will suffice. In this work, only DC potentials will be considered, but the effect also exists with some differences, for AC potential. First, consider a colloidal particle in an electrolytic solution between an electrode and counter electrode and a DC potential driven between them. This particle will have an electric double layer of electrolyte surrounding it to neutralize its intrinsic zeta potential. The applied field moves electrons in the particle to one of its poles, inducing a dipole moment. Ions will rush to the particle's poles to resist this dipole moment and eventually the particle will become fully charged by electrolyte. This in turn creates a nonuniform distribution of (induced) double-layer charge along the equator, generating a tangential electric field that drives a tangential ionic (and thus osmotic) flow. This results in a quadrupolar fluid flow profile with flow towards the particle's poles and out from its equator. The streamlines associated with this flow (u_s) are depicted in Figure 2.3a in the case of an uncharged cylinder.

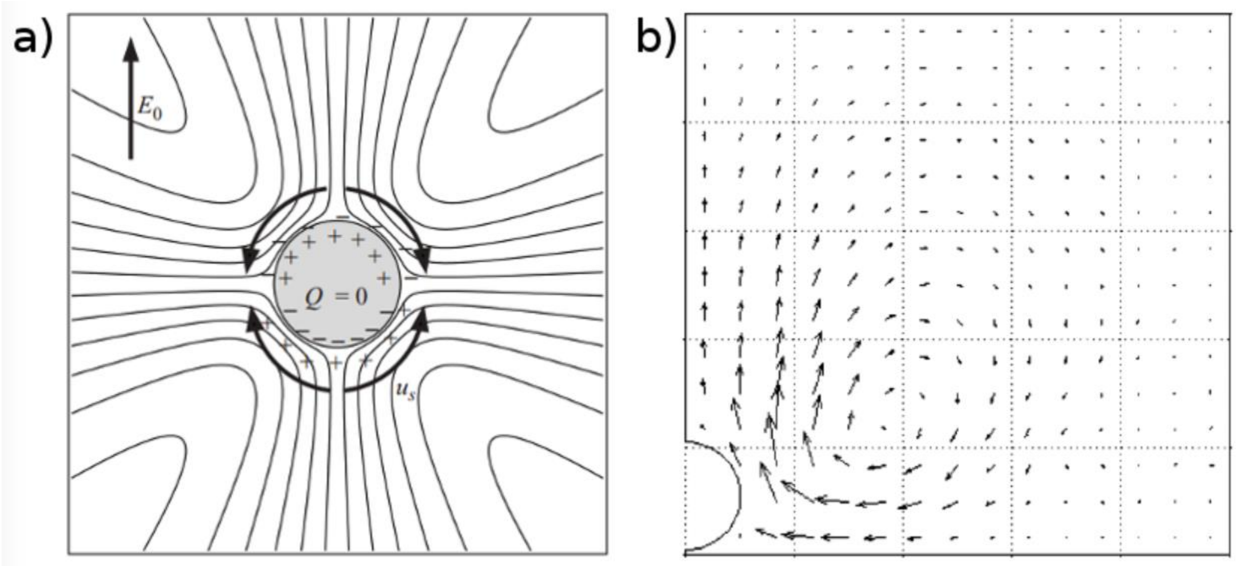

Figure 2.3: a) depicts the EHD flow streamlines around an uncharged cylinder in a homogeneous electrolytic solution adapted from Bazant and Squires with permission from Cambridge [1912] b) depicts EHD flow streamlines around an uncharged cylinder above an electrode, where the double layer thickness equals the particle radius adapted from Sides et al with permission from the American Chemical Society.[192] Copyright 2003

The induced charge electroosmosis described above becomes much more interesting when symmetry is broken. This occurs when a particle is at the electrode surface. Here, fluid cannot flow from the particle through the surface, trapping the particle at the surface. Now the tangential ion buildup will interact with the double layer on the electrode, causing tangential ionic current. The resulting (osmotic) flows are now called EHD flow, and the flow now occurs inward towards the particle's equator and out through the top of the particle. Streamlines associated with this flow are depicted in Figure 2.3b This flow entrains nearby particles and two dimensional assembly of particles proceeds.

2.4 Machine learning

Machine learning, the use of powerful computers to perform statistical inference with extremely efficient algorithms, has long been used to aid in the interpretation of SERS spectra.[166] Deep learning – a modern incarnation of machine learning which has produced disruptively high performance in various artificial intelligence applications – is rooted in the development of artificial neural networks (ANNs). The lens of ANN can also aid in understanding simpler machine learning algorithms that are used in this work. To understand ANNs, consider its fundamental building blocks: The artificial neuron (Figure 2.4).

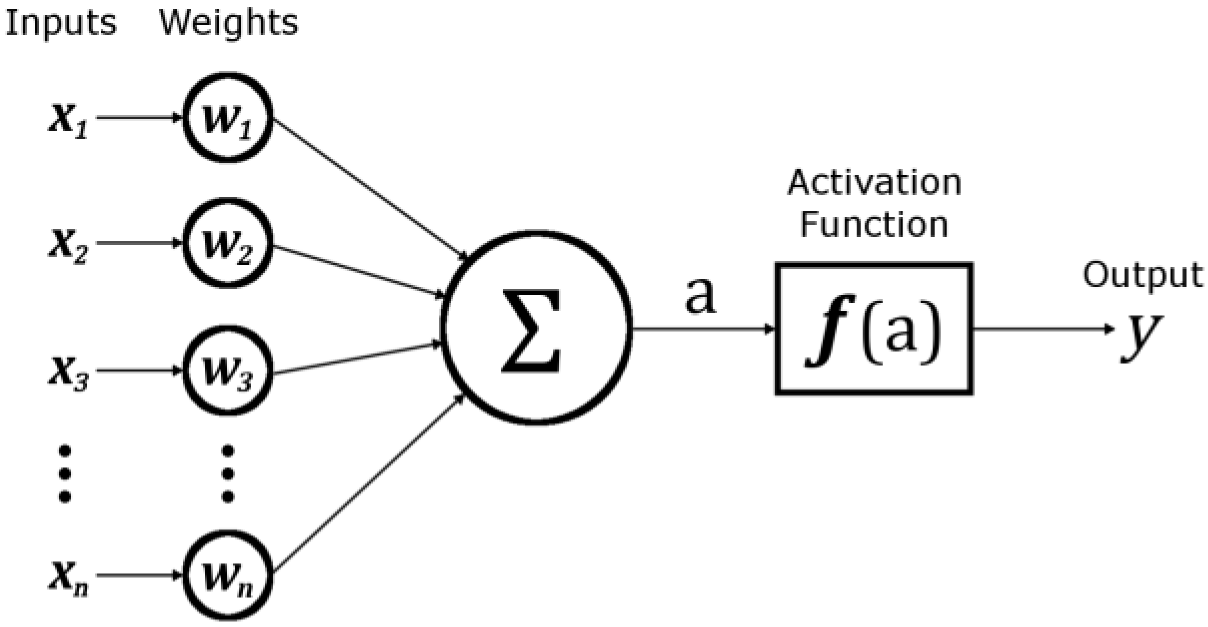

Figure 2.4: Components of an artificial neuron. Adapted from Wikipedia[193]

From Figure 2.4 one can see that a single neuron takes in multiple inputs and outputs to a single value. The output is computed in two steps: (1) the neuron’s activation is calculated as a weighted sum of all inputs and (2) an activation function is then applied on the activation to produce the output, summarized in Equations 2.9 and 2.10, respectively.

$$a = \sum x_j w_j \tag{2.6}$$

$$y = f(a) = f\left(\sum x_j w_j\right) \tag{2.7}$$

Here a is the activation, x_j is the jth input, w_j is the corresponding jth weight, f(x) is the activation function, and y is the output. Activation functions can be chosen from a range of options, such as linear or rectified linear units (ReLU), among others. This output is then compared to the true value with the loss function, which is typically mean squared error (the average deviation of the predicted value from the true value, squared) for regression problems, and binary crossentropy (the amount of entropy (a measure of information carried between the distributions) between the predicted and true values) for categorization problems.[194]

ANNs are built by connecting artificial neurons into networks (directed acyclic graphs), as visually depicted in Figure 2.5. Here each line represents a weight, and each circle represents a node, or a single neuron. Often there are many intermediate layers between the input and output that are termed hidden layers.

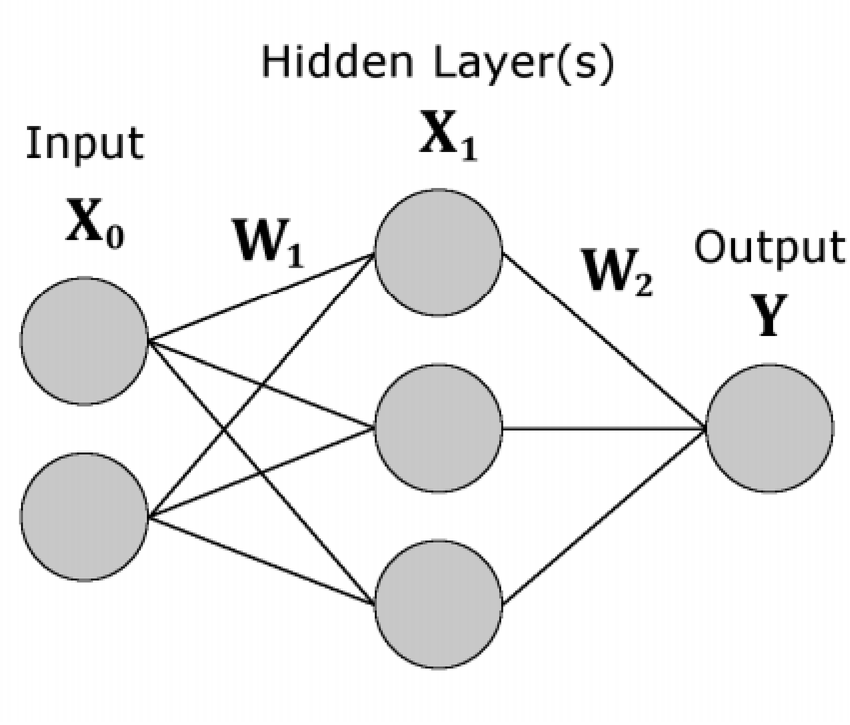

Figure 2.5: Schematic representation of a fully connected neural network with 1 hidden layer.

These considerations yield the following equations:

$$\mathbf{A_1} = \mathbf{X_0 W_1} + \mathbf{B_1} \tag{2.8}$$

$$\mathbf{X_1} = f(\mathbf{A_1}) \tag{2.9}$$

$$\mathbf{A_2} = \mathbf{X_1 W_2} + \mathbf{B_2} \tag{2.10}$$

$$\mathbf{X_2} = \mathbf{Y} = f(\mathbf{A_2}) \tag{2.11}$$

where $\mathbf{A_i}$, $\mathbf{X_i}$, $\mathbf{W_i}$, and $\mathbf{B_i}$ depict the activation vector, output vector, weight matrix, and bias vector of the i^{th} layer, respectively. From this, one can see that ANNs are a series of matrix transformation separated by nonlinear activation functions. During training, the weight matrices $\mathbf{W_i}$ of each layer are iteratively updated—or learned—using backpropagation.[195]

With this framework we can understand several of the machine learning algorithms used in this work. First consider the multilinear regression used in chapter 6. We can reinterpret this algorithm as an ANN with all input wavenumbers fully connected to just one neuron with a linear activation function and mean squared error loss function. Non-negative matrix factorization (NMF), which is used extensively in this work to reduce the dimension SERS spectra into scores that represent isolated analyte signals, can be reinterpreted similarly. For NMF, we would have all input wavenumbers fully connected to nodes (whose number equals the number of NMF components used) with linear activation functions, Mahalanobis distance loss function, and the requirement that all weights and biases be positive. Convolutional neural networks, used in the later chapters of this work, are similar to ANN, but not all nodes are fully connected, and the connection rules are determined by the user for the particular application.

Part Two: Nanomanufacturing Large-Area, High Performance Optical Devices

"The astonishment which I had at first experienced on this discovery soon gave place to delight and rapture. After so much time spent in painful labour, to arrive at once at the summit of my desires was the most gratifying consummation of my toils."

-Mary Shelley, Frankenstein, or the Modern Prometheus

Chapter 3

Driving Chemical Reactions in Plasmonic Nanogaps with Electrohydrodynamic Flow

3.1 Introduction

Since seminal reports demonstrating assembly of superlattices from colloidal dispersions,[196] fundamental understanding of material, geometric tolerances, and process parameters in assembly of nanospheres from colloid has led to long range order in superlattice thin films on the wafer scale,[197,198] marking significant progress in the ability to structure matter from molecular to mesoscopic length scales. In the case of 2D plasmonics and metasurfaces, it is often desirable to fabricate discrete assemblies rather than closed packed films that requires further understanding of hierarchical driving forces (long-range and short-range) for control of resultant architectures.[199] These discrete assemblies are often referred to as oligomers because plasmon hybridization theory intuitively describes the electromagnetic response using an analogy with molecular orbital theory.[130] Nanoparticle colloids thus provide metamolecule building blocks where not only composition, size, and shape of the nanoparticle but the geometry of resultant oligomers, gap spacings and dielectric environment provide additional degrees of freedom for tuning electromagnetic response.[200–203] Our group's early work has demonstrated that metal nanoparticles from colloids can be assembled in oligomers over large areas using self-organized chemical patterns[131] that produce enhanced electromagnetic fields for SERS sensors.[137] (Sub)nanometer gaps between plasmonic nanoparticles has led to unprecedented light-matter interactions including: single molecule surface enhanced Raman scattering (SERS) spectroscopy,[121] room temperature single molecule strong coupling,[204] and second harmonic generation.[205] Yet fabrication of discrete assemblies at nanoscale (sub-100 nm) dimensions with controlled (sub)nanometer gap spacing and at high densities over large areas remains challenging.[206,207]

Long-range driving forces (~µm scale) have been used extensively in colloidal assembly, including capillary forces,[208–210] convection,[211,212] optical tweezing,[213] electrophoresis,[214] and electrokinetic phenomena (electrically driven fluid flow and particle motion).[215–218] Short-range driving forces, such as electrostatic interactions, Van der Waals forces, and chemical crosslinking, are also essential for control gap spacing which has a profound impact on the optical response of plasmonic assemblies. Of these, chemical crosslinks – including DNA origami,[219–221] Cucurbit[*n*]uril,[222,223] small molecules,[224–226] protein linkers,[227,228] and polymer encapsulation[229–231] – provide the most flexibility and control in architectures achieved from colloidal dispersions. Yet assemblies fabricated using chemical crosslinking methods typically do not employ long-range driving forces and rely on diffusion often leading to low oligomer density and/or incomplete assemblies.

In this thesis, we use equilibrium charge electroosmosis, also referred to as electrohydrodynamic (EHD) flow, to drive the chemical crosslinking of nanospheres on a working electrode surface under a DC potential with a counter electrode. EHD flow[191,232,233] is an electrokinetic phenomena that arises from lateral potential gradients within an electrode polarization layer[218] and has been studied for both AC and DC applied potentials.[234] The resultant flow can be radially attractive toward the source of the gradient; in this chapter a nanosphere in the lateral plane of an electrode is used to generate the gradient. EHD flow has been primarily studied in the context of micron scale particles where particles become entrained in flow fields resulting in two-dimensional close-packed assemblies. Resultant structures are transient and are typically imaged *in situ* using confocal microscopy making it difficult to understand assembly behavior at sub-100 nm dimensions.

Interestingly, EHD flow is found here to drive chemical crosslinking of 40 nm Au nanospheres forming anhydride bonds between them. Anhydride bonds are observed spectroscopically with SERS and are consistent with gap spacings observed via transmission electron microscopy. By freezing transient EHD assemblies with chemical crosslinking, oligomer distributions on the surface are analyzed to provide insight into how long range forces (EHD, Brownian motion) and short range forces (chemical crosslinking, electrostatic interactions) affect assembly of nanospheres. Statistical analysis of scanning electron microscopy images shows permanent oligomers composed of ten nanospheres or less result when deposition favors EHD flow over random Brownian motion. Extinction response of assemblies over large area is dominated by oligomers with gap spacings of 0.9 nm. Further probing of large area optical response with a standard SERS analyte demonstrates uniform enhancement factors of 10^9 with a relative standard deviation is 10% over a 1 mm^2 area. Thus these surfaces provide the benefit of uniformity of contemporary lithographic[235] and roll to roll[236] fabrication of SERS surfaces while providing high signal enhancements due to 0.9 nm gap spacings in oligomers. Overall assembly is found to depend on the probability of: 1) nanosphere-nanosphere collisions (influenced by EHD flow), 2) overcoming activation energy barriers (enthalpic effects) and 3) steric interactions (entropic effects) leading to appropriate molecular configurations for reactions to occur. These results demonstrate the ability to enable chemical reactions in plasmonic nanogaps with long range colloidal driving forces. This work represents a step forward for directed assembly of nanoarchitectures to create metasurfaces by demonstrating the influence of both the template and physical driving forces in assembly to control oligomer morphology and uniformity of gap spacing.

3.2 Investigation of Driving Forces in Assembly

Two-dimensional growth of Au nanosphere oligomers over areas as large as 1 cm^2 occurs via a seeded growth method where oligomer growth from Au nanosphere seeds is driven by EHD flow; oligomers are then stabilized by chemical crosslinking that also provides uniform gap spacings. The process of oligomer formation is depicted in Figure 3.1a and the mechanism is as follows: 1) Electrophoretic sedimentation guides Au nanospheres towards a doped Si working electrode coated with a PS-b-PMMA diblock copolymer template. Carbodiimide crosslinking chemistry[237] covalently binds these Au nanosphere seeds on amine functionalized PMMA regions. The PS-b-PMMA diblock copolymer template is composed of PMMA lamella domains with widths of 40 nm and fractional surface coverage of 28%. The template is chosen to distribute Au monomer seeds evenly over the entire surface. 2) The bound Au nanosphere monomers serve as the seeds that nucleate further oligomer growth by generating EHD flow that entrains nearby nanospheres, extending approximately 4 particle diameters,[238] dragging them towards the source of the perturbation in the plane of the electrode. The inset of Figure 3.1a illustrates the tangential, attractive EHD flow on an electrode surface near a covalently bound Au seed on PMMA. 3) As oligomers formed by EHD are transient,[234] e.g., when the voltage is removed the assembled particles will dissolve back into the bulk solution without a strong attractive force with the surface. Here, carbodiimide crosslinking chemistry stabilizes oligomers after formation.

Chemical pathways of carbodiimide crosslinking chemistry are illustrated Figure 3.1b. Au nanospheres are assembled on the surface via two paths. First carbodiimide activates the carboxylic acid on the particle forming the isourea species which then forms an amide bond with

PMMA, labeled Path 1. Amide bond formation with PMMA can also be accelerated with S-NHS via intermediate succinide formation, labeled Path 3. However, the activated isourea may also react with carboxylic acids yielding anhydride bonds as labeled in Path 2.[237] Anhydride bond formation is not frequently observed, and when observed it has been attributed to the close proximity between carboxylic acid groups during polymer crosslinking,[239] and related to the Thorpe–Ingold effect in cyclization reactions. Atomistic simulations show that the equilibrium distance between two gold surfaces crosslinked with anhydride to be 0.85 nm. This prediction is consistent with the gap spacing between nanospheres observed in transmission electron microscopy (TEM) images (Figure 3.2a) which is approximately 0.9 nm. It is also important to note that the gap spacing observed is different when oligomers are assembled in the absence of applied bias; where previous reports show a range from 2-7 nm.[137] Since all substrates are assembled in 23 mM electrolyte solution resulting in an electrostatic potential barrier between nanospheres.

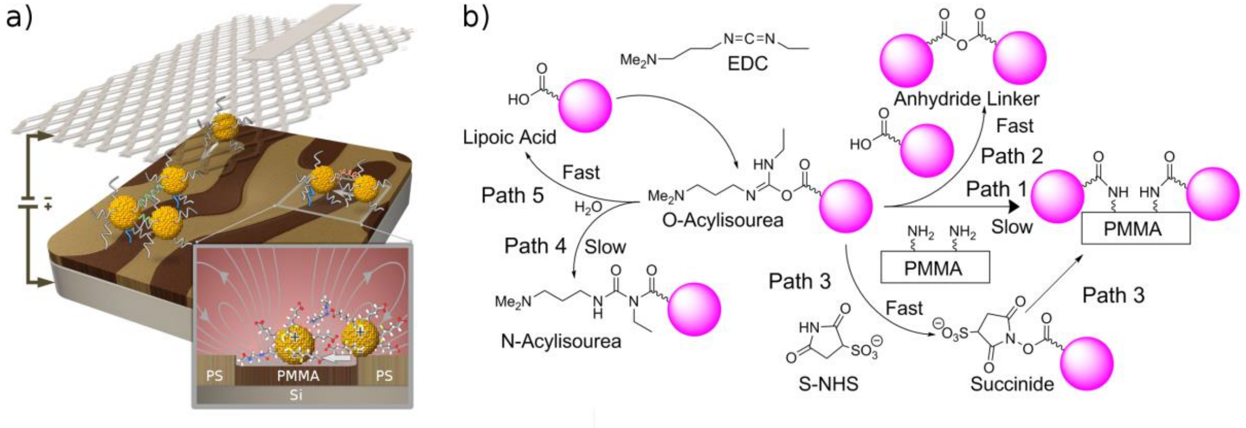

Figure 3.1: (a) Schematic of EHD attractive forces in colloidal solution of carboxylic acid functionalized Au nanospheres near the working electrode composed of a self-organized PS-b-PMMA diblock copolymer template on Si. Inset depicts the EHD flow that entrains nearby particles around a Au seed, increasing the probability of nanosphere-nanosphere collisions (b) Carbodiimide chemistry pathways yield covalent linkages to the amine functionalized surface via paths 1 and 3, while covalent linkages between nanospheres are achieved through path 2.

While the TEM results support anhydride bond formation, it is also possible to spectroscopically probe molecules between nanospheres. Surface enhanced Raman scattering (SERS) spectra provides a means to interrogate molecules near the center of the gap where the electric field intensity is maximum. Figure 3.2b (lower blue curve) depicts a SERS spectra of a sample with 40 nm Au oligomers on the surface prepared as described in the methods section. Vibrational bands associated with anhydride groups are observed at 1011 cm^{-1}, 1230 cm^{-1}, 1380 cm^{-1}, similar to those reported for acetic anhydride. We further probe the ability to monitor local chemistry in the gap by cleaving anhydride linkages via nucleophilic substitution. The orange curve in Figure 3.2b depicts a SERS spectra observed after overnight treatment of a sample with 0.5 mM methylamine, a molecule with weak Raman bands.[66] Notably the vibrations associated with anhydride groups no longer stand out above the background indicating they have been cleaved by methylamine. In order to further probe chemistry in the gaps, nucleophilic substitution of the anhydride linker with aniline, a molecule with a higher Raman cross section than methylamine, is examined. The green curve in Figure 3.2b depicts SERS spectrum observed after overnight treatment of a sample with 0.5 mM aniline. Characteristic peaks are observed at 1000 cm^{-1} and 1173 cm^{-1}. For comparison, subsequent treatment of first methylamine (to cleave anhydride bonds) and then aniline (where selective attachment in the hotspot will now not occur since anhydride bonds have been cleaved) indeed shows a much weaker SERS signal for aniline. This spectrum is shown in the top green curve of Figure 3.2b. Here any indication of aniline in the spectrum would be due to binding to the gold surface via amine-Au interactions. This does not appear to be occurring at a significant rate and this is reflected in the SERS spectrum. The SERS spectra showing the presence of vibrations associated with anhydride groups and then aniline groups after nucleophilic substitution with methylamine indicates molecules can be placed in the gaps via a reaction with the crosslinking molecular group. Thus the crosslinking chemistry used to assemble oligomers also enables sensing of nucleophilic compounds.

The SERS data of Figure 3.2b demonstrates that functional groups used for chemical assembly and other molecules in the gaps are observable in spectra due to high electric field enhancement in the nanometer gaps. Full-wave finite elements method simulations show an electric field enhancement on the order of 600 in the hotspot region of a dimer due to the narrow gap spacing; the electric field profile on a cross section passing through the middle of dimer is shown on a logarithmic scale in the inset of Figure 3.2c. We can also measure field enhancements using a standard Raman reporter molecule, benzenethiol (BZT). Figure 3.2c (lower curve) shows SERS spectra of an oligomer sample after exposure to BZT. Yet if anhydride bonds reside in the 'hottest' region, BZT cannot access this region. An oxygen plasma treatment was then performed to remove ligands from nanosphere surfaces. SERS spectra of an oligomer sample with BZT self-assembled onto the nanospheres' surface after plasma treatment is shown in the upper curve of Figure 3.2c. Here, SERS signal from BZT increase up to 258%. The increased SERS signal is concomitant with increased hotspot occupation volume by BZT and can be attributed to diffusion of BZT into the hotspot that is possible after removal of the anhydride from the gaps between nanospheres. Interestingly, comparison of the 1076 cm^{-1} in plane ring deformation superimposed with C-S stretch and the 998cm^{-1} in plane ring deformation shows greater enhancement of the 1076 cm^{-1} band after removing the anhydride linker; the two peaks show approximately equal intensity in SERS data before plasma etching. This further indicates that the C-S bond has displaced the anhydride previously located in the hotspot as the relative intensity of vibrational modes is related to the molecular orientation in the gap. The ability to remove anhydride groups in the gaps is important for label-free SERS sensing applications where analyte molecules in hotspots will yield lower detection limits.

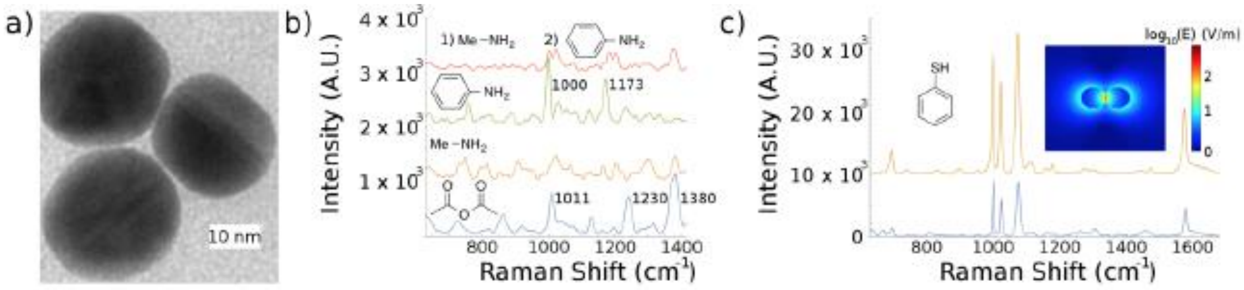

Figure 3.2: (a) A representative transmission electron microscopy micrograph of EHD assembled Au nanospheres deposited on PS-b-PMMA coated graphene membrane grid. Gap spacings observed are consistent with anhydride linkers forming from LA ligands on nanospheres. (b) SERS spectra (each offset from one another by 1x10^3 for visual clarity) of EHD assembled oligomer sample of an untreated sample with anhydride group intact (blue curve), after overnight treatment of a sample with 0.5 mM (orange curve) methylamine - which cleaves anhydride groups - and (green curve) aniline - which cleaves the anhydride and leaves a phenol group, (red curve) after subsequent treatment with methylamine and then aniline and (c) when a BZT self-assembled monolayer is formed on a sample (blue curve) without and (orange curve) with oxygen plasma treatment. The inset depicts the magnitude of electric field in logarithmic scale

Reactive molecular dynamics (MD) simulations illuminate the origin of the EHD driven anhydride crosslinking. The stability of Au nanosphere's in solution with carbodiimide crosslinker indicates that the anhydride pathway is a rare event.[240] We examine this hypothesis by analyzing the probability distribution function of the distance between the reactive carbon atom in acylisourea (OA) and the deprotonated oxygen atom in lipoic acid (LA) – $d_{\text{C-O}}$ – using the transferable ReaxFF potential.[241,242] The exact concentration of LA and OA ligands on the surface may vary from nanosphere to nanosphere so three concentrations are analyzed: one OA to one LA (Low), one OA to four LA (medium) and two OA to four LA (High) concentrations. The concentrations investigated are in the range of LA ligands observed on the surface of gold nanospheres.[243] In simulations, ligands are connected to opposing faces of an Au (111) slab at 11, 15, and 20 Å nanogap separations, d_{Au}. Simulation parameters are further detailed in the methods section. Figure 3.3a-c (dotted curves) depicts the distributions of $d_{\text{C-O}}$ in the various cases, with Figure 3.3d providing an interpolated contour map of the relationship between d_{Au}, $d_{\text{C-O}}$, and ligand concentration. Two features of the nanogap chemistry are salient: 1) 17 Å corresponds to fully extended LA and OA ligands and is thus the minimum d_{Au} for the reaction to occur. This minimum distance for reaction is not observed in TEM images obtained from control samples assembled without EHD flow. This minimum distance for reaction is not observed in TEM images obtained from control samples assembled without EHD flow. 2) Regardless of ligand concentration and d_{Au} the proximity required for reaction – $d_{\mathrm{C\text{-}O}} = 2.9$ Å – is at least 4 standard deviations from the mean $d_{\mathrm{C\text{-}O}}$. These results indicate that EHD flow is necessary for maintaining close proximity between nanoparticles in timeframes reasonable for reaction.

Interesting ligand dynamics also emerge from the MD study, indicating entropy's role in this reaction. Gaussian mixture analysis of $d_{\mathrm{C\text{-}O}}$ trajectory, depicted in Figure 3.3a-c as solid curves, identifies three normal distributions regardless of concentration and nanogap distance. The inset of figure 3.3b visualizes these states in red, violet and cyan in the configuration space. These three states are separated by energy barriers comparable to thermal fluctuations, k_BT, schematically illustrated in Figure 3.3c's inset. MD simulations probe these three degenerate states via a diffusive mechanism, which is entropic in nature. We find that the average $d_{\mathrm{C\text{-}O}}$ decreases with increasing concentration except for the case of high concentration at small nanogap distances, d_{Au}=11 Å. The latter is strongly affected by the presence of large steric forces between OA and LA groups that ultimately hinder the mobility of reaction sites in the nanogap. The minimum average proximity, $d_{\mathrm{C\text{-}O}}=4.1\pm0.3$ Å, occurs at high concentration, when the nanogap separation is 15Å. Yet, this distance is greater than the proximity required for reaction, 2.9 Å, and only the tail end of the probability distribution function samples this space. Therefore, we can conclude that entropic effects play a critical role in bringing reactive cores in proximity of each other – an event that occurs rarely as a result of competition with steric effects. Furthermore, our results confirm that the probability of bringing reactive groups in the proximity of one another is dependent on both the concentration and nanogap size, $P_{entropy}=P_{entropy}(C, d_{\mathrm{Au}})$ hypothesized in previous work to drive anhydride bond formation via carbodiimide crosslinking.[239]

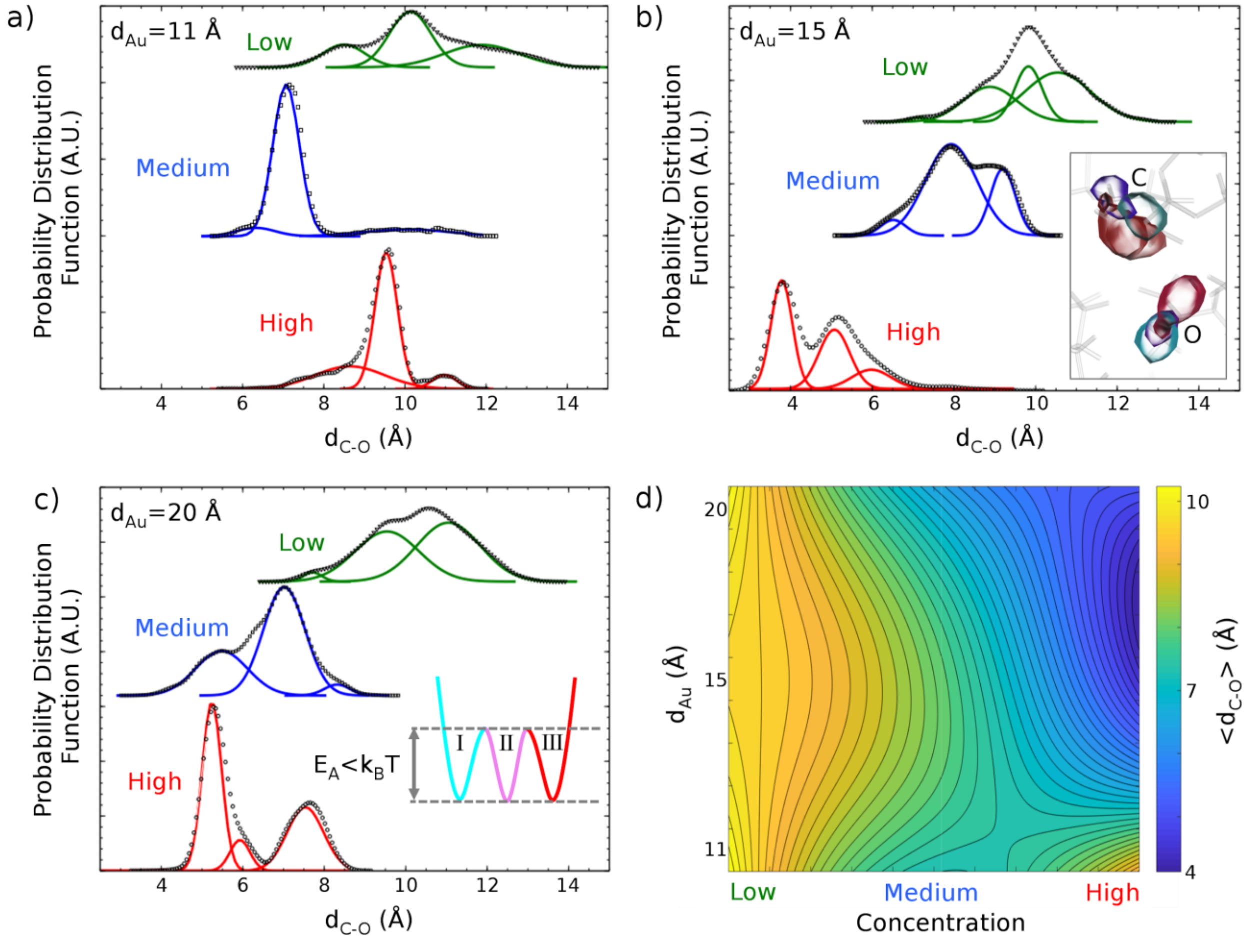

Figure 3.3. Analysis of the distribution of C-O distance, $d_{\mathrm{C\text{-}O}}$, corresponding to the reactive carbon atom in OA and the deprotonated oxygen atom in LA for low (one OA+ one LA), medium (one OA+ four LA) and high (two OA+ four LA) concentrations. OAs and LAs are interacting in the aqueous medium between two Au surfaces that are separated at d_{Au} (a) 11, (b) 15 and (c) 20Å. Gaussian mixture analysis indicates presence of three distinct degenerate $d_{\mathrm{C\text{-}O}}$. The inset in (b) represents superposition of the degenerate states that are depicted by red, violet and cyan regions corresponding to three distributions of $d_{\mathrm{C\text{-}O}}$ averaged at 3.8, 5 and 6Å, respectively. Water molecules are not displayed and OA and LA are shown in gray color for the sake of clarity. The inset in (c) provides a schematic representation of the energy landscape consisting three states separated by energy barriers comparable to thermal fluctuations. (d) The color-map of the average $d_{\mathrm{C\text{-}O}}$ as a function of concentration and d_{Au}. The minimum average distance has an entropic nature as a result of competition between concentration and steric interactions.

With the role of entropy understood, we turn now to consider enthalpic processes. Observation of rare chemical reactions with the MD framework is facilitated via a harmonic bias potential that favors small $d_{\mathrm{C-O}}$. At high temperatures anhydride formation is indeed observed, consistent with the premise that EHD flow is necessary to facilitate this reaction. A nudged elastic band (NEB) technique is employed[244,245] to probe the minimum energy path (MEP), given low $d_{\mathrm{C-O}}$. Figure 4 depicts the MEP of Low concentration LA-OA (I) reaction to form an anhydride (III), exhibiting two enthalpic reaction barriers (TS1 and TS2) and a metastable intermediate complex (II) at different d_{Au}. The configurations on top of Figure 3.4 provide a schematic representation of successive reaction stages extracted from NEB simulations. The enthalpy of reaction is found to vary between -2 and -6 kcal/mol, indicating an exothermic process. The TS1 activation energy barrier is roughly 27 kcal/mol (~ 45 k_BT at ambient conditions) irrespective of d_{Au}. This activation energy is comparable to that of imine carbons on polycarbodiimides measured in ^{13}C CP/MAS NMR spectroscopy.[246] The intermediate complex further dissociates to yield anhydride with TS2 activation energy barrier estimated at 5-8 kcal/mol. NEB calculations indicate that transition through TS1 is the rate-controlling step in the anhydride formation. This enthalpic rate is controlled neither by the concentration, nor by the nanogap size. Therefore, the probability of such enthalpic reaction is just a function of temperature, $P_{enthalpy} = P_{enthalpy}$ (T) x const., in our system. Thus the probability of reaction $P_{Reaction} = P_{entropy}(C, d_{\mathrm{Au}}) \times P_{enthalpy}(\mathrm{T}) \times P_{Collision}$ is tuned on the nanosphere assembly surface by modifying $P_{Collision}$ via EHD flow.

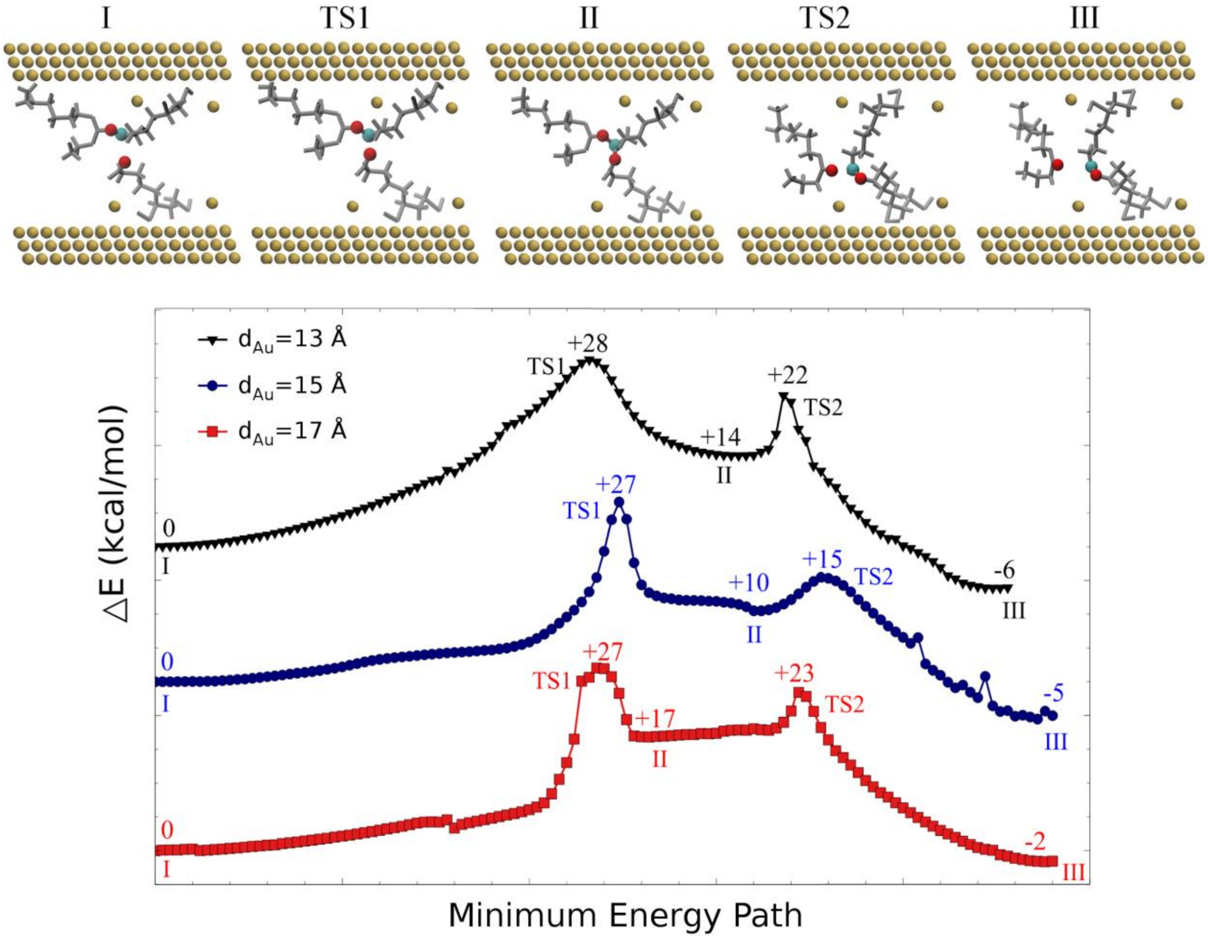

Figure 3.4. Minimum energy path of OA and deprotonated LA reaction obtained from nudged elastic band (NEB) calculations at d_{Au}=13, 15 and 17Å. The initial (I), intermediate complex (II) and final states (III) are separated via two reaction energy barriers, transition states TS1 and TS2. The top schematics provide the snapshot of these states extracted from NEB calculations. The reactive carbon and oxygen atoms are highlighted respectively in blue and red, while the rest of the organic molecules are shown in silver and water molecules are omitted for clarity. The activation energy, E_a, is estimated to be around 27 kcal/mol for TS1 and 5-8 kcal/mol for TS2.

3.3 EHD Driving Forces Influencing Oligomer Assembly

We turn now to investigate how EHD flow induced anhydride crosslinking affects density and geometry of oligomers on substrates; hereafter we refer to samples as EHD-anhydride substrates. This analysis is aided by examining oligomer formation on substrates assembled without applied bias, hereafter referred to as Control substrates. Figure 3.5a,b depicts representative SEM images of substrate surfaces assembled at room temperature without and with an applied bias, respectively. First it is important to consider how oligomers form on Control surfaces. In the case of zero bias-voltage deposition, amine-functionalized copolymer-coated Si substrates are placed in carboxylic acid-functionalized Au nanosphere colloid with carbodiimide crosslinker. Brownian motion drives the nanospheres to randomly collide with the substrate surface. Similar to Au seed assembly, collisions can result in a reaction between an *O*-acylisourea on the nanosphere and an amine on the surface, covalently binding nanospheres to the PMMA domains via an amide bond. An oligomer is formed when another nanosphere reacts with the PMMA surface near a monomer. The tendency to grow beyond a monomer is referred to as oligomerization. Oligomer configuration (measured as number of nearest neighbors), and oligomerization, is examined by measuring percent coverage of various oligomers (monomer, dimer, trimer, etc.) on the substrates' surfaces. Coverage statistic data are shown in Figure 3.5e,f. We find, unsurprisingly, that Control substrates exhibit a steep decrease in percent coverage beyond a monomer, due to electrostatic repulsion between nanospheres, and a monotonic decrease in percent coverage beyond a dimer.

Comparison of statistics in Figure 3.5e and 3.5f show a 336% increase in nanosphere coverage is observed on the EHD-anhydride substrate with respect to the Control. The increase in coverage can partially be attributed to electrophoretic sedimentation due to the applied bias. Yet examination of Figure 3.5e and 3.5f also show dramatically increased oligomerization, few monomers and many close packed oligomers - a signature of EHD flow driven assembly - are observed on the EHD-anhydride substrate. The average number of nearest neighbors per particle increases from 0.66 to 2.2 (where 0 is an isolated particle and 6 is a 2D hexagonally close packed lattice) from the Control to the EHD-anhydride substrate. If electrophoresis was the only driving force, surfaces would exhibit a monotonically decreasing oligomerization distribution similar to the Control substrate. Instead, the distribution is centered around quadrumers in Figure 3.5f and anhydride linking induced by EHD flow leads to over a 2.5-fold reduction in the number of monomers with respect to the Control substrate. This observation is consistent with the seeded growth assembly mechanism where seed monomers are consumed when they oligomerize via anhydride attachment with nanospheres in colloid due to the presence of an applied bias driving EHD flow.

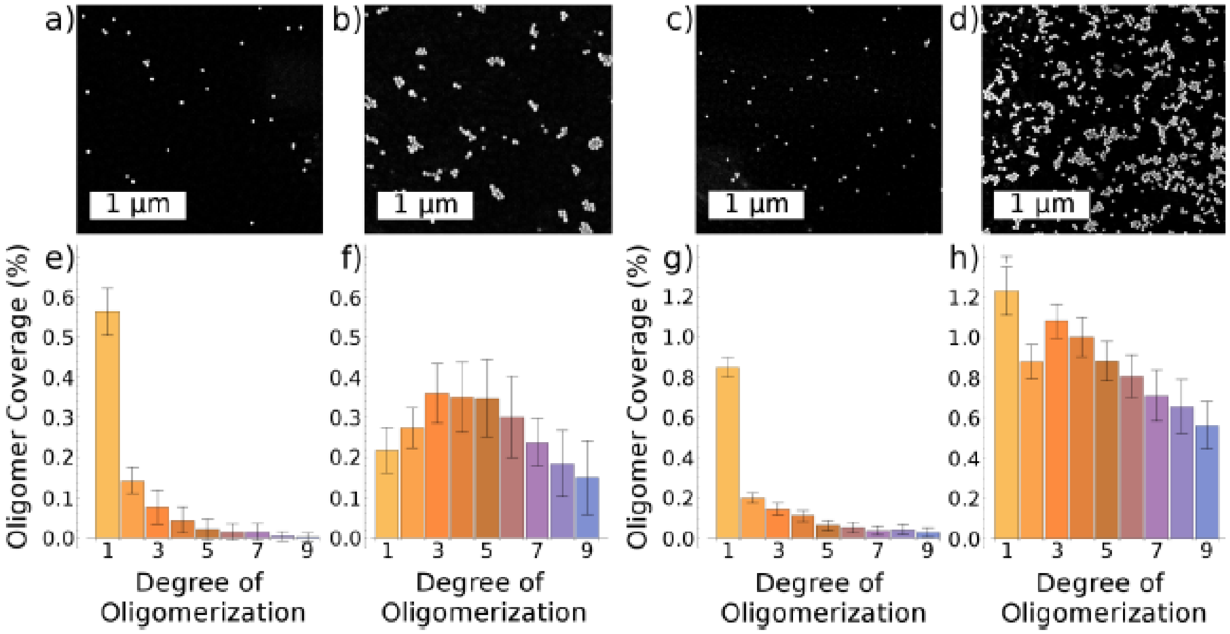

Figure 3.5: Representative SEM images taken from the substrate surfaces with 40 nm Au nanospheres deposited at (a),(b) room temperature and (e),(f) corresponding oligomerization distribution observed on Control substrates and EHD-anhydride substrates, respectively. The close packing of oligomers and reduction of monomer observed on EHD-anhydride substrates are the hallmark of EHD flow driven oligomerization. (c),(d) Representative SEM images and (g),(h) corresponding oligomerization distribution observed on 60 ^{0}C Control substrates and 60 ^{0}C EHD-anhydride substrates, respectively. Nanosphere surface density is seen to increase greatly with increased Brownian motion of the particles. The 60 ^{0}C EHD-anhydride substrate is observed to have a reduced degree of oligomerization compared to the room temperature EHD-anhydride substrate.

It is also important to consider how Brownian motion competes with EHD flow during oligomer formation. The effect of Brownian motion can be evaluated by examining oligomer configuration statistics after assembly when the colloid is heated to 60 ^{0}C during deposition; representative images are shown in Figure 3.5c,d for both Control and EHD-anhydride substrates, respectively. Without voltage, Figure 3.5c, one observes a nearly two-fold increase in the number of nanospheres attached to the surface in comparison to room temperature deposition (compare with Figure 3.5a). With voltage, Figure 3.5d, one observes a dramatically increased density of oligomers, with respect to all other assembly conditions. The total nanosphere surface coverage for the 60 ^{0}C EHD-anhydride substrate is 15%, while the surface coverage the room temperature EHD-anhydride substrate is 3.7% (shown in Figure 3.5b). Figure 3.5g,h depicts the oligomerization statistics obtained from the Control and EHD-anhydride substrates, respectively, to understand the more complex behavior at elevated temperature. First Brownian motion is expected to increase the number of nanosphere collisions with the surface increasing the number of monomer seeds. Monomers indeed dominate in frequency on 60 ^{0}C Control substrate, as observed in the SEM image in Figure 3.5c and statistics in Figure 3.5g. The EHD-anhydride substrate deposited at 60 ^{0}C shows a reduced degree of oligomerization, see Figure 3.5h, when compared to the room temperature EHD deposition (shown in Figure 3.5f). There is a slight shift to smaller oligomers on average; trimers are observed with the higher frequency over quadrumers when deposition occurs at 60 ^{0}C versus room temperature. Oligomers are still observed to be close packed, with an average of 2.06 nearest neighbors per particle, indicating EHD is still a major long range driving force in assembly at 60 ^{0}C. This data unsurprisingly indicates that EHD flow competes with Brownian motion, yet Brownian motion does not completely overwhelm EHD flow in determining oligomer morphology during assembly of nanospheres with 40 nm diameter. Indeed, the increase in temperature is a means for limiting the size of oligomers and increasing their density on the surface by increasing the number of seed monomers on the template surface. For many applications, for example, SERS sensors, the higher surface coverage of the EHD-anhydride substrate prepared at elevated temperature is desirable for high hotspots density. The increased oligomer density and reduced degree of oligomerization from a slight increase in deposition temperature synergistically improve optical response uniformity as we shall see in the following section.

The relative contributions of EHD flow, Brownian motion, and electrophoretic sedimentation can be further elucidated by comparing oligomerization of particles with different dimensions.[247] While long range forces associated with EHD flow (Brownian motion) increases (decreases) with increasing particle diameter, electrophoretic mobility is primarily dependent on zeta potential.[248] Figure 3.6a,b are representative SEM images and oligomerization statistics from room temperature and 60 ^{0}C depositions, respectively, performed with 20 nm diameter Au nanospheres. All other deposition parameters are identical to the 40 nm diameter Au nanosphere EHD-Anhydride depositions. Oligomerization statistics depicted in Figure 3.6c,d are determined from 25 SEM images with dimensions of 4.0 μm × 2.7 μm. The room temperature sample has significantly greater total surface coverage than its 40 nm diameter counterpart, 17% compared to 3.7%, likely due to the reduced Stokes drag on the smaller particles.[248] Figure 3.6d shows that increasing heat to 60 ^{0}C greatly reduces EHD flow induced oligomerization, this is likely due to the significant increase the average velocity of the 20 nm particles at these temperatures. This data shows that increased Brownian motion competes with EHD flow and the faster moving nanospheres have a reduced probability of being entrained by EHD flow. Overall the 20 nm diameter substrates are significantly less oligomerized than the 40 nm substrates, which is consistent with the fact that EHD flow force density scales with particle size; 20 nm diameter appears to be the lower limit in which EHD flows are an important driving forces for colloidal assembly, consistent with observations using optoelectrokinetic flows.[217]

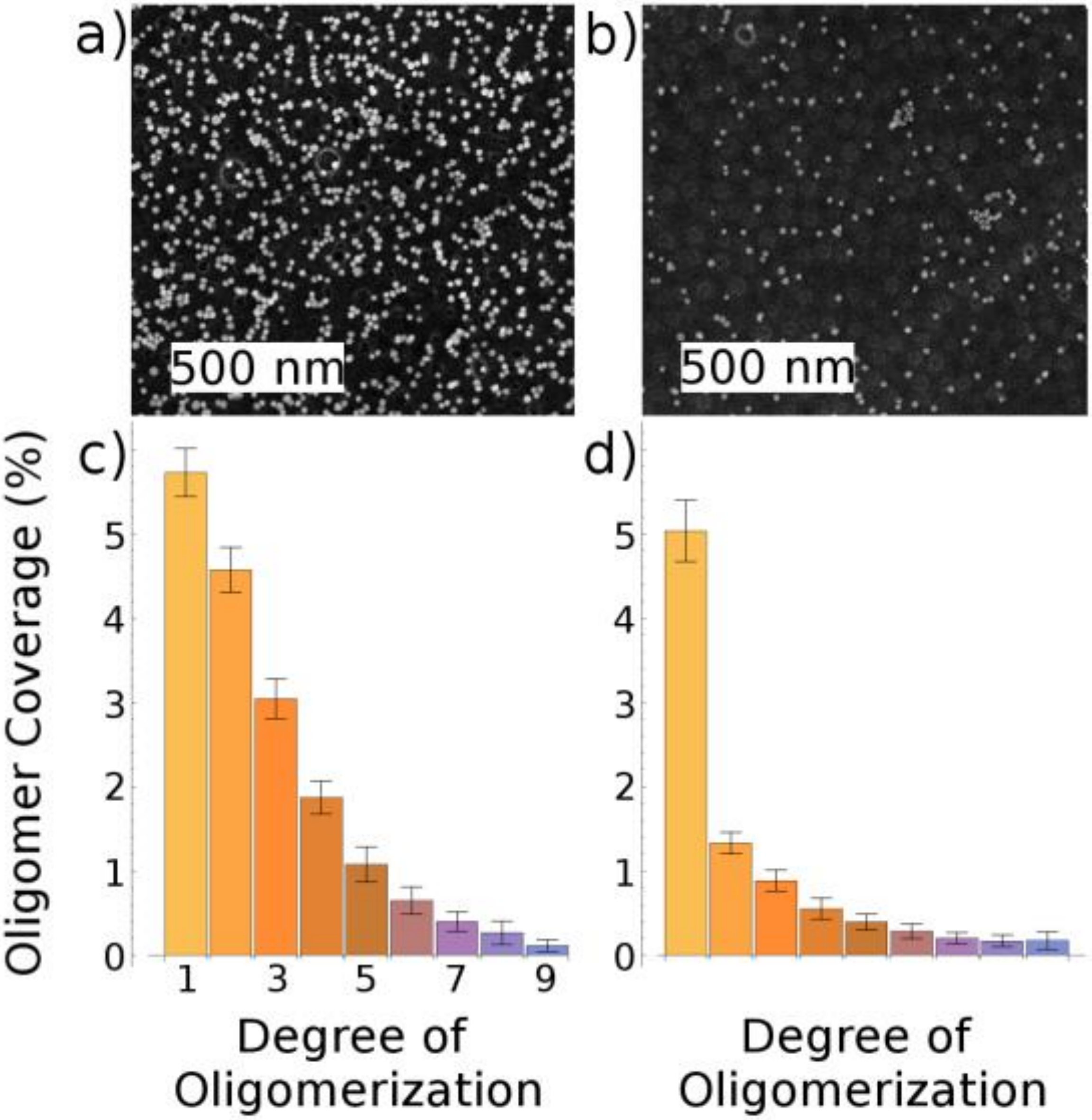

Figure 3.6 depicts nanosphere oligomer substrates assembled with 20 nm diameter Au nanospheres. (a),(b) are representative SEM images taken from the substrate surfaces and (c),(d) depict the oligomerization distribution observed on room temperature and 60 ^{0}C EHD-anhydride substrates, respectively.

3.4 Plasmon Resonance Response of EHD-Anhydride Substrates

The motivation for utilizing a long range driving force to induce chemical crosslinking is uniformity of coverage over large areas. UV-vis spectroscopy is used to probe the large area optical response of EHD-anhydride substrates with 40 nm Au nanospheres. An area of 9 mm x 0.5 mm was measured and normalized attenuation spectrum of 60 ^{0}C EHD-anhydride and, for comparison, Control substrates are plotted in Figure 3.6a,b, respectively. Samples are prepared on transparent conductive substrates as detailed in the methods section and immersed in water during measurements. The absorption spectra of differing oligomer geometries with 0.9 nm gap spacings are simulated with the full-wave finite elements method and overlaid as dotted and dashed curves with the measured attenuation spectrum in Figure 3.6a. This allows us to examine how different geometries contribute to the measured attenuation spectrum.

Four distinct peaks are readily observable within the 60 ^{0}C EHD-anhydride substrate's attenuation spectrum. Simulations identify these peaks to be associated with monomers (536 nm, in good agreement with Mie scattering theory), dimers (686 nm), trimers (740 nm), and quadrumers (782 nm) with 0.9 nm gap spacings. These results indicate that the large area optical response is dominated by anhydride linked oligomers with characteristic 0.9 gap spacing. This is further corroborated by examining the attenuation spectrum of the 60 ^{0}C Control substrate, shown in Figure 3.7b. Unsurprisingly the spectrum is dominated by monomer attenuation. Lines representing the absorption maximum, determined from full-wave simulations of dimers, trimers, and quadrumers are also depicted on Figure 3.7a,b for visual clarity. No discrete oligomer peaks can be observed in the UV-Vis spectrum from the Control sample. Consider that the absorption maximum of a 2 nm gap dimer geometry, calculated in simulations and plotted with a dashed line in Figure 3.7a, shifts the dimer absorption peak by nearly 50 nm from the 0.9 nm dimer case. The blurring of the oligomer response for the Control substrate is attributed to the lack of anhydride mediated control of gap spacings, which results in a continuous distribution of spectral position for any given oligomer geometry.

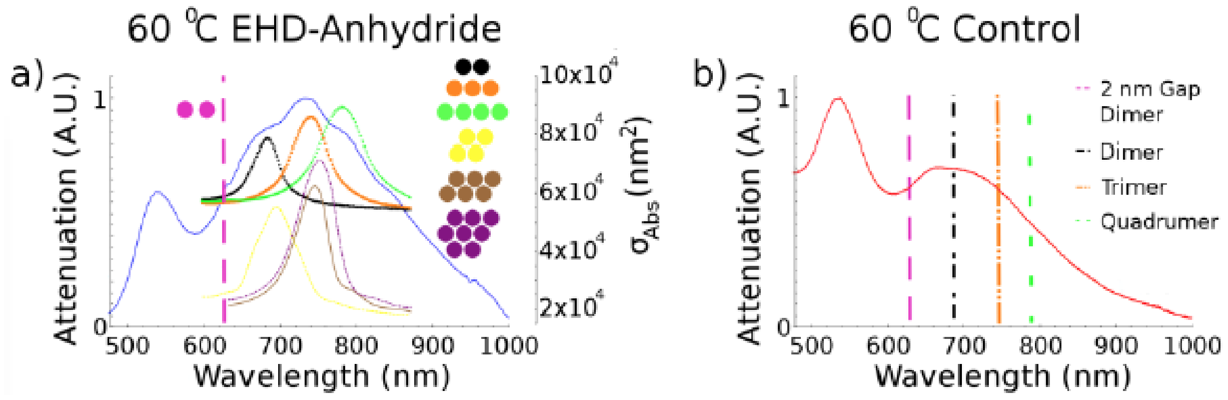

Figure 3.7: Normalized attenuation spectra of a 40 nm (a) 60 ^{0}C EHD-anhydride substrate (solid blue curve) and (b) 60 ^{0}C Control (solid red curve). Spectrum in (a) is overlaid with calculated absorption cross section of observed oligomers with various geometries with color-coded schematic. Close packed absorption cross-sections offset by 3.9x10^4 for visual clarity. Clearly resolved dimer, trimer, and quadrumer peaks within the attenuation spectrum are indicative of the anhydride crosslinking which maintains gap spacings within oligomers. In (b) the calculated maximum value of absorption cross section of a linear dimer, trimer and quadrumer with 0.9 nm gap spacings are shown with a black, red, and green lines, respectively.

Interestingly, UV-vis data and simulations demonstrate that oligomer gap spacings have a more profound impact on the spectral shift of an oligomer resonance than oligomerization. These results are unsurprising when one considers that the plasmon mode is only slightly perturbed by the addition of a nanosphere on an oligomer when it is not in the polarization direction of the incident excitation beam.[249] Consider the spectral position in the simulated absorption spectra of Figure 3.7a of a linear trimer that is shown in the curve composed of orange squares is only blue shifted by 4 nm from a hexamer, shown in the brown curve. Adding two more particles to this configuration, shown in purple, results in a further 3 nm spectral shift. This is a much smaller broadening than occurs when the gap spacing increases from 0.9 nm to 2 nm. This observation is key to understanding why the optical response is narrower than one might expect from the oligomerization statistics in Figure 3.7d. As EHD flow drives the formation of close packed oligomers most oligomers have between three and four particles along any given polarization axis. Consider that the first circular perfectly close packed oligomer to have greater than four particles in a row contains 21 particles; 98% of oligomers observed in the 60 ^{0}C EHD-anhydride substrate contain fewer than 21 particles. More significantly, the distribution statistics by number show that 93% of the surface is composed of oligomers of nine nanospheres or less. Thus, the clear majority of oligomers can be excited with a laser wavelength between the linear trimer and linear quadrumer resonance wavelengths. The peak maxima observed in the UV-vis absorption data of Figure 3.7a is associated with a trimer. By exciting a surface with densely packed oligomers, observed for example in Figure 3.5d,h, we can expect to excite a reproducible distribution of oligomers in any given laser spot diameter.

3.5 SERS Response

In order to further examine the optical uniformity, a SERS enhancement factor (EF) map was acquired over 1 mm^2 of an oxygen plasma treated 60 ^{0}C EHD-anhydride substrate when excited with a 785 nm laser source at 76 μW for 0.1 s exposure times, shown in Figure 3.8. The small gap spacings in oligomers enabled by the (cleavable) anhydride crosslinker provided by template-seeded EHD flow should yield a model system for SERS. The EF is determined by assembling a self-assembled monolayer of BZT onto the substrate and observing the intensity of 1573 cm^{-1} vibration band. The raw SERS data is converted into an EF using the following equation:[137]

$$EF = (I_{SERS}/N_{SERS})/(I_{neat}/N_{neat}) \tag{3.1}$$

Where I_{SERS} is the SERS intensity from the sample, I_{neat} is the neat Raman intensity from the bulk solution, N_{SERS} is the number of molecules participating in the SERS, and N_{neat}, is number of molecules participating neat measurements. The number of BZT molecules participating in the SERS measurements is determined through the average surface coverage of nanospheres assumed to be coated with a self-assembled monolayer of BZT over the mapped region, obtained using

SEM images. The average coverage is used for all EF calculations. The SERS map shows an average SERS EF value of $1.4x10^9$. The data shows a 10% relative standard deviation (RSD) of the 1573 cm^{-1} peak across the mapped area. SERS EF is approximately proportional to the electric field enhancement taken to the fourth power, indicating that the anhydride linked oligomers achieve extraordinarily large electric field enhancements while providing point to point uniformity in the spot size of the optical beam. The low signal deviation is a result of the seeded growth oligomerization protocol that is optimized by controlling the temperature, block copolymer template, and EHD flow.

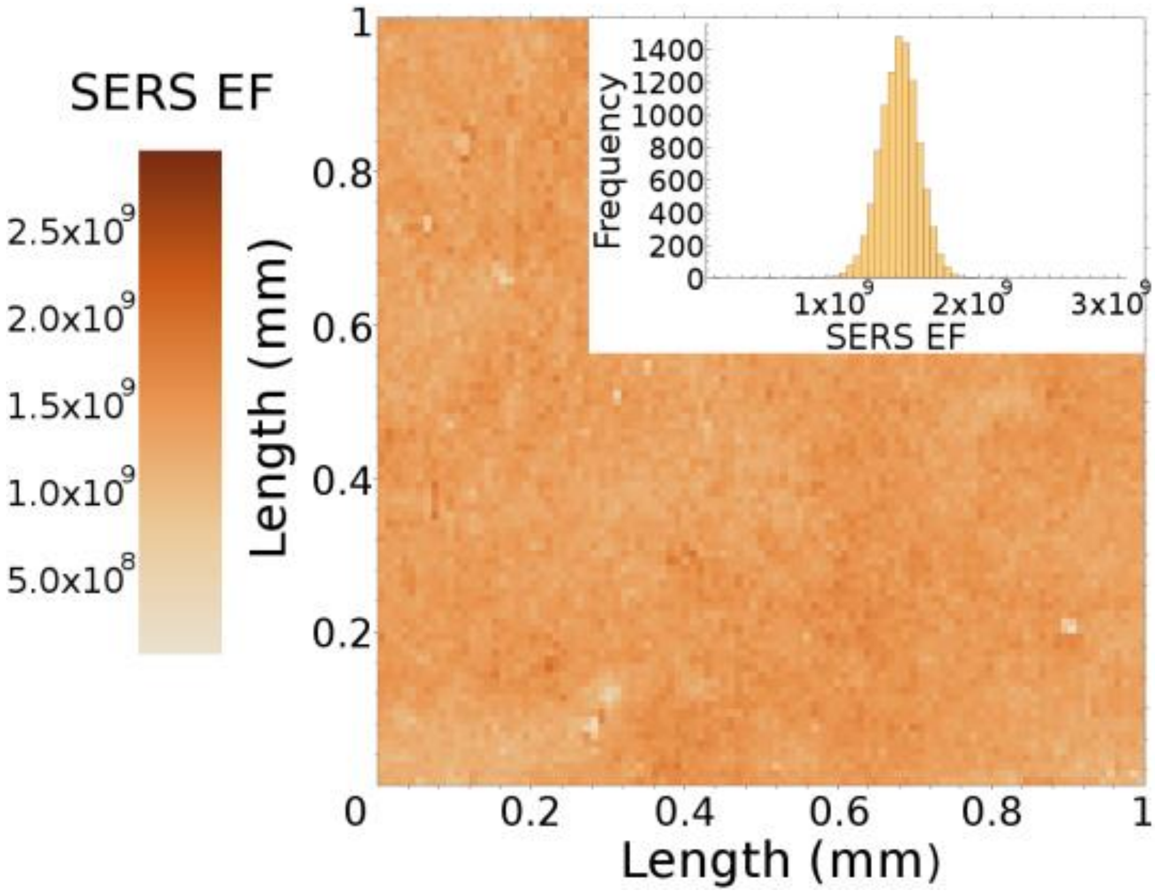

Figure 3.8**:** SERS EF map of an oxygen plasma treated 60 ^{0}C EHD-anhydride across a 1 mm x 1 mm area where each measurement is separated by 10 µm, corresponding to 10,000 measurements. The laser power is 76 µW with a 0.1 s exposure time. Inset: Histogram of measured SERS EF values with 57 bins depicting a mean of 1.4×10^9, full width at half maximum of 3.5×10^8, and RSD of 10%.

3.6 Conclusion

In this chapter, we propose a novel method for the seeded chemical assembly of discrete oligomers with high density over large areas. Electrohydrodynamic (EHD) flow is found to serve as a long-range driving force to bring Au nanospheres together, enabling local chemical crosslinking at an electrode surface using carbodiimide chemistry. We confirm the anhydride linker assembly mechanism with SERS spectroscopy, and show with transmission electron microscopy that the resulting gap spacings between particles are 0.9 nm only when an applied bias is present to generate EHD flow. Parameters affecting assembly are elucidated by atomistic simulations. The probability of forming an anhydride bond between nanospheres is the product $P_{Reaction} = P_{entropy}(C, d_{\mathrm{Au}}) \times P_{enthalpy}(T) \times P_{Collision}$, where EHD flow is a key factor to increase the probability of nanosphere-nanosphere collisions. The effect on oligomer bulk optical response is observed with UV-vis microscopy. Individual oligomer contributions to optical response is elucidated with electromagnetic full-wave simulations where it is found to depend more strongly on nanogap dimensions versus oligomer geometry. Finally, the electric field enhancement produced by oligomers, as determined with surface enhanced Raman scattering (SERS) spectroscopy, is shown to be relatively uniform over a 1 mm^2 area, with a relative standard deviation in signal of 10%. This is significant as the SERS enhancement factor has an average value of 1.4×10^9. There are typically tradeoffs between uniformity and high electric field enhancements in SERS measurements. The results demonstrate that the use of long range driving forces to drive chemical crosslinking between nanospheres represents a new tool for scientists and engineers to fabricate plasmonic devices using colloidal assembly. In chapter 4, this tool will be extended to templated assembly, permitting precisely defined EHD flow fields for fabricating uniform oligomer structures and metasurfaces.

3.7 Methods

3.7.1 Materials

Random copolymer Poly(styrene-co-methyl methacrylate)-α-Hydroxyl-ω-tempo moiety (PS-r-PMMA) (M_n = 7,400, 59.6% PS) and diblock copolymer poly(styrene-*b*-methyl methacrylate) (PS-*b*-PMMA) diblock copolymer PS-*b*-PMMA (M_n = 170-*b*-144 kg mol^{-1}) were purchased from Polymer Source, Inc. (Dorval, Canada). Gold nanospheres diameter of 20 nm and 40 nm with lipoic acid functionalization were purchased from Nanocomposix (San Diego, CA).

Si(001) wafers with resistivity of 0.004 ohm-cm were purchased from Virginia Semiconductor (Frederickburg, VA. Hydrofluoric acid (HF) was purchased from Fisher Scientific (Pittsburgh, PA). 2-(*N*-morpholino)ethanesulfonic acid (MES) 0.1M buffer, 1-ethyl-3-[3-dimethylaminopropyl] carbodiimide hydrochloride (EDC), and N-hydroxy sulfosuccinimide (s-NHS) were purchased from Pierce (Rockford, IL). Dimethyl sulfoxide (DMSO), ethylenediamine, benzenethiol, toluene, ethanol, isopropanol (IPA), aniline, methylamine, potassium carbonate, and 52-mesh Pt gauze foil were all purchased from Sigma Aldrich (St. Louis, MO). Nanopure deionized water (DI) (18.2 MΩ cm^{-1}) was obtained from a Milli-Q Millipore System. Single layer graphene on ultra-fine mesh copper TEM grids and 50 mesh copper TEM grids were purchased from Ted pella.

3.7.2 Nanoantenna Oligomer Substrate Fabrication

Lamella PS-b-PMMA block copolymer is spin-coated onto a HF-cleaned, heavily doped Si wafer and annealed as described in previous work.[137] PMMA regions are selectively functionalized with amine end groups by first immersing the entire substrate in DMSO and then in ethylenediamine/DMSO solution (5% v/v), both for 5 minutes without rinsing between steps. The Si substrate coated with functionalized copolymer is then washed with IPA for 1 minute and dried under nitrogen.

Au nanosphere solution (0.1mg/mL, 3 mL) is added to a 10 mL beaker. Freshly prepared s-NHS (20 mM) in a MES (0.1 M) buffer (35 μL) is added to the beaker and swirled. Next, freshly prepared EDC (8 mM) in a MES (0.1 M) buffer (35 μL) is added to the beaker and swirled. For heated samples, the beaker is placed on a hotplate and brought to 60 ^{0}C°C. Similar to previous work[141] a 1 cm x 1 cm functionalized copolymer-coated Si substrate is placed into the solution vertically, held in place as the anode with alligator clips that do not contact the nanosphere solution.

1 mm away from the substrate, a 1 cm x 1 cm Pt mesh is placed into the solution vertically, held in place as an anode with alligator clips that do not contact the nanosphere solution. A DC Regulated Power Supply was used to apply a voltage of 1.2 V for 10 minutes. The substrate, Pt mesh, and beaker are rinsed with IPA for 1 minute and dried under nitrogen. This process is repeated with the same substrate and fresh nanosphere solution as described above, but with 25 μL of EDC and s-NHS solution. Two growth steps are necessary to obtain oligomers with enough Au nanospheres to be on resonance at the 785 nm illumination wavelength used for surface enhanced Raman scattering (SERS) measurements. The second growth step is performed with reduced EDC and s-NHS concentrations to promote anhydride formation, which has been shown to increase with decreased concentration.[250] Oxygen plasma treated samples are treated with a 50W oxygen plasma etch for 120 seconds.

3.7.3 Assembly on Alternative Substrates

Substrates fabricated for TEM characterization are fabricated as above, but on copper TEM grids for field-free Control substrates and graphene membrane TEM grids for EHD-anhydride samples. PS-b-PMMA block copolymer templates on ITO-coated glass are fabricated identically, but using ITO-coated glass that is oxygen plasma etched at 100W for 1 minute instead of the Si wafer.

3.7.4 Characterization

After nanoantenna oligomers are assembled onto the block copolymer-coated Si substrate, images are collected with a Magellan XHR SEM (FEI), and a CM-20 TEM (Philips).

UV-Vis absorption spectra are taken of nanoantenna oligomer on ITO-coated glass substrates taped (away from the beam path) onto a quartz cuvette. The cuvette is then filled with water and imaged with a Shimadzu UV-1700 absorption spectrometer.

Raman spectroscopy measurements are conducted using a confocal Renishaw InVia micro Raman system with a laser excitation wavelength of 785 nm. All surface enhanced Raman scattering (SERS) measurements are taken at 73 μW with exposure of 0.1 seconds for map measurements and 1 second for individual spectra on 60 ^{0}C°C EHD-anhydride substrates. All measurements use a 60X water immersion objective with a 1.2 NA, immersed in DI water. SERS enhancement factors (EF) are calculated according to the method of Cai et al.[251] Briefly, N_{SERS} is estimated by determining the average surface area of Au via SEM and using an experimentally determined density REF of benezenethiol (BZT) on Au surfaces. N_{Neat} is determined by aliquoting neat BZT onto a Si wafer and a glass coverslip affixed on top. The effective height of the BZT – 4.07μm – is determined by measuring the Raman intensity – depth profile and the bulk density of BZT is used in the calculation. As both the neat and SERS measurements are obtained with the same microscope objective the EF is not a function of spot size. I_{SERS} and I_{NEAT} are normalized by laser power and acquisition time. Neat BZT benzenethiol (BZT) measurements are taken with the same laser and 60X water immersion objective but with 73 mW laser power and 10 second exposure time. The number of molecules probed in SERS measurements is determined by finding the average nanosphere coverage with SEM and calculating the BZT density using values from Aggarwal et al.[252]

3.7.5 Finite Element Simulations:

Full-wave simulations (frequency domain finite elements method solver) are implemented in CST Microwave Studio (CST AG). Absorption spectra are simulated for several nanosphere oligomers: dimer, linear trimer, linear quadrumer, close-packed quadrumer, close-packed hexamer, and close-packed octamer. Au nanospheres with diameter of 40 nm are used with permittivity from the Drude model with parameters extracted from Grady et al.[253] We use a 0.9

nm gap between nanospheres, consistent with both observed and the modeled length of an anhydride linker. The nanospheres were previously shown to be partially embedded in PMMA,[141] with the PMMA layer thickness set to 40 nm, and the center of the nanospheres 8 nm above the layer. Below the PMMA layer is 150 nm layer of ITO on top of a 2 μm layer of glass. The relative electric permittivity of water, PMMA, glass and ITO used in the simulations are 1.77, 2.47, 2.3207, respectively, and adapted from Moerland et al.[254] The permittivity in the gap region is uncertain as the excitation source will probe a volume composed of the anhydride linker, aqueous solution, and copolymer environment. In order to account for this, a parameter sweep of the gap permittivity is performed using the dimer configuration, and the permittivity which best corresponds to the observed dimer structure peak at 686 nm is determined to be 2.25. This parameter is reasonable based on an estimate of the average permittivity of the materials[255] and further verified by the correspondence between the trimer configuration's simulated plasmon resonance and the trimer peak observed in the EHD-anhydride sample.

Oligomers are excited with plane wave illumination at normal incidence with electric field polarization along the axis of the linear oligomers, and the absorption cross section of the structure is determined.

3.7.6 Atomistic Simulation:

We construct a periodic supercell containing a total of 288 gold atoms arranged in six layers of (111) plane. The simulation box dimensions in *xy* plane are roughly 20 Å × 17 Å. The size of the simulation cell in *z* dimension varies between 23 Å and 32 Å to provide different nanogap distances. We change the number of water molecules from 30 to 120 to fill the free space between organic molecules in the nanogap. The number of water molecules depends on the concentration of organic molecules and gold surface-to-surface distance. We first construct all organic molecules, both lipoic acid (LA) and O-acylisourea (OA), in AVOGADRO software[256,257] and minimize the configurations using UFF forcefield[258]. We subsequently insert these organic molecules inside the nanogap such that sulfur atoms are next to the gold surface. We perform our simulations using ReaxFF potential,[242] a reactive force field designed based on the notion of bond-order parameter[82,83] and the electronegativity equalization method[84] to update variable charges at each step of the atomistic simulation. Here, we use a transferable set of ReaxFF parameters that were trained for bio-molecules and their interaction with gold surfaces.[241] We use LAMMPS simulation package in all MD calculations.[259] MD simulations are carried out in the canonical (NVT) ensemble using Nosé-Hoover thermostat[240] with a relaxation constant of 0.01 ps. The time step in MD simulations is set to 0.1 fs.

In simulations for carbon-oxygen distance analysis, we construct nine periodic simulation boxes with gold-gold surface distances of 11, 15, and 20 Å, each containing three levels of OA-LA concentrations. In low and medium concentrations respectively, we attach one OA to one side of Au slab, while we place one and four LA molecules at the other surface. For the high concentration case, two OA molecules in one side face four LA groups on the other side of the gold slab. We adjust the number of water molecules in each simulation to achieve ambient conditions at equilibrium. We first perform potential energy minimization to relax the structure. We subsequently relax the position of water molecules in a 20-ps long MD simulation, while the rest of atoms are fixed in their position. Afterwards, we further relax the system in canonical ensemble for 100 ps. A 400 ps-long production phase produces the MD trajectory saved at intervals of 0.1 ps.

For MEP simulations, we first construct three simulation boxes with nanogaps of 13, 15, and 17 Å. We place one OA molecule on one surface and one LA group on the other. We adjust the number of water molecules to obtain ambient pressure at equilibrium. The purpose of these simulations is to capture a reasonable initial and final state on each side of the energy barrier for NEB calculations. To accelerate the formation of intermediate state, we apply a bias harmonic potential between reactive carbon and oxygen atoms. This spring should be strong enough to keep the reactive core in the proximity of the transition state. Also, a large spring constant biases the configuration space by preventing the organic molecules to properly relax to accommodate short C-O distances. Here, we determine the value of the spring constant to be roughly 100-200 kcal/mol x Å^2. The equilibrium C-O distance is found to be around 3 Å. Since the number of degrees of freedom corresponding to atoms in organic molecules is large, it is critical to capture proper initial configurations in z(I, II, and III) states to obtain a smooth MEP. We take the initial and final states for NEB calculations right before and after transition states during a small window of time (less than 100 fs) in biased MD simulations.[244,245] We find that the metastable intermediate complex (II), Figure 3.4, is at least stable for 20 ps in MD simulations at room temperature. Therefore, we divide the reaction path into two separate stages, one from state I to state II, and the other from state II to state III, see states in Figure 3.2. Since the reaction does not directly involve water molecules, we do not include them in NEB calculations. This also helps with NEB's difficulty in determining transition state in the presence of soft degrees of freedom. For the second stage of the chemical reaction, it is enough to elevate the temperature to 600 K to reach state III. To perform NEB calculations, we use the standard LAMMPS implementation of a two-stage NEB procedure.[260] First, we use the standard NEB[245] by constructing replicas with linear interpolation, and imposing inter-replica forces to find MEP. We subsequently use a barrier-climbing technique to find the true transition state.[245]

3.8 Supplemental Information

Au nanospheres were assembled on PS-b-PMMA diblock copolymer templates on a Cu TEM grid without an applied bias to examine gap spacings between nanospheres in the absence of electrohydrodynamic flow. Figure 3.9a depicts a TEM micrograph of 40 nm radius Au nanospheres assembled on the copolymer coated copper TEM grid. The gap spacing of the dimer observed in Figure 3.9 is approximately 3.6 nm. Without EHD flow, gap spacings are not controlled by anhydride crosslinking, leading to a high degree of variability in observed values. Figure 3.9b depicts a SERS spectra acquired from a Control substrate assembled as described in the main text. Control substrates are sparsely populated with oligomers and have increased gap spacings with respect to EHD substrates. Thus higher laser excitation intensities are necessary to collect SERS spectra with similar intensity as the EHD substrates; here 7.3 mW laser power is used compared to the 73 μW used for EHD substrates. No significant peaks are observed that correspond to the anhydride vibrational modes identified in the main text and highlighted with a dashed line in Figure 3.9b. The scattered intensity near 1600 cm^{-1} is associated with carboxylate vibrational modes from the lipoic acid.

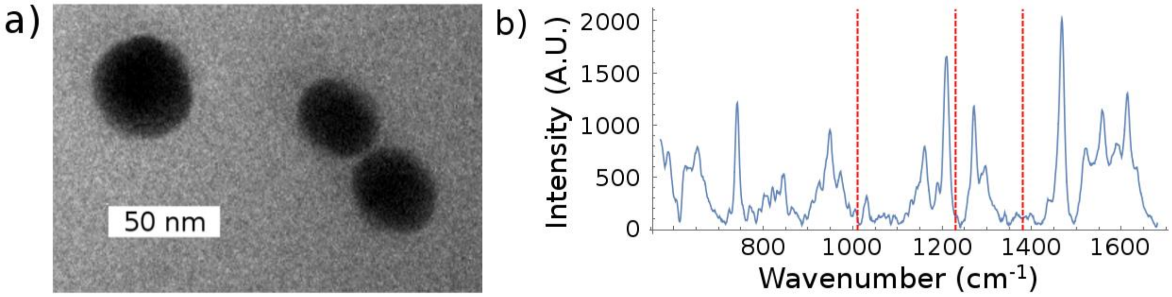

Figure 3.9: (a) TEM micrograph of an assembled oligomer without an applied bias Control substrate. (b) Surface enhanced Raman scattering (SERS) spectra from a similar Control substrate acquired at 785 nm with a laser power of 7.3 mW for 1s. Dashed red lines indicate spectral positions associated with anhydride linked samples.

To analyze the degree of oligomerization of a given oligomer and the number of nearest neighbors of a given Au nanoparticle, Wolfram Mathematica^{TM} is implemented for image analysis. SEM images, an example shown in Figure 3.10a, are first binarized, and oligomers are distinguished from one another as being separate collections of foreground pixels, or morphological components. Single Au nanoparticle are identified through their circularity, defined as the ratio between the equivalent disk perimeter length and the perimeter length of a polygon formed by the centers of each perimeter element, only morphological components over a certain threshold of circularity are determined to be single nanoparticles. The remaining morphological components are then divided into component nanoparticles using a modified Euclidean distance transform approach, which is designed for implementation in SEM images, where edge effects can make identification of small nanoparticles in close packed structures difficult. For each morphological component, the original image is again binarized using local adaptive binarization, shown in the left side of Figure 3.10b. A Euclidean distance transform (figure distance transform) then reveals the nanoparticle centers as local maxima, shown in the right side of Figure 3.10b. The maxima are then used to determine the center of each nanoparticle and a distance threshold is used to determine the number of nearest neighbors of a given nanoparticle. From these images the number, density, and nearest neighbor statistics are obtained as shown in Figure 3.10c and d, respectively.

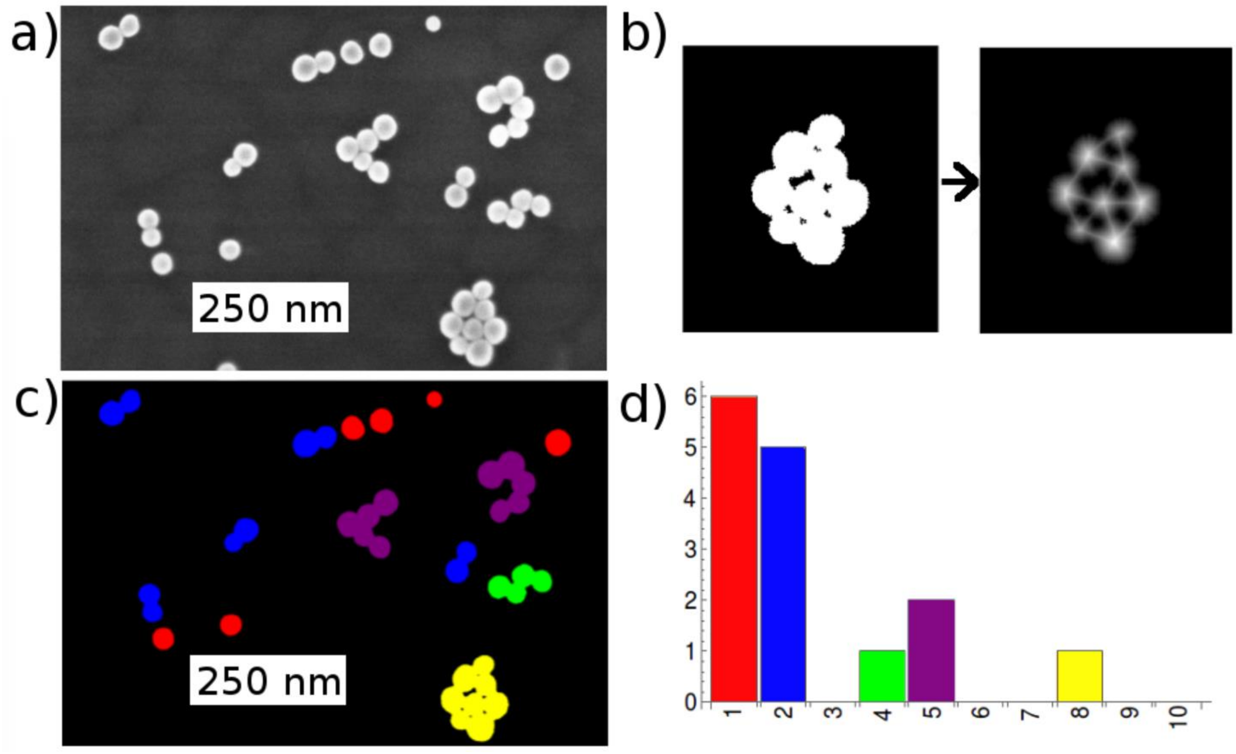

Figure 3.10 (a) SEM image of 40 nm EHD-anhydride substrate (b) left: local adaptive binarized oligomer, right: distance transform of the left image (c) Colorized image of the SEM image with oligomers identified (d) Histogram of the number oligomers identified for each degree of oligomerization.

Figure 3.11 depicts a TEM micrograph of a dimer assembled on a 20 nm EHD-anhydride substrate. The pictured substrate was assembled at room temperature. The gap spacing is consistent with anhydride linking, indicating that EHD flow induced anhydride crosslinking still occurs with particles as small as 20 nm.

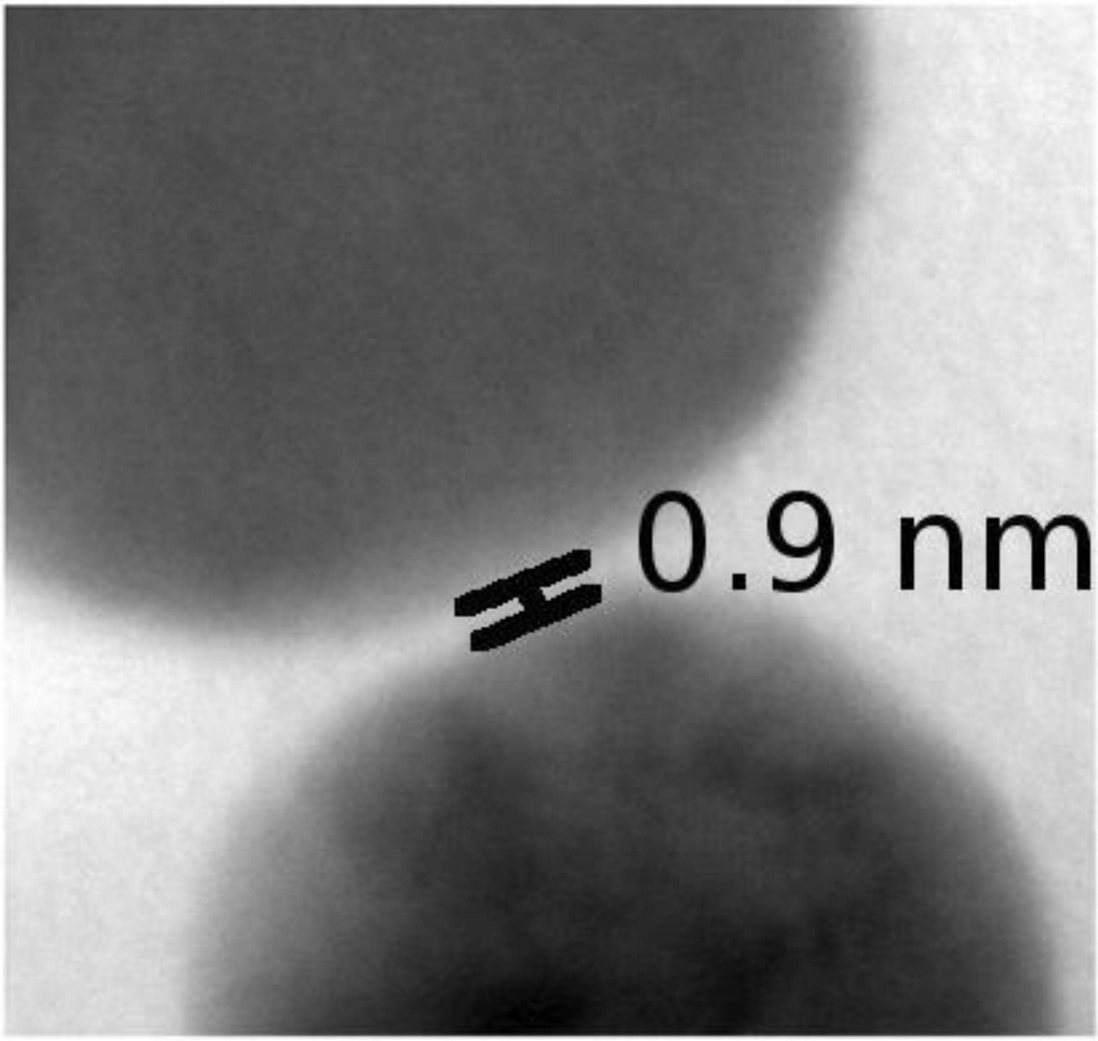

Figure 3.11: TEM micrograph of a dimer assembled on a 20 nm EHD-anhydride substrate.

Multiple deposition steps may be used to increase the degree of oligomerization on a sample. Figure 3.12 depicts an SEM image of a 60 ^{0}C EHD substrate after four deposition steps, leading to greater oligomer density than the two deposition step samples presented in the main text. Here, we begin to see the formation of three dimensional oligomers as electrophoresis drives Au nanospheres onto existing oligomers.

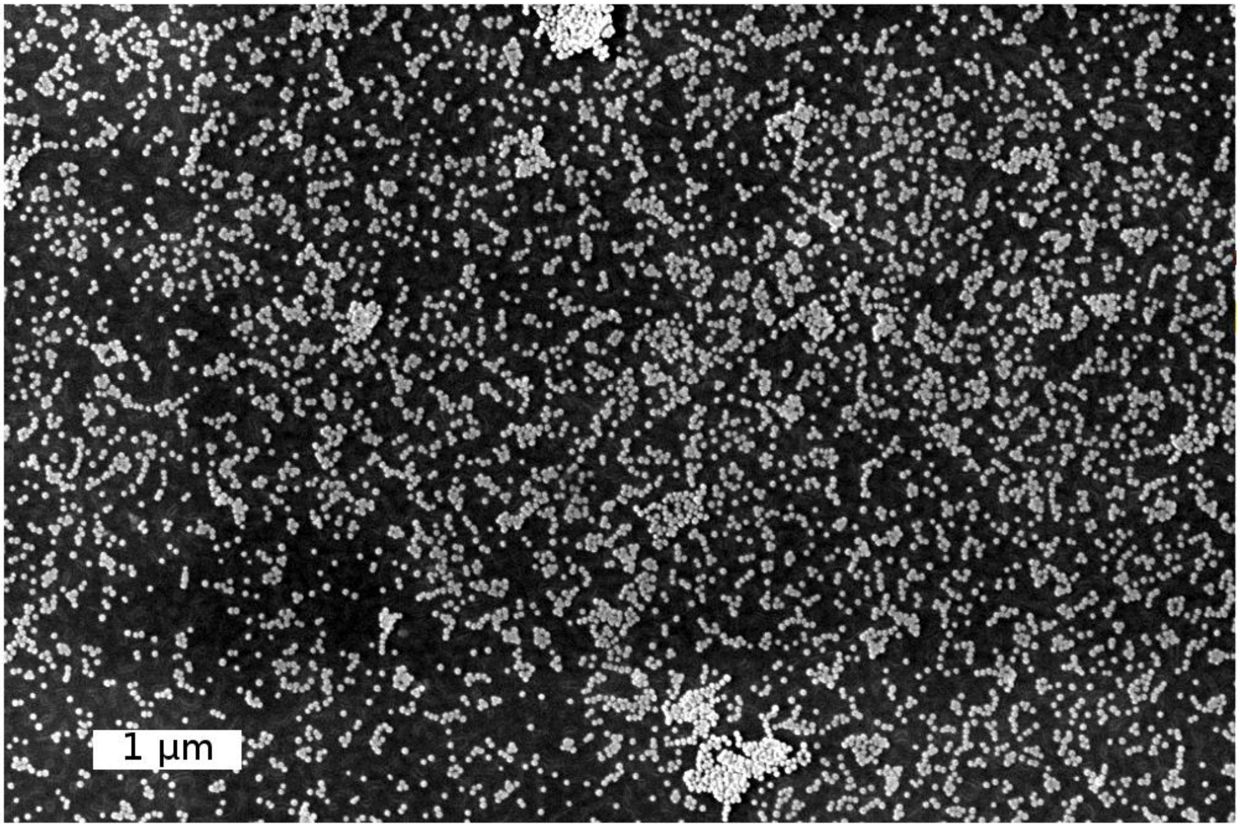

Figure 3.12: SEM image acquired from a 60 ^{0}C EHD substrate after four depositions with 35 µL EDC and s-NHS solution in each deposition.

Figure 3.13 depicts oligomers on a 60 ^{0}C EHD substrate with the underlying block copolymer template visible in the imaging conditions. Au nanospheres within the oligomers are observed on the amine-functionalized PMMA domains (dark with bright borders) and on PS regions. Anhydride crosslinking of Au nanospheres to one another, where one or more of the nanospheres in the oligomer is bound to the PMMA domain, will result in some nanospheres on the PS region. PMMA domains appear darker with bright edges due to topography changes during the ethylenediamine functionalization. Thus they can be observed in SEM images when using lower current and accelerating voltage.

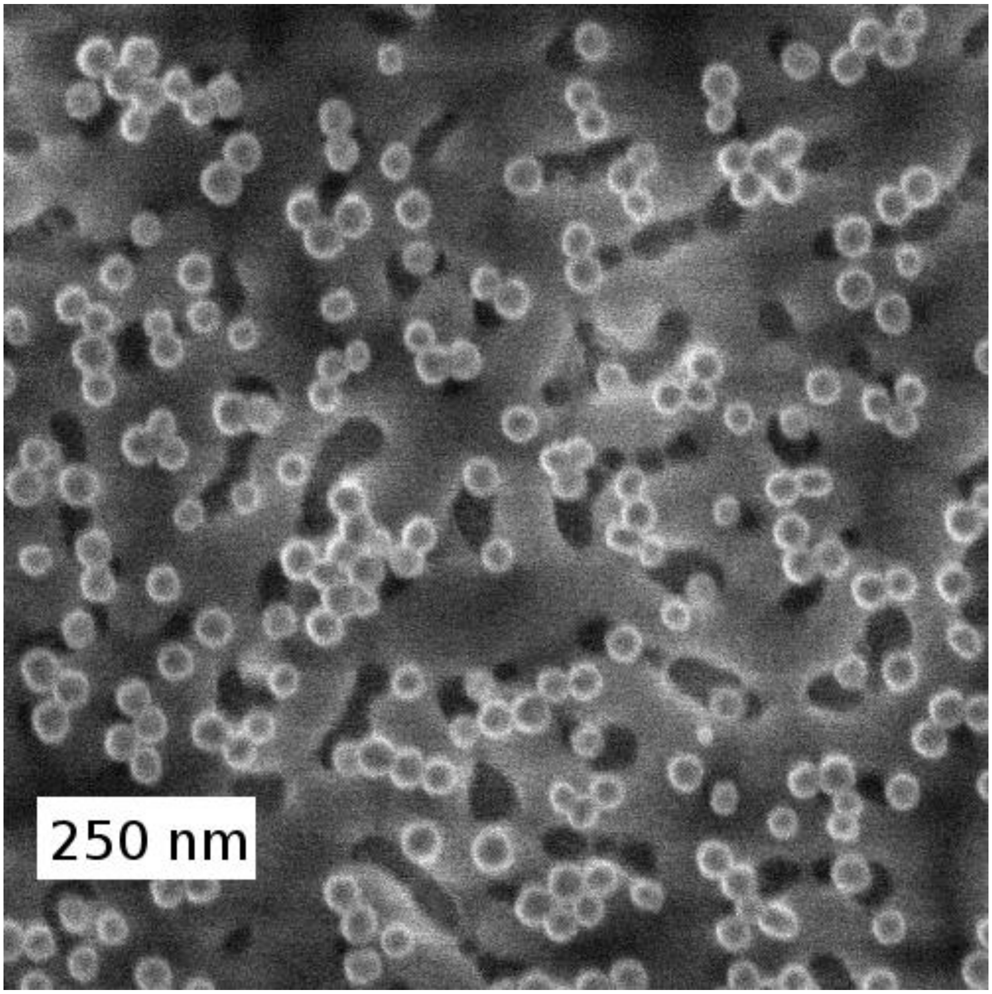

Figure 3.13: SEM image acquired at 2 KeV and 13 pA from a 60 ^{0}C EHD substrate to highlight the PS-b-PMMA block copolymer template.

Additional SERS data is also provided here to further demonstrate reproducibility and stability of SERS response from EHD anhydride substrates. Figure 3.14 includes SERS spectra acquired from samples undergoing the same treatments discussed in section 3.2 in Figure 3.2b,c. The five different spectra in each panel come from 5 different samples. Reproducible SERS signal is clearly observed for the (a) anhydride, (c) aniline and (d) benzenethiol. SERS signal after treatment of (b) methylamine has peaks associated with carbonate groups in lipoic acid in the region around 1600 cm^{-1}. The data is slightly noisier that would isbe consistent with less ordered molecular orientations in the hotspot. Figure 3.15 depicts SERS spectra before (lower) and after (upper) soaking samples in DI water for 7 days after undergoing treatments discussed in the main body text in Figure 3.2b,c. No additional treatment is performed after the 7 day soak. The similarity of spectra observed after soaking the samples demonstrates that the Au nanosphere oligomers are stable in aqueous solution.

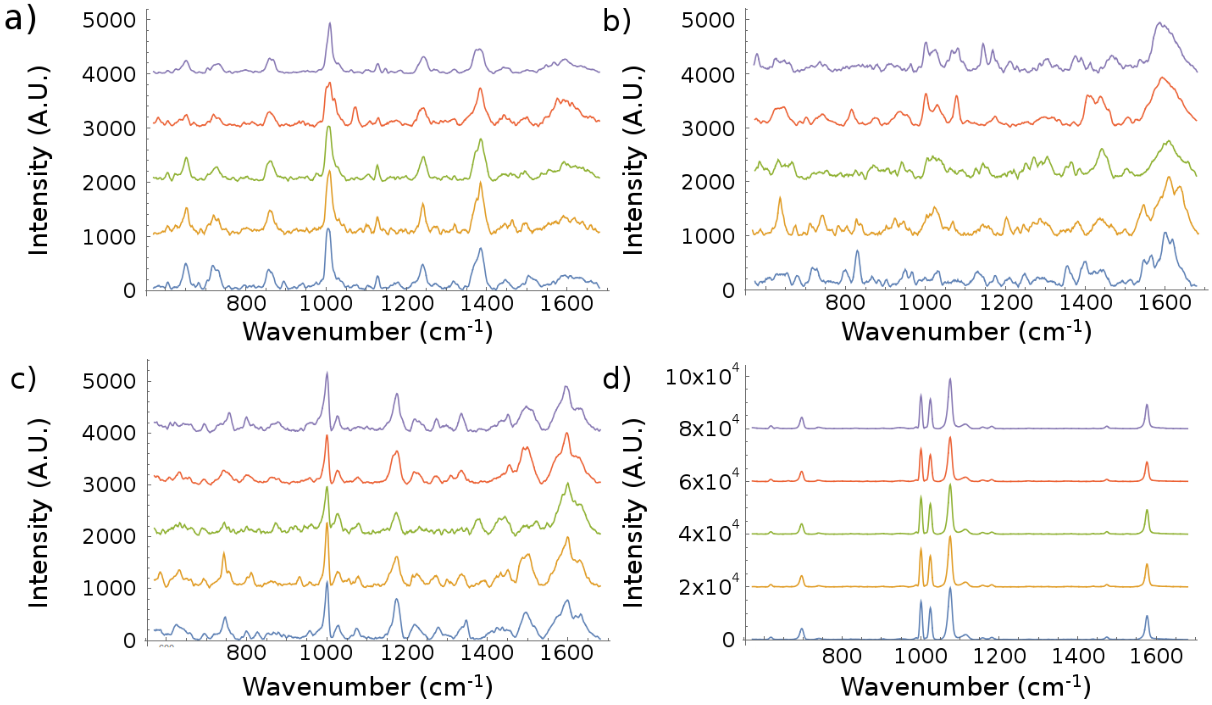

Figure 3.14: SERS spectra acquired from 5 different EHD-anhydride samples in the following conditions: a) untreated sample b) methylamine treated sample c) Aniline treated sample d) plasma cleaned benzenethiol treated sample.

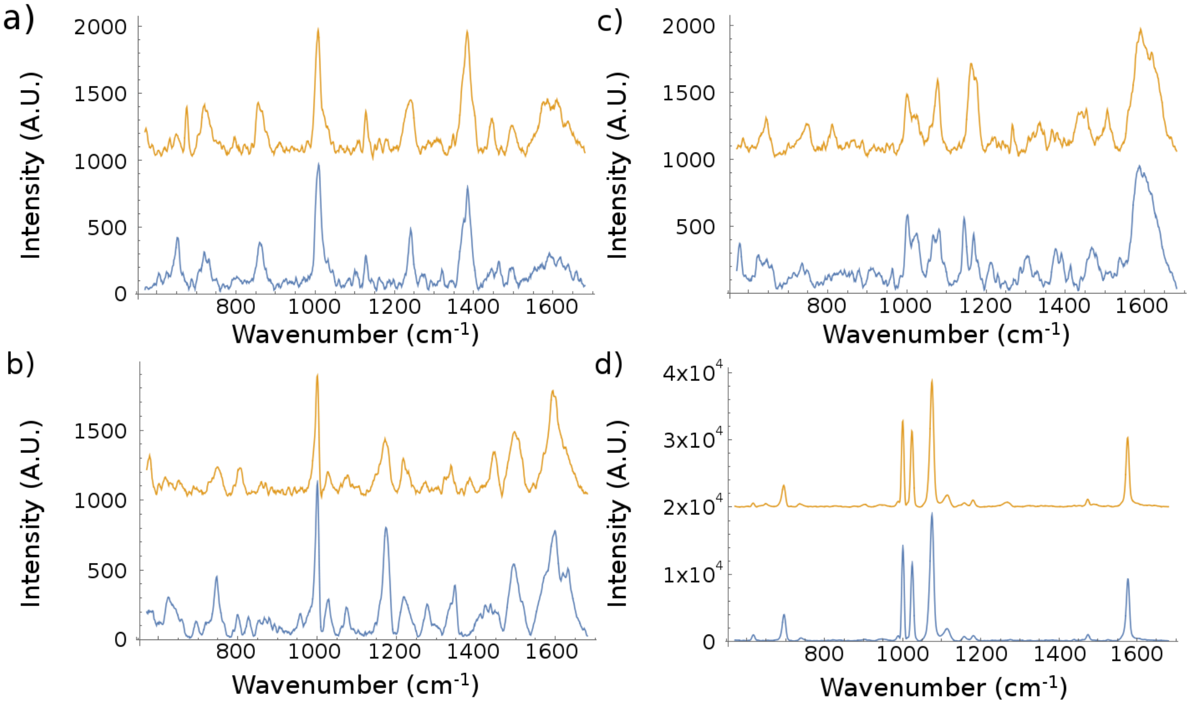

Figure 3.15: SERS spectra acquired from: an a) untreated sample, b) methylamine treated sample, c) aniline treated sample, and d) plasma cleaned benzenethiol treated sample after soaking the sample in DI water for 7 days.

Chapter 4

Templated Electrokinetic Directed Chemical Assembly for the Fabrication of Close-Packed Plasmonic Metamolecules

4.1 Introduction

Metal architectures using colloidal nanoparticles as meta-molecule building blocks have shown great promise as a scalable self-assembly method to control light matter interactions in large area devices. The magnetic interaction of light and matter at optical frequencies is negligible compared to its electric counterpart[249,261–263] as natural magnetism fades away at infrared and optical frequencies. By forming discrete sub-wavelength clusters, devices that rely on narrow-band resonances — i.e., Fano resonances — based on dark (i.e., low scattering) electric and magnetic resonances can be realized.[201,203] Conventional split-ring resonators, that in principle could provide narrow band resonances, are prohibitively difficult to scale down to optical wavelengths,[264] whereas coupled nanospheres can be scaled.[265] As these architectures are composed of subwavelength building blocks, traditional optical lithography methods cannot be utilized for large area device fabrication. Optical magnetic resonators composed of nanoparticle building blocks have shown that large magnetic dipoles, and suppression of electric dipoles are achievable.[266,267] These systems lead to highly confined, extremely large magnetic fields which are designed to probe magnetic transitions in molecules at optical frequencies.[268] Optical frequency magnetic resonator structures are of special interest for templated self-assembly (directed assembly) because they can be composed of close-packed nanoparticles.[208] Still, control of gap spacings and particle geometries within metamolecules is an ongoing challenge in directed assembly.

In this chapter, in order to assemble plasmonic metamolecules, chemical crosslinking, using carbodiimide based chemistry, between nanoparticles and templates is selectively performed on surface using electrohydrodynamic (EHD) flow as a driving force. Templates are composed of Au pillars on the substrate's surface. These pillars serve as perturbations that drive EHD flow and electrophoresis, directing the assembly of nanoparticles around the pillars. Induced by an electric field in colloidal solution, EHD flow promotes lateral motion and close-packing of particles in colloid at an electrode surface, thereby, assembling transient close-packed structures.[269,270] We freeze in transient structures with local carbodiimide[131,137,271] crosslinking. The EHD flow increases the nanoparticle residence time at small interparticle spacings; this is necessary to enable the O-acylisourea – lipoic acid pathway of the carbodiimide crosslinking to occur in high yields. This chemical pathway leads to an anhydride bridge between ligands on Au nanoparticles, yielding permanent close-packed nanoparticle oligomers.[272] The metamolecules produced with this technique are simulated and shown to deliver large magnetic resonances at optical frequencies. This work represents a step forward for directed assembly of nanoarchitectures to create metasurfaces by demonstrating the influence of both the template and physical driving forces in assembly to control oligomer morphology and uniformity of gap spacing. Close-packed plasmonic metamolecules fabricated using the EHD – anhydride assembly method have already been demonstrated as effective biosensors using their electric field enhancements,[273,274] now we pave the way forward to their use in sensing magnetic transitions in biomolecules.

4.2 Results and Discussion

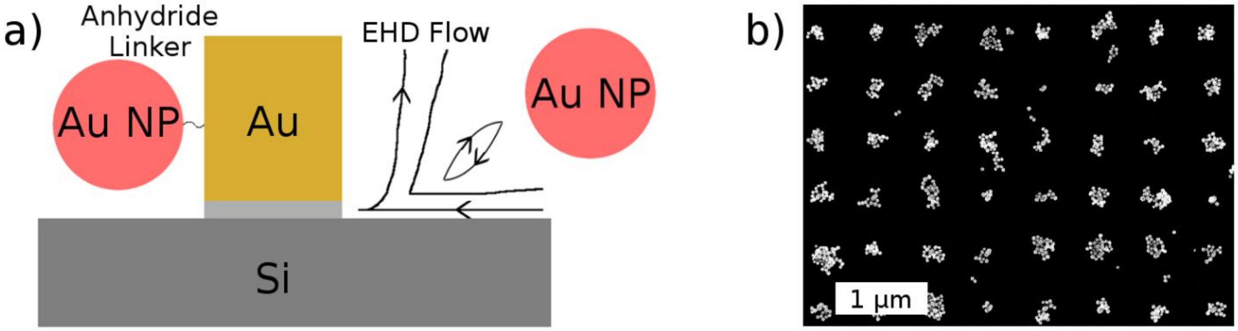

Figure 4.1: a) Schematic of electrohydrodynamic nanoparticle assembly on block copolymer template. Inset depicts electrohydrodynamic flow fields generated by nanoparticles bound to the surface. b) Close-packed plasmonic metamolecule array before chemical removal of excess nanoparticles.

Metamolecules composed of close-packed nanoparticles are assembled via a directed assembly approach. Electrohydrodynamic (EHD) flow enables carbodiimide crosslinking of nanoparticles to a template composed of Au nanopillar arrays. The EHD flow – crosslinking assembly system, depicted in figure 4.1a, works as follows: a conductive substrate with Au nanopillars and counter electrode are placed in a carboxylic acid functionalized Au nanoparticle electrolytic colloid solution and a bias is applied across the electrodes. At the substrate surface, the ion double layer around a nanopillar is deformed by the applied bias resulting in an ionic concentration gradient. This concentration gradient results in lateral osmotic flows towards the nanostructure (equilibrium charge electroosmosis) and the action of the double layer on the substrate’s double layer also causes lateral osmotic flows towards the nanostructure (induced charge electroosmosis), the combined osmotic flows are called EHD flow. An additional induced charge electroosmotic flow is generated by the hydrophobic/hydrophillic interface between the amine terminated Au nanopillars and the alkane terminated Si surface. The Au nanoparticles are driven to the templated substrate due to electrophoretic sedimentation and are entrained by EHD flows generated by the nanopillars. This results in assembly around the pillar to form close-packed assemblies. When the field is removed, assemblies will rapidly dissolve into solution, so crosslinking is necessary to “freeze-in” these assemblies for use in plasmonics.

In order to prevent the dissolution of the assembled metamolecules, chemical crosslinking is used to covalently bind the nanoparticles and nanopillars together. The carbodiimide activates the carboxylic acid moieties on the nanoparticles, forming O-acyliosureas that participate in crosslinking reactions. Two crosslinking pathways are possible: 1) the O-acylisourea attacks the amine groups on the nanopillars, linking the nanoparticles via an amide bond. 2) the O-acylisourea attacks the carboxylic acid group on the nanoparticles, linking together nearby nanoparticles via an anhydride bond. The latter process does not occur in solution due to the long timeframes necessary for reaction and is selectively enabled by the extension of the residence time of nanoparticles in assemblies generated by EHD flow. The result is permanent, anhydride linked, close-packed assemblies on a conductive substrate. The anhydride linker is a key advantage of this assembly system. Plasmonic systems with small gap spacings, or hotspots, are incredibly sensitive to variations of the gap spacings and lead to detuning of resonances that are a significant fraction of the resonance linewidth. Here, these gap spacings are controlled to the molecular level by chemical crosslinks.

Figure 4.1b depicts close-packed metamolecules as assembled according to the methods section before methylamine treatment. The diameter of pillars ranges from 40-100 nm diameter with 50 nm height pillar arrays and 1 μm pitch. One may observe that the nanoparticle deposition is strongly affected by the EHD flow producing Au nanopillars. These assemblies are significantly larger than one ring around a pillar due to continued entrainment of nanoparticles into EHD flows and anhydride crosslinking of nanoparticles onto existing nanoparticles. The primarily two-dimensional character of these assemblies is due to the relative strength of the EHD flow fields which are lateral. Nanoparticles observed in the plane above the surface can be observed to be directly linked to pillars and are likely caused by brownian motion or electrophoresis. Assemblies such as these contain magnetic resonances, but are clearly not uniform enough for plasmonic metasurface applications; further treatment of the substrate is necessary to achieve large area functional magnetic resonator surfaces.

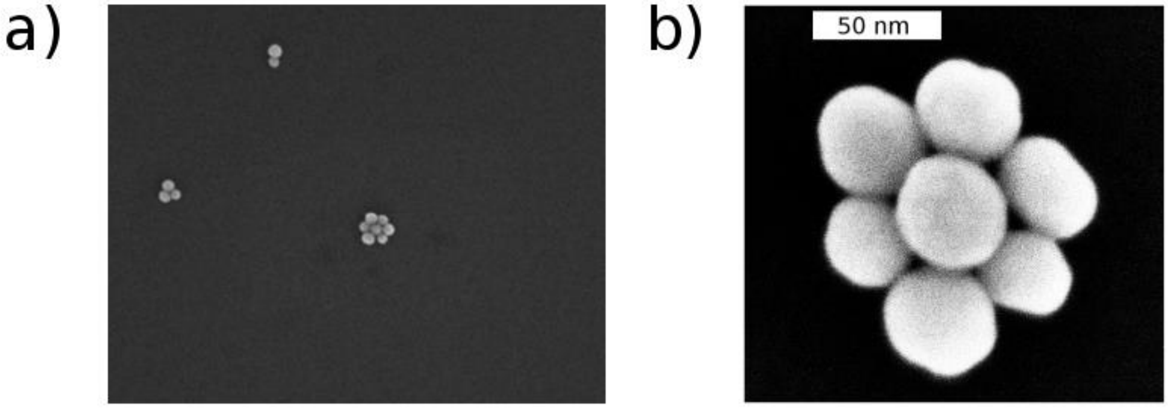

Figure 4.2: a) SEM micrograph of a close-packed metamolecule after methylamine treatment. b) Large magnification image of a close-packed metamolecule.

The dual pathway carbodiimide crosslinking enables the selective removal of nanoparticles that are not bound to pillars. Over-assembled metamolecules occur on the surface due to the continued entrainment of nanoparticles into EHD flow fields generated by the metamolecules and further anhydride crosslinking. Fortunately, the anhydride group linking nanoparticles may be attacked via nucleophilic substitution and cleaved, unlinking the nanoparticles, while the amide bond between particles and pillars is robust to nucleophiles. Figure 4.2a depicts metamolecules after an assembled substrate is sonicated in 5% methylamine solution for 10 minutes. The process leaves pillar defined metamolecules significantly increasing the uniformity of the assemblies. Due to the close-packed assembly provided by EHD flow, proper design of the center pillar ensures uniform numbers of nanoparticles in a metamolecule, the gap spacing control – which may be observed in Figure 4.2b – further ensures consistent spectral position of magnetic resonances exhibited by the metamolecules.

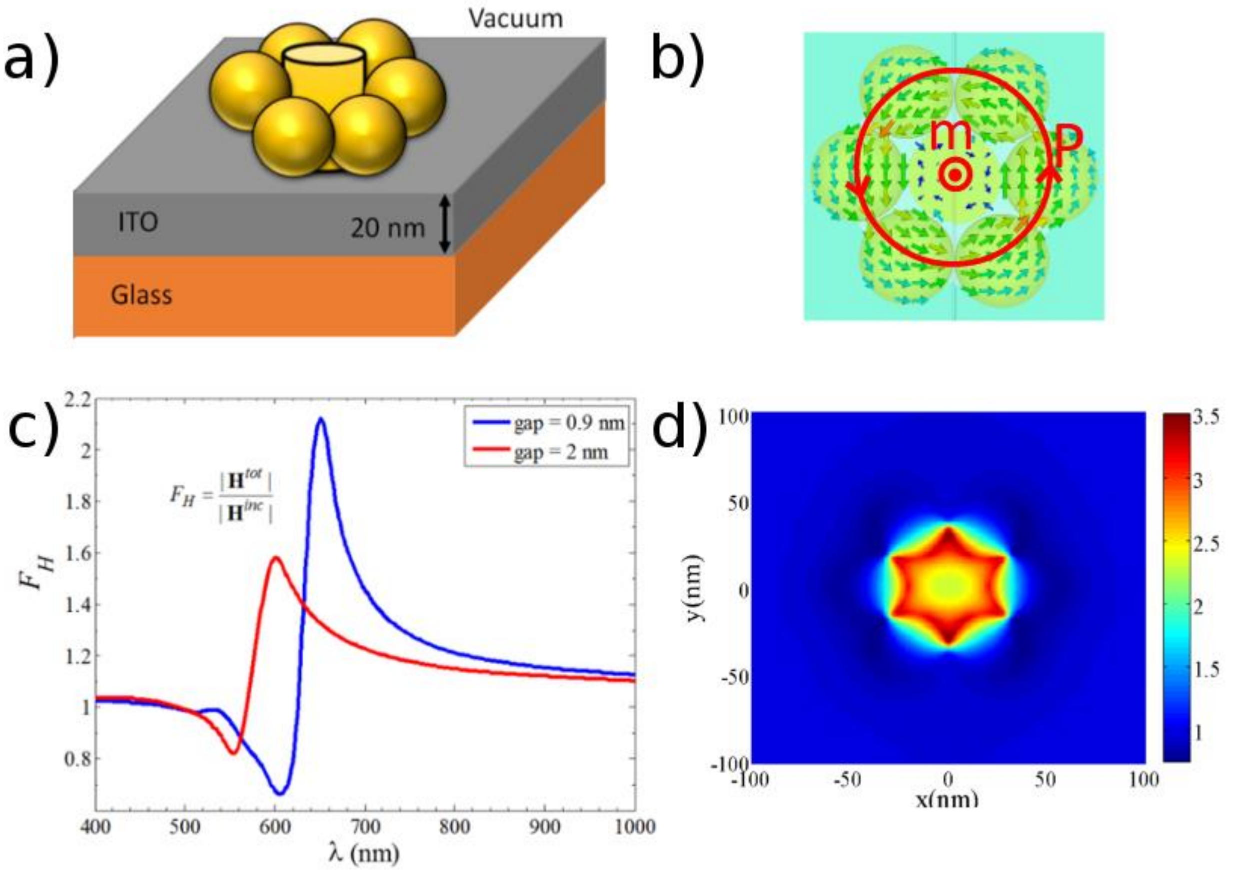

Figure 4.3: a) Schematic depicting nanoparticles assembled around a Au pillar defined with electron beam lithography on an ITO coated glass substrate. b) Full wave simulation plotting the current density distribution of the structure with a schematical representation of the polarization and magnetization produced by the close-packed metamolecule structure. c) Plot of the magnetic field enhancement F_H in the gap between two nanoparticles as a function of wavelength at two gap spacings. The small gap spacings that result from anhydride crosslinking are shown to significantly increase F_H. Both structures are shown to have magnetic resonances in the visible regime d) Full wave simulation depicting the magnetic field enhancement of the structure in Figure 4.2b at a crossectional slice taken from the halfway point of the pillar height at the structure’s resonance wavelength.

Having observed close-packed metamolecules, we turn now to consider their properties as magnetic resonators. Reflections from Si disrupt magnetic resonances so Figure 4.3a shows six gold nanoparticles with diameter of 40 nm are placed around a pillar with diameter 40 nm and height of 50 nm on ITO coated glass, this substrate is compatible with the assembly method outlined above. With excitation from two oblique plane waves at +/- 45^0 we observe, in Figure 4.3b, circular currents that give rise to an orthogonal magnetism. This effect is a result of the six fold rotational symmetry where electric dipoles in individual Au nanoparticles hybridize with their neighbors to generate current distributions with the continuous symmetry of the metamolecule.

The magnetic field enhancement is defined as the ratio of the magnitude of the total magnetic field in presence of nanoparticles and pillar, to the magnitude of the incident magnetic field in their absences at the same location. Figure 4.3c shows the magnetic field enhancement versus wavelength in visible and NIR region, for the structure shown in Figure 4.2b. Two gap spacings are analyzed to demonstrate the efficacy of using molecular linkers to build magnetic resonators. While artificial magnetism is an inductive effect and small gap spacings are typically used as capacitors in the context of metamolecules, we still observe a large increase from ~1.6 to ~2.1 of the magnetic field enhancement in the gaps between nanoparticles. This effect is due to the field confinement. Smaller gap spacings yield more strongly hybridized modes and thus lower energy, larger wavelength resonances. With the mode volume reduced somewhat with the smaller gap spacings and the wavelength made larger, the field confines more than compensates for the increased capacitance leading to larger magnetic field enhancements. In either case, one can observe a Fano lineshape of the response due to interference between the narrow magnetic resonance and the broad electric resonance. The narrow Fano-type magnetic resonance demonstrates the necessity of fine control over gap spacings between nanoparticles and the pillar, the small deviation from 0.9 nm to 2 nm leads to low overlap between the modes which would reduce device performance if the two metamolecules coexisted on the same surface.

Magnetic field enhancement profile, in the transverse cross section of the structure for 0.9 nm gap, has been shown in Figure 4.3d. As it is clear from the figure, in the all gap spacings between nanoparticles and the pillar, magnetic field enhancement is strong. Unlike metamolecules for enhancing electric fields, we see that the magnetic field enhancement occurs over a large fraction of the metamolecule area. This enables sensors based on this technique to probe a significantly greater number of molecules than electric field enhancing resonators which may help compensate for the somewhat smaller field enhancements achieved by magnetic resonators.

4.3 Conclusion

In this chapter, we have demonstrated the importance of combining short range and long range driving forces in self-assembly of Au nanospheres from colloid. Nanoantenna surfaces composed of oligomers are formed using local chemical reactions and EHD flow that drives nanospheres together for the formation and retention of anhydride bridges. The anhydride linker is observed with surface enhanced Raman scattering spectroscopy and also shown to be cleavable via nucleophilic substitution. The process by which anhydride linkers are selectively promoted between particles on a substrate surface is elucidated via molecular dynamics simulations that demonstrate long residence times at small nanoparticle gap distances are necessary. The EHD flow – anhydride crosslinking directed assembly method is then used on Au nanopillar arrays to control the deposition of close-packed plasmonic metamolecules. The magnetic resonance of an individual metamolecule is investigated and shown via full wave electromagnetic simulations to enhance the local electric field by a factor of 3.5. These results represent a step towards sensing of molecules with optical magnetic moments over a large area.

4.4 Methods

4.4.1 Chemically Functionalized Template Fabrication

Arrays of Au nanopillars on silicon are prepared using electron beam lithography (EBL), depicted in Figure 4.4. First, a highly doped Si wafer (0.004 ohm-cm) is cleaned via a standard cleaning protocol: 1) Piranha acid treatment for 15 minutes 2) Sonication in Acetone for 5 minutes 3) Sonication in IPA for 5 minutes 4) Nitrogen drying 5) Vacuum baking at 200 ^{0}C for 5 minutes. 45 nm of Au with a 5 nm Cr adhesion layer is electron-beam evaporated (Angstrom) onto the clean Si wafer. The Au on Si wafer is diced and cleaned again. Ma-N 2400 negative tone photoresist (Microchem) is spin-coated onto the clean Au on Si substrates and baked at 90^{0}C for 1 minute. Pillars are defined by exposing the photoresist with an electron beam (FEI) and a 1 minute development in Ma-D 525 (Microchem). The Ma-N pillars are used as a mask for ion milling (IntIvac), where the samples are etched until reaching the Cr layer, leaving Ma-N on Au pillars. The Cr layer is etched with Cr etchant (Sigma Aldrich) to avoid any overetching of the pillars with ion milling. The Ma-N is removed via piranha cleaning and samples are rinsed with deionized water and dried with nitrogen. Next two self-assembled monolayers are formed on the substrate. The Au pillars are selectively functionalized with Amine terminated ligands by overnight treatment of 0.5 mMol 11-Amino-alkane-thiol (Sigma Aldrich) in ethanol overnight. Samples are vigorously rinsed in methanol and dried with nitrogen. The Si wafer is then selectively functionalized via a vapor phase with trimethoxy(octyl)silane (Sigma Aldrich) by placing a wafer in a desiccator with an open vial of trimethoxy(octyl)silane under low vacuum for three hours. The substrates are then sequentially rinsed with toluene, acetone, and isopropyl alcohol and dried with nitrogen.

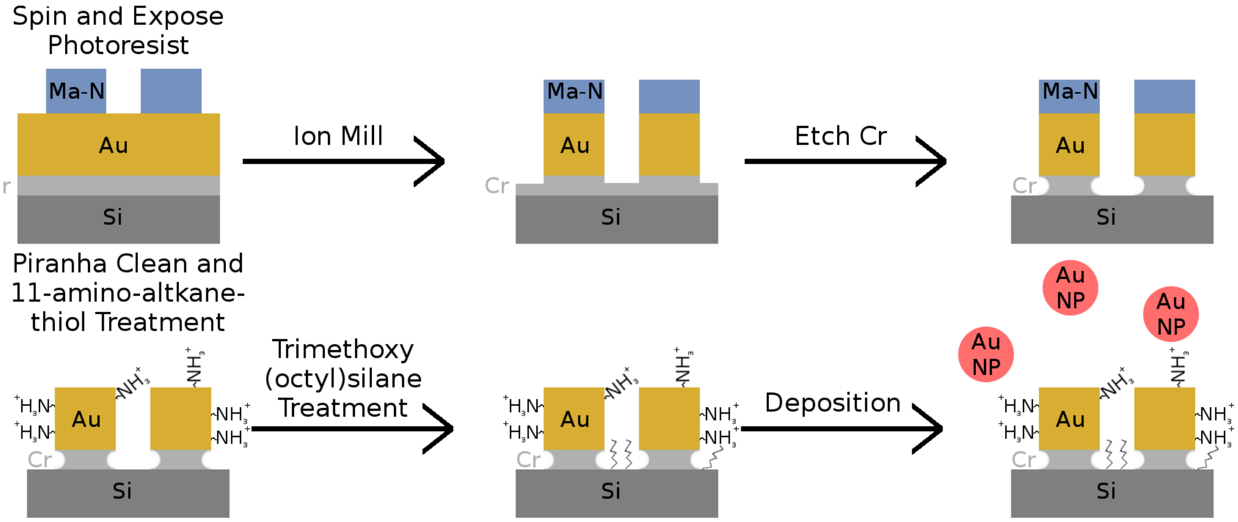

Figure 4.4: Schematic of the fabrication process for seeded growth of close-packed metamolecules

4.4.2 EHD Flow Assisted Chemical Assembly of Nanoparticles onto Pillar Templates

Au nanoparticles with 40 nm diameter in colloid are synthesized using a seeded growth method[275] at a concentration of 0.2 mg/ml. Nanoparticles are functionalized with lipoic acid (Sigma Aldrich) by replacing the citrate solution with pH 11 water and adding 100 µM of lipoic acid in ethanol. After overnight treatment the nanoparticle solution is then washed with DI water and the concentration is maintained at 0.2 mg/mL. 3 mL of this solution is added to a clean 10 mL beaker and 60 µL of 20 mMol s-NHS (Sigma Aldrich) in DI water with 100 mMol MES buffer (Sigma Aldrich) is added to this solution and swirled. Subsequently 60 µL of 8 mMol EDC (Sigma Aldrich) in DI water with 100 mMol MES buffer is added and swirled.

The functionalized Au pillars on Si substrate, as the working electrode, is then placed in the nanoparticle solution 0.5 mm apart from a Pt mesh counterelectrode with 1.2 V bias across them for 1 hour at room temperature. The bias is then removed and the substrate is moved to the bottom of the beaker where subsequent dilutions of the nanoparticle solution are performed with 10 mMol NaCl solution until the solution is diluted 1000 fold. This is performed to minimize aggregation of nanoparticles on the surface due to drying effects. The substrate is then rinsed thoroughly with isopropyl alcohol as it is removed from solution and then quickly dried with nitrogen. Excess nanoparticles are then selectively removed by submerging the substrate in 5% methylamine solution and sonicated for 10 minutes.

4.4.3 Characterization

Scanning electron microscopy images are obtained with a Magellan (FEI) SEM. The optical responses of close-packed oligomers are understood with full wave electromagnetic simulations (CST).

Part Three: Machine Learning Enabled Biosensing with Surface Enhanced Raman Scattering Spectroscopy

“Nothing contributes so much to tranquillize the mind as a steady purpose- a point on which the soul can focus its intellectual eye”

-Mary Shelley, Frankenstein, or the Modern Prometheus

Chapter 5

Longitudinal Monitoring of Biofilm Formation via Robust Surface-Enhanced Raman Scattering Quantification of *Pseudomonas aeruginosa*-Produced Metabolites

5.1 Introduction

During biofilm formation, differential gene expression is regulated through a cell density-dependent mechanism called quorum sensing (QS).[43,276] Soon after surface attachment, bacteria begin producing extracellular polymeric substances (EPS) and QS signaling molecules.[277] Once formed, a combination of physical mechanisms and genetic and metabolic adaptations within biofilms impart extreme antibiotic tolerance or resistance to constituent cells,[278] which can withstand up to 1000 times higher doses of antibiotics than their free floating planktonic counterparts.[279] While new antimicrobial strategies are being developed to combat antibiotic resistance, here we investigate a promising parallel strategy, sensing bacterial metabolites associated with QS for early detection of *Pseudomonas aeruginosa* biofilm formation at a stage where antibiotic treatment has higher efficacy. *P. aeruginosa* is a biofilm-forming, opportunistic pathogen that is associated with contamination of medical devices, respiratory infections in immunocompromised patients,[280,281] and is one of the most common bacteria isolated in chronic wounds.[282] Among the many virulence factors and QS compounds that *P. aeruginosa* produces is pyocyanin,[283] a redox-active secondary metabolite which can act as a terminal signaling factor in the QS process.

Interestingly, recent Raman spectroscopy-based studies have demonstrated the ability to distinguish between species of bacteria,[284,285] identify molecules enriched in particular species, and even detect specific local metabolic activity in environmental samples using stable isotope probing techniques.[286] Raman spectra provides “molecular fingerprints” composed of the vibrational spectrum of molecules serving as a label free detection method. Biologically relevant concentrations of metabolites often range from nM to mM levels in clinical samples,[287,288] and have enormous potential for enabling personalized medicine. The limit of detection and quantification range of metabolite concentration needed at physiologically significant levels in biological samples, though this may vary with molecules and environments,[289] is often in the μM range,[290] and thus requires large and reproducible enhancements of Raman signals. Surface enhanced Raman scattering (SERS), when employing plasmonic nanogaps, is capable of providing necessary enhancements to achieve detection at biologically relevant concentrations. SERS has been used to distinguish between colony biofilms of bacteria species during maturation phases[291] and even detect pyocyanin in spiked saliva down to 2.1 $\mu g \cdot mL^{-1}$ (10 μM)[292] and spiked in subcutaneous implants in mice down to 0.1 μM. Yet challenges still exist in incorporating nanogaps with large and uniform enhancement factors in device architectures with nM detection limits and quantification spanning the range of metabolite concentrations found in biological samples.

SERS enhancement factors due to plasmonic nanogaps are highly dependent on the distance between plasmonic nanoantennas, increasing monotonically with decreasing gap size. Statistical analysis of various size controlled nanogaps using DNA-tethering observed single molecule SERS intensity when nanogaps are on the order of 0.5 – 0.9 nm.[207,293] At nanogap distances below approximately 0.5 nm, depolarization effects attributed to quantum tunneling reduce enhancements.[294] Reaching sub-nanometer nanogap dimensions over large area without large variations is difficult and thus, SERS substrates often exhibit tradeoffs between reproducibility and large enhancement factors. Yet it is necessary to have both uniform and large enhancement factors across SERS substrates to reproducibly achieve low detection limits in quantitative sensing applications. Consider that at extremely low concentrations, analyte molecules will not be uniformly distributed across the surface.[295,296] SERS measurements on a mixture of two different analytes determined that single molecule sensing events occur at nM concentration as not every molecule in the scattering volume will reside in a hotspot.[297,298] Inherent variances in SERS substrates' enhancement factors only worsen any analyte's location-dependent signal variations and will lead to large SERS intensity fluctuations. While such surfaces may produce a low limit of detection due to the presence of a few hotspots on the surface, the limit of quantification will be much higher. Thus increasing the uniformity and density of nanogaps with high enhancements will lead to a higher probability that an analyte will adsorb on a hotspot in the illuminating laser spot size and thereby reproducibly contribute to the SERS signal at low concentrations.

Here, large scale uniformity in hot spot intensity from nanoantennas on SERS substrates is achieved using a hierarchical self-assembly method for fabricating Au nanospheres into close-packed structures, referred to as oligomers, on a diblock copolymer template. Assembly combines long range, electrohydrodynamic (EHD) flow,[269,270,299] and short range, chemical crosslinking, driving forces to provide a high density of oligomers with sub-nanometer nanogap spacing. Induced by an electric field in colloidal solution, EHD flow promotes lateral motion and close-packing of nanospheres in colloid at the electrode surfaces, assembling transient close-packed oligomers. Transient oligomers are frozen using EDC, a carbodiimide crosslinker, to form an anhydride bridge between the carboxylic acid ligands on Au nanospheres. We have previously reported that this process achieves a high yield of close-packed oligomers composed of ten nanospheres or less with gap spacing of 0.9 nm, having an electric field enhancement on the order of 600 in the hotspot region.[247] Close-packing is also advantageous as it allows for variability of oligomer orientation with respect to polarization of incident light. In a close packed oligomer, the plasmon resonance will most closely reflect that of a linear oligomer along the polarization axis.[272]

When using benzenethiol as the analyte, here we report that SERS signal exhibits a relative standard deviation of 10% across a 100 μm × 100 μm area with over 1000 different measurements. This achievable uniform SERS response allows for incorporation in microfluidic device architectures and spectral data analysis using multivariate machine learning algorithms.[177,300,301] A large number of training data sets and the full spectra collected from SERS substrates are necessary for accurate quantitative analysis by accounting for signal variance inherent to SERS. By combining uniform SERS substrates having high signal enhancements with multivariate statistical analysis of SERS spectra, we are able to differentiate pyocyanin in the complex soup of biological media at concentrations down to 1 $ng \cdot mL^{-1}$ (4.8 nM) and robustly quantify concentrations spanning five-orders of magnitude. Moreover rapid SERS analysis in solutions negates the need to grow biofilm directly on SERS substrates or to drop cast plasmonic clusters on static biofilms, thus allowing instantaneous quantification of pyocyanin for detection of biofilm formation on any surface. Longitudinal monitoring of supernatant from bacterial cultures in microfluidic devices exhibits SERS signal from pyocyanin as early as 3 h after *P. aeruginosa* culture inoculation. This is before surface-attached bacteria exhibits a decreased susceptibility to bactericidal antibiotic carbenicillin that was measured at 10 h. Low cost fabrication of nanoantennas with nanometer scale nanogaps over large areas enabled by self-assembly and statistical spectral analysis of SERS data demonstrate the capacity for fabricating device architectures capable of early detection of biofilms enabling effective antimicrobial treatment.

5.2 Large-Area Uniformity of SERS Substrates

Directing assembly of Au oligomers with uniformly distributed electromagnetic hotspots over large areas is achieved by utilizing electrohydrodynamic (EHD) flow to drive chemical crosslinking to form nanogaps with sub-nanometer gap spacing as reported previously.[272] Chemical assembly is facilitated using a two-step growth process: (1) electrophoretic sedimentation driving chemical crosslinking of Au nanoparticle seeds onto a working electrode and (2) growing oligomers via EHD flow, and stabilizing via chemical crosslinking. During electrophoresis, depicted in Figure 5.1a, lipoic acid functionalized Au nanoparticle monomer seeds are chemically assembled on amine functionalized PMMA regions of a diblock copolymer poly(styrene-b-methyl methacrylate) (PS-b-PMMA) coated working electrode using 1-Ethyl-3-(3-(dimethylamino) propyl)carbodiimide/N-hydroxysulfosuccinimide (EDC/sNHS) carbodiimide chemistry (inset). In the second deposition step, an applied field deforms the ionic double layer around the Au seeds, shown on the right of Figure 5.1a, inducing an osmotic flow towards the seeds' equator, termed EHD flow.[191,232] These flow fields entrain nearby nanospheres and drives them toward Au seeds to form transient close packed oligomers, which are subsequently stabilized through an acylisourea – carboxylic acid reaction that forms anhydride bridges between nanospheres, shown in Figure 5.1b. These bridges result in oligomers with uniform interparticle gap spacings of approximately 0.9 nm, corresponding to the sulfur-sulfur distance of an anhydride bridge calculated from atomistic simulations.[272]

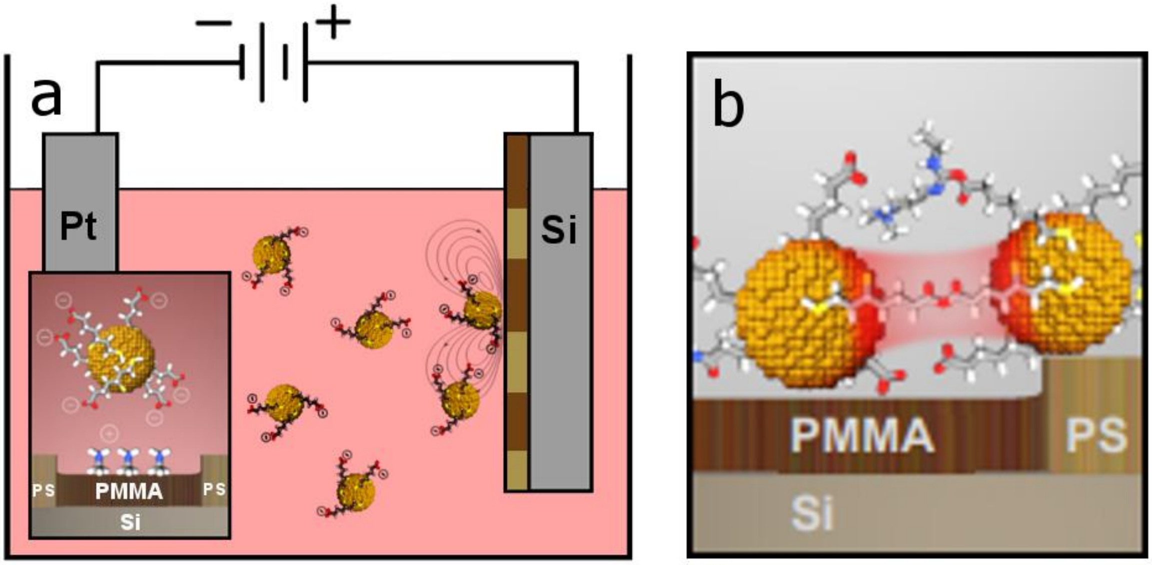

Figure 5.1. Schematic overview of chemical assembly of SERS substrates. (a) Lipoic acid-functionalized Au nanoparticles are electrophoretically driven toward a working electrode. Au seeds chemically crosslink with amine functionalized PMMA regions on self-organized diblock copolymer PS-b-PMMA template (inset). The resultant EHD flow field around the seed entrains nearby nanoparticles forming transient oligomers. (b) Illustration of anhydride bridge that forms between carboxylic acid and acylisourea on neighboring nanoparticles that stabilizes oligomers when the external field is turned off.

A scanning electron microscopy (SEM) image of a SERS substrate composed of self-assembled oligomers is displayed in Figure 5.2a. The size distribution in percent area (Figure 5.2b inset), determined from image analysis of SEM images using Wolfram Mathematica^{TM} indicates that trimers are observed with the highest probability, with quadrumers being the second. The absorption cross section of a dimer, linear trimer, and linear quadrumer with a nanogap size of 0.9 nm – corresponding to the calculated sulfur-sulfur distance of an anhydride bridge – was calculated from full-wave simulations and shows good agreement with the fine structure observed in UV-visible measurement as shown in Figure 5.2b. This agreement suggests that measured attenuation is dominated by frequencies near the trimer resonance with contributions from frequencies associated with dimers and quadrumers, blue and red shifted, respectively. Larger oligomers observed in the SEM image of Figure 5.2a contribute to the broad shoulder at higher wavelength.

This observation can be understood by considering the effect of the close packed arrangement of nanospheres in oligomers on their plasmon resonance. It has been previously shown using dark field microscopy and full-wave simulations that an oligomer's plasmon mode is only slightly perturbed by the addition of a nanosphere when the added nanosphere is unaligned with the incident beam's polarization. For example, a hexamer made up of two trimers on top of each other has a plasmon mode that is slightly redshifted from that of a linear trimer.[272] Consider that for a close packed oligomer to have more than 4 nanospheres in a row, it must be composed of more than 21 nanospheres. From the oligomer size distribution, Figure 5.2b inset, it was determined that, by number, 98% (88% by area) of oligomers are composed of 21 nanospheres or less. It is then unsurprising that the resonance is dominated by dimers, trimers and quadrumers. Full-wave simulations of the electric field enhancement – $|E_{olig}|/|E_0|$, where $E_{olig\ and}$ E_0 are the plane wave field with and without oligomers, respectively – was performed for nanogaps of 0.9 nm and 2 nm in dimers and linear trimers and are shown in Figure 5.2c. As expected the resonance red shifts as the gap distance decreases. Furthermore, increasing gap spacing from 0.9 to 2 nm reduces the calculated field enhancement from 621 to 240 (762 to 314) in a dimer (linear trimer). The SERS enhancement can be estimated as the fourth power of the field enhancement, $|E_{olig}|/|E_0|^4$, and thus even slight variations in field enhancement will lead to large variations in SERS intensity. Overall the plasmon resonance is less affected by close packed oligomer size than the resultant field enhancement in the nanogaps when spacing decreases below 2 nm. Thus chemically controlled gap spacing enables uniform SERS intensity, observed in Figure 5.2d.

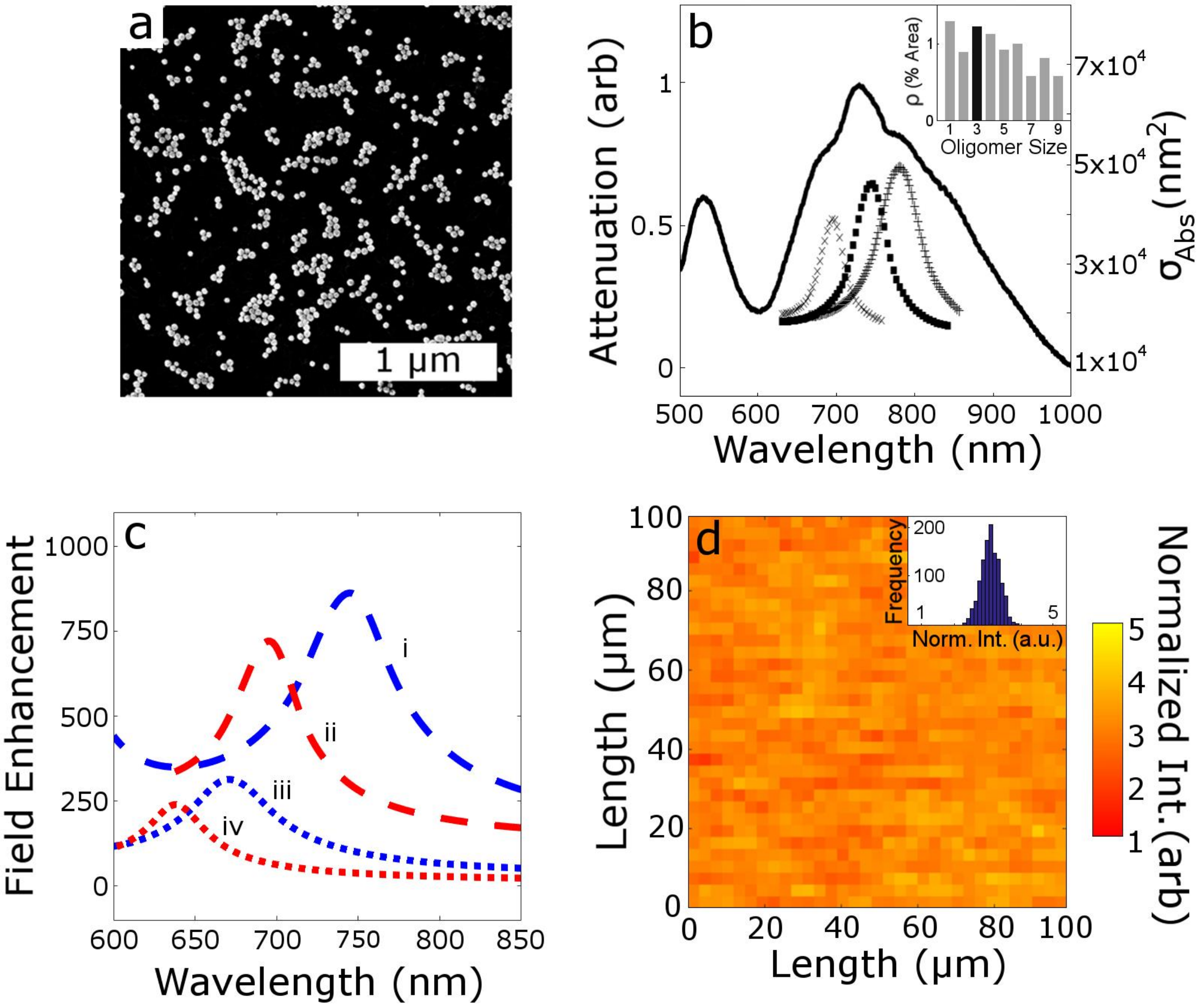

Figure 5.2. (a) SEM micrograph of self-assembled Au oligomers. (b) Attenuation curve of SERS substrates in DI water. Dotted curves represent absorption cross section of a dimer (×), linear trimer (squares), and linear quadrumer (+) from full-wave simulations. Inset shows oligomers distributions in percent area on SERS substrate calculated from SEM images acquired over 10 μm^2. (c) Calculated field enhancement of Au (i) linear trimer with 0.9 nm gap, (ii) dimer with 0.9 nm gap, (iii) linear trimer with 2 nm gap, and (iv) dimer with 2 nm gap from full-wave simulations. Curve (i) and (ii) are offset by 100 for clarity. (d) Normalized SERS intensity map of benzenethiol's 1573 cm^{-1} vibrational band. Inset shows distribution of normalized intensity with a RSD of 10.4%.

Figure 5.2d displays the normalized SERS intensity of a benzenethiol vibration band, 1573 cm^{-1}, acquired over a 100 μm × 100 μm area. The SERS intensity has a relative standard deviation (RSD) of 10.4%. While nanogap spacing of 0.9 nm is important to achieve large and uniform SERS enhancements, at the same time, the assembly method provides a sufficiently broad band response, ranging over a window of 625 nm to 875 nm in the attenuation measurement, to enhance signals at both the plasmon excitation wavelength and Raman scattered wavelength. The uniform SERS response with large enhancements over large area enables the use of these SERS substrates in device architectures. Of further significance, it also enables the acquisition of large datasets needed for statistical analysis enabling quantitative detection.

5.3 Quantification and Detection of Pyocyanin in Aqueous Media

Pyocyanin exhibits a broad absorption band from 550 nm to 900 nm,[302] thus using a 785 nm laser to excite nanoantennas results in surface enhanced resonance Raman scattering (SERRS). SERRS spectra of aqueous pyocyanin from 1 $\mu g \cdot mL^{-1}$ (4.8 μM) to 100 $\mu g \cdot mL^{-1}$ (480 μM) (Figure 5.3a) displays clear Raman bands similar to pyocyanin spectra reported using surfaces with Ag and Au nanorods[293,303] at 552 cm^{-1}, 1353 cm^{-1}, 1602 cm^{-1}, and 1620 cm^{-1}. These bands rise at a concentration as low as of 100 $pg \cdot mL^{-1}$ (480 pM), as exemplified in Figure 5.3b for the band 552 cm^{-1}. The log-log dose-response curve of pyocyanin at 552 cm^{-1} (Figure 5.3c) reveals a linear regime between 1 $ng \cdot mL^{-1}$ and 10 $\mu g \cdot mL^{-1}$, consistent with Langmuir adsorption kinetics.

Within this linear regime, pyocyanin concentration can be quantified using the formula shown in equation (1) with $R^2 = 0.951$.

$$\log C = a \log I + b \qquad (5.1)$$

Here a and b are fitting constants with value of 3.623 and 2.924, respectively, while C and I represent pyocyanin concentration and normalized SERRS intensity at 552 cm^{-1}, respectively. For concentrations below the limit of quantification (LOQ), 1 $ng \cdot mL^{-1}$, SERS substrates detection of pyocyanin was determined by comparing signal and background at 552 cm^{-1} per guideline EP17 of the Clinical and Laboratory Standards Institute.[304] The limit of blank (LOB) is calculated by adding the mean background signal (DI water) $\bar{I}_{bg}$ at 552 cm^{-1} to 1.645 its standard deviation σ_{bg} , shown in equation (2):

$$\text{LOB} = \bar{I}_{bg} + 1.645\sigma_{bg} \qquad (5.2)$$

Detection is defined in the standard manner where the mean signal, $\bar{I}_D$, at 552 cm^{-1} is at least 1.645 standard deviations, σ_D, larger than or equal to the LOB, shown in equation (3):

$$\bar{I}_D - 1.645\sigma_D \geq \text{LOB} \qquad (5.3)$$

From this analysis, SERS substrates exhibit detection of pyocyanin in aqueous media at concentration of 100 pg·mL^{-1}, above which one can observe pyocyanin signals above the background in Figure 5.3c. We did not measure pyocyanin at lower concentration since our interest was in establishing a LOD in biological media. These results are discussed in the following section.

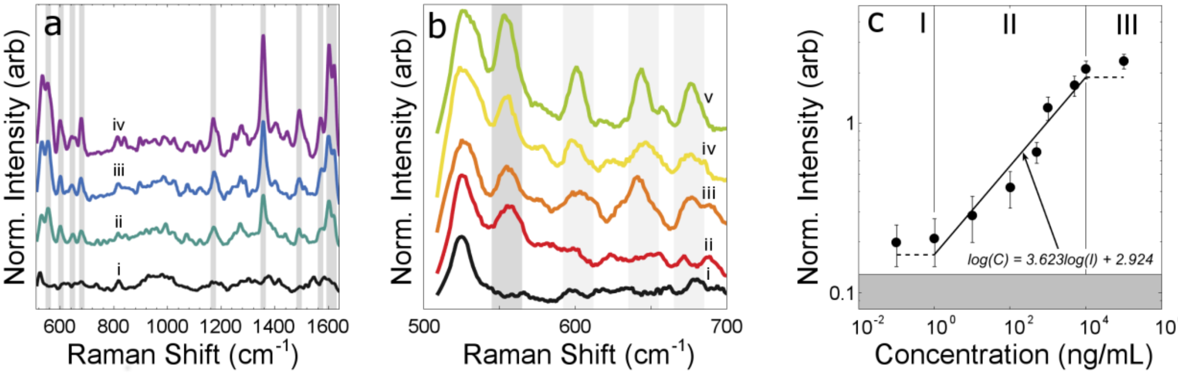

Figure 5.3. (a) SERRS spectra of (i) DI water, (ii) 1 µg·mL^{-1}, (iii) 10 µg·mL^{-1}, and (iv) 100 µg·mL^{-1} pyocyanin in water. Grey bars indicate Raman bands of pyocyanin. (b) SERRS spectra of of (i) DI water, (ii) 100 pg·mL^{-1}, (iii) 1 ng·mL^{-1}, (iv) 10 ng·mL^{-1}, and (v) 100 ng·mL^{-1} pyocyanin in water. Dark grey bar highlights pyocyanin's 552 cm^{-1} band while light grey bars indicate other pyocyanin Raman bands. (c) Dose-response relationship of pyocyanin and normalized SERS intensity at 552 cm^{-1}. Error bars depict standard deviation of 100 measurements.

5.4 Training Data Acquisition and Building Multivariate Predictive Model

Using linear regression on a single Raman band allows for facile comparison with other SERS surfaces in the literature as it is widely used for quantitative calibration.[171,172,305–307] While this method is sufficient for analysis of a pure analyte in water, it discards the remaining rich information in each spectrum and can lose sensitivity when other molecules are present in solution as is the case when monitoring biofilm growth. Alternatively multivariate analysis, specifically partial least square (PLS) regression, analyzes the full spectra collected from SERS substrates and improves limit of quantifications in complex media.[308] Thus we capitalize on the uniform SERS response of our surfaces to acquire necessary training data and employ PLS analysis to quantify pyocyanin in more complex biological media, an important requirement for diagnostic applications. Training data sets, composed of SERRS spectra from known concentrations of pyocyanin spiked in LB media, generate a robust PLS predictive model for pyocyanin concentration in the increased background noise of the biological media. The multivariate model predictive capability is subsequently demonstrated by quantifying pyocyanin production from *P. aeruginosa* planktonic cultures during in vitro growth.

First we demonstrate the suitability of our substrates for detecting pyocyanin in complex media by comparing SERRS spectra of blank LB media, LB media spiked with 10 $\mu g \cdot mL^{-1}$ of pyocyanin alongside spectra from of cell-free supernatants collected from mid-stationary phase cultures of wild type *P. aeruginosa* PA14 and its phenazine-deficient mutant strain Δphz1/2 as a control (Figure 5.4a). The Δphz1/2 strain does not produce phenazines, hence its SERRS spectrum (multiplied a factor of 5 for clarity) indeed lacks the vibrational fingerprint of pyocyanin; it is similar to the spectrum obtained for blank LB media. Meanwhile, SERRS spectra of wild type PA14 and pyocyanin in LB broth exhibit similar features, including distinct pyocyanin vibrational bands, and thus confirming the suitability of our substrates for the task. While pyocyanin signals are clearly seen here, signal interference from other molecules is amplified at low concentration, requiring more the more sophisticated analysis.

The calibration datasets that were acquired by collecting SERRS spectra of LB media spiked with 100 $pg \cdot mL^{-1}$ to 100 $\mu g \cdot mL^{-1}$ pyocyanin generated a predictive model of pyocyanin concentration in LB media using PLS regression. From the 400 spectra collected for each pyocyanin dose, 380 were randomly selected as training set and the remaining 20 withheld as testing set. To optimize the model, over- and under-fitting are avoided by using 10 PLS components where a minimum in RMSE of cross validation (RMSECV) is observed, displayed in inset of Figure 5.4b. With the testing set, the model demonstrates accurate prediction between 1 ng·mL^{-1} and 100 µg·mL^{-1} as shown in Figure 5.4b. Fitting the predicted versus actual concentration with a line having a slope of 1 – representing perfect predictive capability – gives a R^2 value of 0.956.

5.5 Pyocyanin Quantitative Detection in Complex Media

The generated predictive model was used to quantify pyocyanin from SERRS spectra of cell-free conditioned growth medium of wild type *P. aeruginosa* and Δ*phz1/2* as a function of incubation time. For each incubation time point, 400 SERRS spectra were collected within the span of less than 4 minutes, and their corresponding pyocyanin concentrations were calculated using the PLS model; the average concentration for incubation times ranging from 0 to 24 h are shown in Figure 5.4c. The performance of SERRS-PLS was compared with UV-vis absorption spectroscopy typically employed to quantify pyocyanin. Pyocyanin concentrations were calculated from UV-vis absorption peaks using the reported molar absorptivity $\varepsilon = 4.31 \times 10^3 \, mol^{-1} \, cm^{-1}$ at $\lambda_{max} = 690 \, nm$.[309]

The pyocyanin-deficient Δ*phz1/2* strain were used to calculate the instrument noise and contributions from the background media. This yielded the UV-vis LOD as 197 ng·mL^{-1} (0.94 µM), equivalent to 3 standard deviations above the background of the Δ*phz1/2* strain. This value agrees well with the reported LOD of 1 µM for UV-vis.[310] The UV-vis data is also plotted in Figure 5.4c for comparison with SERRS data. The LOD of SERRS observed in Figure 5.4b is 1 ng·mL^{-1} (4.8 nM) using the same definition for differentiating from background noise and PLS model discussed above. Thus SERS substrates are able to detect pyocyanin as early as 2 h of shaking culture growth and quantify the concentration as 2.5 ng·mL^{-1} (12 nM), as opposed to after 8 h of shaking culture growth when using UV-vis absorption. From 8 h onward, quantitative performance of the two methods is comparable. The sharp increase in pyocyanin concentration detected by SERRS from 2 to 8 h correlates with the exponential growth phase of *P. aeruginosa* in shaking culture, which is observed in Figure 5.4d as an increase in optical density and thereby accumulation of biomass. Overall this data validates SERS substrates robustness toward detection of pyocyanin in complex media.

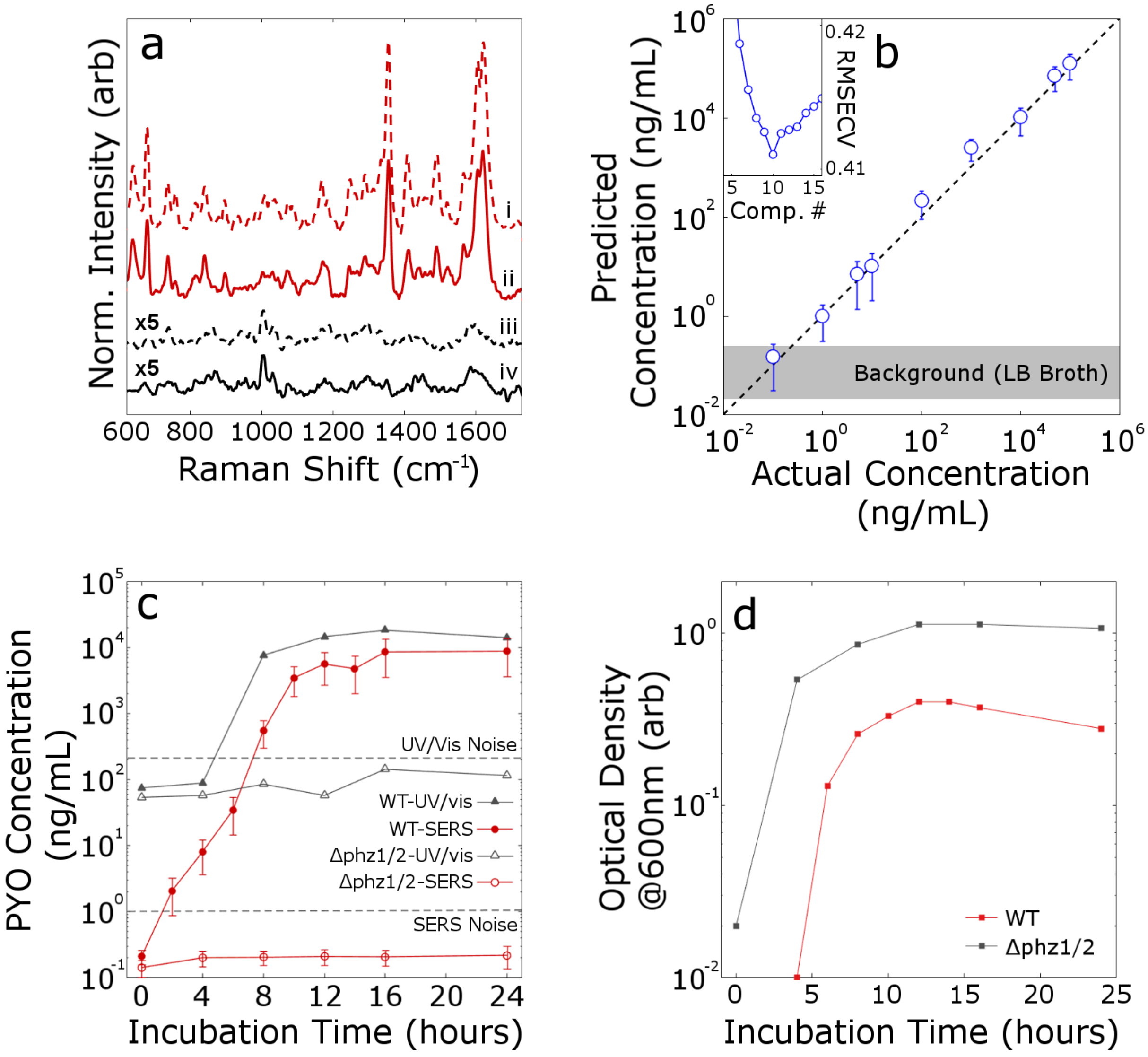

Figure 5.4. (a) SERRS fingerprint of (i) bacteria-free supernatant from mid-stationary phase cultures of wild-type *P. aeruginosa* PA14, (ii) 10 μg·mL^{-1} pyocyanin in LB broth, (iii) phenzine-null mutant strain Δphz1/2 and (iv) LB broth. Spectra in (i) and (ii) are multipled by a factor of 5. (b) Pyocyanin concentration predicted by PLS model for spectra in testing set. Error bars depict standard deviation of 20 measurements. Inset shows RMSECV of the PLS model with respect to the number of components used. (c) Pyocyanin concentration in bacteria-free supernatants of wild-type *P. aeruginosa* PA14 and its phenazine-null strain over 24 hours. Concentrations from SERS and UV-Vis measurements are calculated using trained PLS model and reported molar absorptivity $\varepsilon = 4.31 \times 10^3\ mol^{-1}\ cm^{-1}$ at $\lambda_{max} = 690\ nm$,[47] respectively. Error bars show standard deviation of 400 measurements. (d) Growth curves for planktonic cultures of *P. aeruginosa* PA14 and its phenazine-null mutant strain as determined from optical density.

5.6 Monitoring Biofilm Formation *via* Pyocyanin Quantification

Here we integrate microfluidic channels with SERS substrates to perform inline sampling of biofilm effluent, thus enabling rapid quantitative detection of pyocyanin as a means to longitudinally monitor biofilm growth. *P. aeruginosa* biofilms were grown in tryptone broth (TB) medium as previously described.[311] Effluent from the biofilm growth channel is delivered to a microfluidic channel with Au oligomers assembled on the surface as illustrated in Figure 5.5a. Every three hours, 200 SERRS spectra are collected in the effluent with a total acquisition time of 2 minutes, thus eliminating the need for performing chloroform extraction or incubating on SERS substrates[293] or evaporating solvent on SERS surfaces.[303] To quantify pyocyanin, a suitable predictive model was generated by repeating PLS analysis on training dataset acquired using the appropriate collection parameters for TB media (as oppose to LB in planktonic cultures growth) as reported in our methods. The resulting LOD from this analysis is 10 ng·mL^{-1}. Using the model, pyocyanin was detected in the SERRS signal above the LOD starting between 6 h - 9 h after inoculating as observed in Figure 5.5b. Thus the time of quantification (TOQ) falls within this range. Pyocyanin concentration after 9 h was determined to be 24 ng·mL^{-1} (115 nM), above the determined LOD. Thus, the calculated concentrations are indicative of the instantaneous production of pyocyanin in biofilm growth channel. Investigation of results generated from the model shows that some spectra at 3 and 6 h reveals Raman bands associated with pyocyanin above the background. By analyzing the individual spectra it was determined that 7.8% and 14.3% of the spectra at 3 and 6 h, respectively, show a pyocyanin concentration above the background; representative spectra are plotted in Figure 5.5c. This fraction increases to 70.3%, 96.7%, and 100% for the subsequent time points. This is consistent with non-uniform distribution of pyocyanin on the substrate surface at low concentrations. So while the averaged concentration of pyocyanin calculated for the set of Raman spectra at 3 and 6 h of biofilm growth lies below the LOD, a fraction of those spectra shows that pyocyanin is detected at those early time points. Thus 3 hours is set as the time of detection (TOD). The calculated concentrations of pyocyanin correlate with the time-dependent accumulation of biofilm biomass obtained from analysis of fluorescence images of the microfluidic growth channels (Figure 5.5d). Representative confocal fluorescence images of the growth channels show the initial stages of bacterial cell adhesion (6 h), microcolony formation (10 h), and three-dimensional growth of biofilms above the channel surface (16 h). The biomass and imaging data indicates that biofilm formation can be monitored reliably after TOQ. More significantly, it reveals that the earliest detection of biofilms, corresponding to TOD, occurs during the initial stages of bacterial cell adhesion, earlier than observed in fluorescence confocal images.

Bacterial biofilms impart antibiotic resistance and tolerance on constituent cells via several distinct mechanism.[278] These mechanisms are characteristic of bacterial phenotypes and properties of mature biofilms, so early treatment with antibiotics can be more effective than equivalent treatment of mature biofilms. The potential for early detection to provide a therapeutic improvement for infection outcomes was assessed by measuring the antibiotic susceptibility of surface-attached bacteria at different stages of biofilm growth. Bacteria exposed to a bactericidal antibiotic, carbenicillin, were differentially susceptible after 10 h of growth, after which their susceptibility decreased significantly with increasing growth time as they transition into a more protected state (Figure 5.5d inset); this transition is also captured as the sharp increase in pyocyanin concentration measured from SERRS in Figure 5.5d between 9 h and 12 h. A control biofilm was grown for 24 h and exposed to PBS instead of Carb. *n.s.* indicates no statistically significant difference between the antibiotic susceptibility of cells grown in flow channels for 6 and 10 h. Consequently, detection of a *P. aeruginosa* infection and intervention at early stages of colonization (< 10 h growth) show potential to substantially improve the effectiveness of antibiotic treatment.

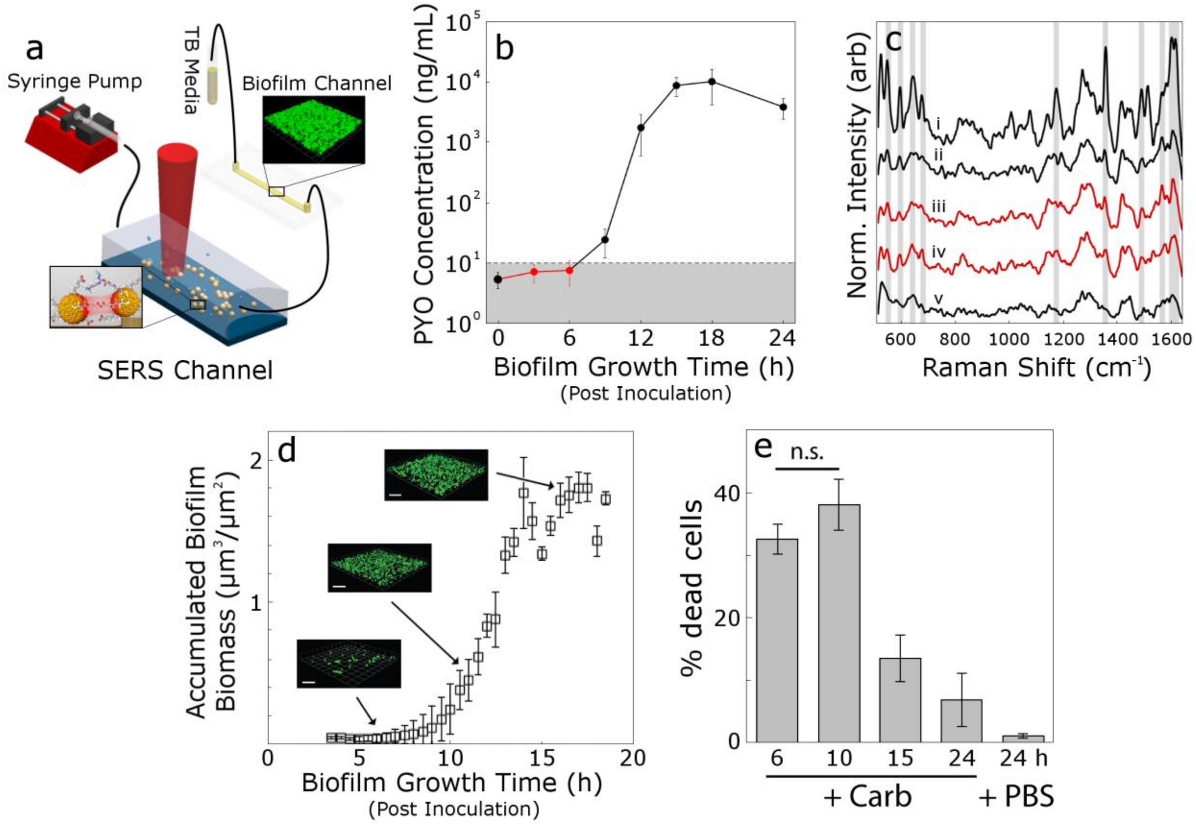

Figure 5.5. (a) Schematic of the biofilm inline measurements setup. (b) Pyocyanin concentration predicted by PLS model from SERS spectra of various time. Each data point represents the averaged of predicted concentrations and error bars show standard deviation from 200 measurements. (c) Individual representative SERS spectra at (i) 12h (ii) 9h, (iii) 6h, and (iv) 3h contrasted with spectrum of (v) TB media. Grey bars indicate pyocyanin vibrational bands. (d) Accumulated biomass of biofilm growth in the flow cell over time. Error bars depict standard deviations from n = 3 independent growth channels. Confocal fluorescence microscopy images show representative bacterial accumulation on the glass surface at the time points indicated by the arrows (scale bars are 20 μm). (e) Susceptibility of surface-associated cells to carbenicillin treatment at the specified growth times. Error bars depict standard deviation of dead cell fraction from n = 3 biological replicates.

6.7 Conclusion

Chemical assembly of SERS substrates, using electrohydrodynamic flow to initiate chemical crosslinking, yields nanogap spacings of 0.9 nm. The resultant nanogaps provide high and reproducible SERS enhancements over μm^2 areas when excited at 785 nm with power as low as 7.3 μW. This allows for integration in a device platform and the rapid acquisition of large data sets for statistical analysis. SERRS measurements yielded a limit of detection of 100 pg·mL^{-1} of pyocyanin in aqueous solution by tracking the linear relationship between the SERRS intensity at a characteristic pyocyanin molecular vibrational mode. Longitudinal studies demonstrated that pyocyanin concentration is correlated with accumulation of biomass. In more complex media, SERS substrates were able to robustly quantify analyte concentration in biologically relevant levels when using an established machine learning algorithm for spectral data analysis. Using PLS regression to analyze SERRS spectra collected from cell-filtered supernatant in a microfluidic channel, before a decreased susceptibility to bactericidal antibiotic carbenicillin is observed for surface-attached bacteria at 10 h, the ability to detect biofilm formation as early as 3 h after inoculation was demonstrated. Thus detecting microbial production of metabolites associated with quorum sensing with chemically assembled SERS substrates is a promising strategy for early intervention of bacterial infections.

5.8 Methods

5.8.1 Materials

Random copolymer Poly(styrene-co-methyl methacrylate)-α-Hydroxyl-ω-tempo moiety (PS-r-PMMA) (Mn = 7,4 kg mol^{-1}, 59.6% PS) and diblock copolymer poly(styrene-b-methyl methacrylate) (PS-b-PMMA) (Mn = 170-b^{-1}44 kg mol^{-1}) were purchased from Polymer Source, Inc. (Dorval, Canada). Gold nanospheres with diameter of 40 nm and lipoic functionalization were purchased from Nanocomposix (San Diego, CA). Si(001) wafers with resistivity of 0.004 ohm-cm were purchased from Virginia Semiconductor (Frederickburg, VA). Hydrofluoric acid (HF) was purchased from Fisher Scientific (Pittsburg, PA). 2-(N-morpholino)ethanesulfonic acid (MES) 0.1M buffer was purchased from Pierce (Rockford, IL). 1-ethyl-3-[3-dimethylaminopropyl] carbodiimide hydrochloride, N-hydroxy sulfosccinimide (s-NHS), dimethyl sulfoxide (DMSO), ethylenediamine (ED), toluene, ethanol, isopropanol (IPA), potassium carbonate, and 52-mesh Pt gauze foil were purchased from Sigma Aldrich (St. Louis, MO). Nanopure deionized (DI) water at 18.2Mohm cm^{-1} was obtained from Milli-Q Millipore System.

5.8.2 Self-Assembly of SERS Substrates

Si wafers were cleaned in 10% HF to remove the native oxide. *The potential of HF to cause severe injury mandates extreme caution during usage.* Random copolymer PS-r-PMMA and diblock copolymer PS-b-PMMA solution in toluene (1 wt%) were spun coat on the flat Si surface and annealed at 198^{o}C to form thin films as described in previous work.[137] PMMA regions are selectively functionalized with amine end groups by first immersing in DMSO and then in ED/DMSO solution (5% v/v) for 5 minutes each step and without rinsing in between. The substrate is then washed with IPA for 1 minute and dried under nitrogen.

Lipoic acid functionalized 40 nm Au nanospheres in aqueous solution is concentrated twofold by adjusting the pH to 8 with potassium carbonate and centrifuging for 25 minutes at 1700 RCF and redispersed in DI water. 3 mL of concentrated Au nanosphere solution is added to a 10 mL beaker, followed by 35 uL of freshly prepared 20 mM s-NHS in 0.1 M MES buffer and 35 uL of freshly prepared 8 mM EDC in 0.1 M MES buffer. The solution is swirled and moved on a hotplate to be heated to 8o°C. A 1 cm x 1 cm Pt mesh and 1 cm x 1 cm Si substrate are stabilized using alligator clips and placed into the solution vertically as the anode and cathode, respectively. A DC Regulated Power Supply was used to apply a voltage of 1.2V for 10 minutes to drive electrophoretic sedimentation of Au nanospheres to the surface. EDC initiated carbodiimide crosslinking chemistry is used to covalently bind carboxylic acid functionalized Au nanospheres to the selectively amine-functionalized PMMA regions of the template via an amide bond.[131] The substrate, Pt mesh, and beaker are rinsed with IPA for 1 minute and dried under nitrogen. The process is repeated on the same substrate with a freshly concentrated Au nanospheres solution, but with 30uL of EDC and s-NHS solution. Here electrohydrodynamic flow around previously attached nanospheres drives nanospheres to chemically crosslink via EDC coupling chemistry to nanospheres assembled in the first step. Finally, the diblock copolymer layer is etched under oxygen plasma (50W, 60s) in the PC2000 Plasma Cleaner from South Bay Technology (San Clemente, CA) to remove organic groups from nanoparticle surfaces. Assembly on indium tin oxide (ITO)-coated glass are carried out identically, with the exception of the surfaces undergoing oxygen plasma etching at 100 W for 1 min before spin coating PS-r-PMMA.

5.8.3 Characterization

Images of SERS substrates are collected with a Magellan XHR scanning electron microscope (FEI). UV-vis absorption measurements of SERS substrates are carried out using a Shimadzu UV-1700 absorption spectrometer. Absorption spectra are taken of ITO-coated glass substrates taped (away from the beam path) on to a quartz cuvette filled with DI water.

5.8.4 P. aeruginosa cell free supernatant preparation

Wild type *P. aeruginosa* (strain PA14[312]) shaking culture supernatant was used to measure pyocyanin production over time. PA14 was streaked onto lysogeny broth (LB, EMD Millipore) agar plates from frozen glycerol stocks and grown overnight at 37 °C. Shaking cultures were inoculated from single colonies on the LB plate into 5 mL of liquid LB and grown at 37 °C on an orbital shaker overnight. 50 µL of the shaking culture was diluted into 24 mL of 10 g/L tryptone media (Bacto tryptone, BD Scientific) and this subculture was grown at 37 °C on an orbital shaker. Aliquots were taken periodically from the shaking subculture for optical density measurements. To measure pyocyanin production, *P. aeruginosa* conditioned medium was isolated by centrifuging the 2 mL cultures at 15000 x *g* and passing the resulting supernatant through a 0.2 µm

PES vacuum filter (Corning). The filtered supernatants were dropped onto the SERS substrates and measured within 1 hour from their collection time.

5.8.5 Fluidic Device Fabrication and Biofilm Growth

The microfluidic device consists of two channels, one for the biofilm and the other for the SERS substrate. Microfluidic channels for the biofilm were made by bonding the plasma-activated surfaces of PDMS with a glass slide. SERS substrate and a glass slide are separated by an adhesive spacer of 100 μm thickness (3M 415) with channel design laser cut (Epilog Fusion Laser Cutter) in the adhesive layer.

P. aeruginosa biofilms for in-line detection of pyocyanin were grown in the above microfluidic devices as previously described using a *P. aeruginosa* strain constitutively expressing yellow fluorescent protein (YFP).[311] *P. aeruginosa* cells were seeded with no flow in the biofilm growth channel for 2 h. The fluid feed was then switched to sterile tryptone media and the outlet of the channel was connected to the SERS channel input. The 0 h SERRS measurements were collected once the in-line device in fully connected, prior to start of flow. Media was then pulled through the in-line detection device at 10 $\mu L \cdot h^{-1}$ and SERRS spectra were collected at the indicated time points. Identical biofilms were grown and imaged in the flow cells to measure the time-dependence of biofilm growth. The biofilm volume was obtained from Volocity imaging analysis software (PerkinElmer) of confocal fluorescence images of the YFP-producing biofilms. SERRS collection parameters are described below.

5.8.6 Antibiotic Susceptibility Measurements

P. aeruginosa biofilms were grown for 10, 15, and 24 h on glass coverslips submerged in 2 mL of TB supplemented with 3 $g \cdot L^{-1}$ NaCl in sterile, six well tissue culture plates (Fisher Scientific). After the indicated growth time, the growth media was aspirated and the biofilms on the coverslips were rinsed once with PBS while still in the wells. The PBS rinse solution was then aspirated and replaced with 600 $\mu g \cdot mL^{-1}$ carbenicillin (minimum inhibitory concentration in *P. aeruginosa* strain PA14 128 $\mu g \cdot mL^{-1}$)[313] in PBS or PBS only for the control. Biofilms were soaked in the antibiotic or control solutions for 3 h, rinsed again with PBS, and then stained with live/dead cell viability assay stains, propidium iodide and Syto 9 (Fisher Scientific) at 2 μM final concentration each, for 15 min. Coverslips were rinsed once more in PBS, removed from the wells and placed face down on a microscope slide for confocal fluorescence imaging. The fraction of dead cells was calculated using the biovolumes of each color channel (red and green) obtained from Volocity imaging analysis software (PerkinElmer) of confocal fluorescence images of biofilms.

At 6 h biofilm growth, the washing steps described above removed all cells from the coverslips. Instead, the antibiotic susceptibility of surface-associated cells was measured from cells grown in microfluidic devices. After 6 h of growth, as described above for pyocyanin detection experiments, the input line was changed from TB to 600 $\mu g \cdot mL^{-1}$ carbenicillin in PBS for 3 h. The input line was then changed again to BacLight for 15 min and finally PBS for 30 min, after which cells attached to the channel surface were imaged and the dead cell fraction quantified by counting individual cells of each color channel.

5.8.7 Spectroscopic, Measurements, Instrumentation, and Procedure

Surface enhanced Raman spectroscopy (SERS) measurements are conducted using a Renishaw InVia Raman Microscope system. A 785 nm continuous wave laser is chosen to excite near the plasmon resonance of Au nanoparticle assemblies as determined from electromagnetic simulations and ultraviolet-visible absorption measurements.[272]

In droplet measurements, as referred to in the main text, a 60X water immersion objective with 1.2 NA is used for illumination and collection. Approximately 150 μL of solution of interest is transferred onto SERS substrates and the measurements are acquired with laser power and acquisition time of 7.3 μW and 0.5 second, respectively, over area specified for each case. SERS substrates are cleaned with IPA and DI water for 1 minute each and dried under nitrogen between measurements. One substrate was used per set of measurements (concentrations or time points).

For in-line measurements, a 50X objective is used. Measurements are taken with laser power and acquisition time of 14.6 μW and 0.1 second, respectively. Illumination and collection were done through SERS microfluidic channel.

5.8.8 Spectra Processing and Analysis

Raman scattering spectra processing and analysis were performed off-line using MATLAB R2016b (The MathWorks Inc, Natrick, MA). Each spectrum undergoes baseline correction, smoothing with Savitzky-Golay, and normalization to the average intensity of Si second-order vibrational band, reported between 920 – 1045 cm^{-1},[314] from the substrate. This allows for comparison of different samples where slight intensity variations may arise due to deviations in optical collection in the experimental setup. While this band was previously only the range between 920 – 970 cm^{-1} was used due to the appearance of pyocyanin ring stretching vibrational band at 975 cm^{-1}. When preparing SERRS spectra for full-spectrum partial least squares (PLS) regression, a constant $k = 1.0067$ was added to the processed signals to eliminate negative values in the calibration matrix associated with variations due to noise. This is necessary as signals are then log-transformed before analysis with PLS regression. PLS regression combines characteristics of principal component analysis with multiple linear regression to predict a set of dependent variables from a large set of independent variables.[315]

5.8.9 Simulations

Full-wave simulations (frequency domain finite elements method solver) are implemented in CST Microwave Studio (CST AG). We calculate the absorption cross sections of several nanosphere oligomers: dimer, linear trimer, and linear quadrumer. The structure identical to that considered in chapter 3. Au nanospheres of 40 nm in diameter with 0.9 nm interparticle gaps are partially embedded in a 40 nm PMMA layer with their centers 8 nm above the PMMA. The PMMA layer is directly above a 150 nm ITO layer, which is directly above a 2 μm glass layer. The upper medium surrounding nanospheres is water. The total size of the structure in x-y plane (transverse plane) is assumed to be 4 μm × 4 μm, which is approximately four times larger than the largest excitation wavelength of 900 nm, and we applied software's open boundary condition. Au permittivity was calculated from Drude model with parameters extracted from Grady.[316] The relative electric permittivity used for water, PMMA, glass, and ITO are 1.77, 2.47, 2.3207, and adapted from Moerland,[254] respectively. The permittivity in the gap region was unknown due to anhydride bonding, and thus was approximated by performing a parameter sweep of the gap permittivity and calculating the absorption of the dimer configuration to determine which best corresponds with the observed dimer peak at 686 nm. The resulting permittivity of 2.25 was used in simulations. Oligomers are excited with plane wave illumination at normal incidence with electric field polarization along the axis of the linear oligomers, and the absorption cross section of the structures is determined.

Field enhancement was also calculated from full-wave simulations of dimer and linear trimer with 0.9 nm and 2 nm interparticle gaps using the same conditions. Field enhancement is defined in equation (4).

$$\mathrm{FE} = |E_{\mathrm{olig}}|/|E_0| \tag{6.4}$$

$|E_{olig}|$ represents the electric field magnitude at the center of the gap in the oligomer (dimer or trimer), and $|E_0|$ is the electric field magnitude at the same location in the absence of the structure.

Chapter 6

SERS-based Odor Compass: Locating Multiple Chemical Sources and Pathogens

6.1 Introduction

From October 2015 through February 2016 100,000 tons of methane and 2.5 tons of benzene were leaked from the Aliso Canyon natural gas storage facility, making it the second worst natural gas leak in US history.[317] Rapid and early detection and identification of hazardous gas leaks is essential to reduce the damage of these disasters, dangerous industrial gas leaks, and contamination of virulent bacteria. Yet most gas sensors merely alert of the presence of gasses and give no indication of the gas source direction. Often, leaks are identified using a mobile odor detector and gas distribution mapping, an incredibly time consuming process.

To improve upon existing methods, the odor compass was developed in 1991.[318] The canonical odor compass is composed of two semiconductor gas sensors mounted to a rotating stage with a fan mounted to draw analyte containing air towards the gas sensors.[319,320] Heuristics are used to identify odor source direction by comparing the relative signal from the gas sensors at various rotation angles.[321,322] These odor compasses are mounted onto mobile robots so that they can travel through the gas concentration gradient and eventually find the odor source, a process called chemotaxis in the biological context. Often, simple algorithms inspired by Braitenberg vehicles[323] are used, but sophisticated methods for odor source localization have been developed using genetic algorithms,[324,325] fuzzy logic,[326,327] and infotaxis.[328] Mobile odor compasses have now been used on land, air,[329,330] and sea.[331,332]

While most odor compasses rely on semiconductor gas sensors,[333,334] other gas sensors have been used such as conductive polymers,[335,336] quartz crystal microbalance gas sensors,[337,338] and even silkworm antennae.[339] Yet (aside from silkworm antennae), these detectors cannot effectively differentiate different types of gasses, limiting their usefulness in many applications. More sophisticated sensors are thus necessary for odor identification. Classification of odors during odor localization was first achieved with an electronic nose,[340,341] composed of an array of semiconductor gas sensors with differing composition. Electronic noses have been shown to be very effective at identifying odors in a range of applications, particularly food[342] and health applications.[343] Nevertheless, electronic noses suffer from two limitations: 1) they require a seperate electronic readout for each sensor in the array limiting the number of unique sensors to often as little as 4,[340] making classification of similar odors challenging, and 2) a limit of detection (LOD) of analyte in the parts per million range is typical for electronic noses used as an odor compass.[344,345] In comparison, many biological systems have detection limits reaching the parts per trillion range.[346] The achievable limits of detection of typical odor compasses have limited their applications to plume detection. This necessitates the use of fans to draw in odor plumes, mechanical rotation[347] of the odor compass apparatus, and robotic system to travel through large areas in order to identify the true odorant direction, greatly slowing the odor localization process.[321,322]

We present the use of surface enhanced Raman scattering (SERS) sensors to obviate many of the challenges associated with odor source direction identification. SERS brings three main benefits: 1) the use of vibrational spectroscopy enables superior differentiation of even extremely similar analytes,[348] 2) SERS has extremely low limits of detection of molecules, with typical operation in the parts per billion, and can reach LODs as low as 1 part per trillion with the use statistical analysis and appropriate nanoarchitectures that can achieve single molecule SERS,[349] and 3) each sensing element does not need a seperate electrical readout. In this chapter, we demonstrate that a passive (i.e., without fans or motors) SERS sensor array can identify the direction of multiple odor sources with similar vibration spectra relying only on diffusion of analytes to reach the array. We use a machine learning approach to analyze the SERS spectra, which has been shown to be useful in this context.[177,187,188,273] Raman spectra are first analyzed with non-negative matrix factorization to differentiate the analytes, benzenethiol and 3-methoxybenzenethiol. Then, various machine-learning classifiers are compared in terms of their ability to correctly identify the odor source direction via evaluation of SERS data from the sensor array. Machine learning interpretation of signals from sensor arrays realize a marked improvement of well over 20% classification accuracy compared to heuristics typically used for odor compasses. The best in class models, support vector machine classifier and convolutional neural networks, correctly identify the direction of one or two odor sources over 90% of the time vs 4.2% chance. Models are also evaluated with k-folds cross validation and categorical accuracy achieves a cross validation standard deviation of less than 0.2%. Finally, we demonstrate that this approach can be used to detect the presence of *Escherichia coli* biofilms as well as identify its source direction by tracking volatile organic compounds from the bacteria.

6.2 SERS Sensor Array Fabrication and Validation

Passive odor compasses relying on diffusion of analytes to sensor surfaces require significantly better sensitivity, limit of detection (LOD), and precision compared to typical plume based detection due to both lower gas concentrations and smaller differences in analyte concentration from point to point. Uniformity of signal response across a passive sensor is necessary to ensure variations are related to analyte concentration rather than variable sensor response. In order to fabricate nanosensors with spatially uniform response we use 2-dimensional, physically activated chemical assembly (2PAC), to produce surface enhanced Raman scattering (SERS) sensors with billion-fold Raman scattering enhancements over large areas.[272] 2PAC assembly involves seeded growth of close-packed gold nanosphere assemblies, described in detail elsewhere.[272,273] Briefly, nanosphere seeds are driven to deposit on select regions of a copolymer coated Si electrode by electrophoresis. After seeding, the interaction between the electrical double layers of the electrode and the nanospheres drives electrohydrodynamic (EHD) flow. This lateral, attractive flow entrains nearby particles, forming close-packed nanosphere assemblies with carbodiimide-mediated crosslinking between carboxylic acid groups on adjacent nanoparticles. 2PAC results in uniform gap spacings of approximately 0.9 nm, which leads to reproducible SERS enhancements and thereby signal. A scanning electron microscopy (SEM) image of a 2PAC assembled SERS surface is depicted in Figure 6.1a. One may observe discrete, close-packed nanosphere assemblies. Discrete assemblies have greater field enhancements than 2-D close packed films, while still having a relatively dense collection of hotspots for analyte molecules to reside. The reproducibility of the gap spacings due to chemical crosslinking and the rotational invariance of the close-packed oligomers enables the surfaces to be optically uniform within the laser spot size across large sample surfaces.[350]

In this work, we focus on using SERS sensors to not only to detect but also to identify the direction of analyte sources. We show that 2PAC fabricated SERS sensors in the vicinity of liquid phase analyte sources are sensitive to gradients arising from diffusion based gas transport of the vapor. Analytes from the gas phase will chemisorb onto nanosphere assemblies and the spatial concentration gradient measured across sensors is used to locate the source of the analyte. In order to ensure that diffusive gas transport of analytes dominates the measured concentration gradient (as opposed to fan driven plume transport typically used in odor compasses), we place SERS sensors in a 3 x 3 array into a sealed desiccator at a pressure of 0.5 atm. The arrangement of SERS sensors and analyte source is depicted schematically in Figure 6.1b. A 5 µL drop of neat analyte is placed on one side of the array. Analytes used in this chapter, benzenethiol (BZT) and methoxybenzenethiol (MBZT), are chosen due to the similarity of their SERS spectra and affinity to chemisorb on Au. Both analytes are in the liquid phase and have a vapor pressure of 0.13 kPa and 1.33 kPa at room temperature for BZT and MBZT, respectively. Analytes are exposed to SERS sensors for 15 minutes before characterizing the sensor response *ex situ*. For SERS data acquisition, nine Raman maps (one from each sensor in the array) are acquired in 128 µm x 128 µm regions across each sensor surface. The point to point reproducibility across a sensor allows for acquisition of a higher quality data sets than repeated illumination of a single spot. It is known that repeated, prolonged laser illumination of a diffraction limited spot on a SERS surface can lead to photodegradation of nanosphere assemblies due to the high temperature in the hotspots.[351] Also, when using sensors having hotspots with high signal enhancements, it is essential to sample many regions to properly train the model to account for small signal variations which may arise due to variations in analyte chemisorption on hotspots.

Characteristic SERS spectra are plotted in Figure 6.1 c and d. Preprocessing of SERS spectra before machine learning analysis is essential to aid in distinguishing analyte signals from other molecular signals that may emerge from the sample environment and will limit generalizability of machine learning models for environmental sensing. The processed spectra are then divided into training and validation sets in an 80/20 split. Next, the training dataset is used to train non negative matrix factorization (NMF) dimensional reduction (when closure, unimodality, and local rank constraints are used with alternating least squares (ALS) optimization, NMF is called multivariate curve resolution (MCR-ALS), which is also used in spectroscopy).[352,353] The dimension of spectra are reduced from 1011 wavenumber features to 3 NMF scores. Finally, a spectrum from each sample in the 9 sample grid are bundled into a 3 x 3 matrix for each NMF score; the value of each spectra bundle are scaled to a mean of 0 and variance of 1 to improve model convergence.[354] These NMF score matrices are depicted schematically in Figure 6.1 e.

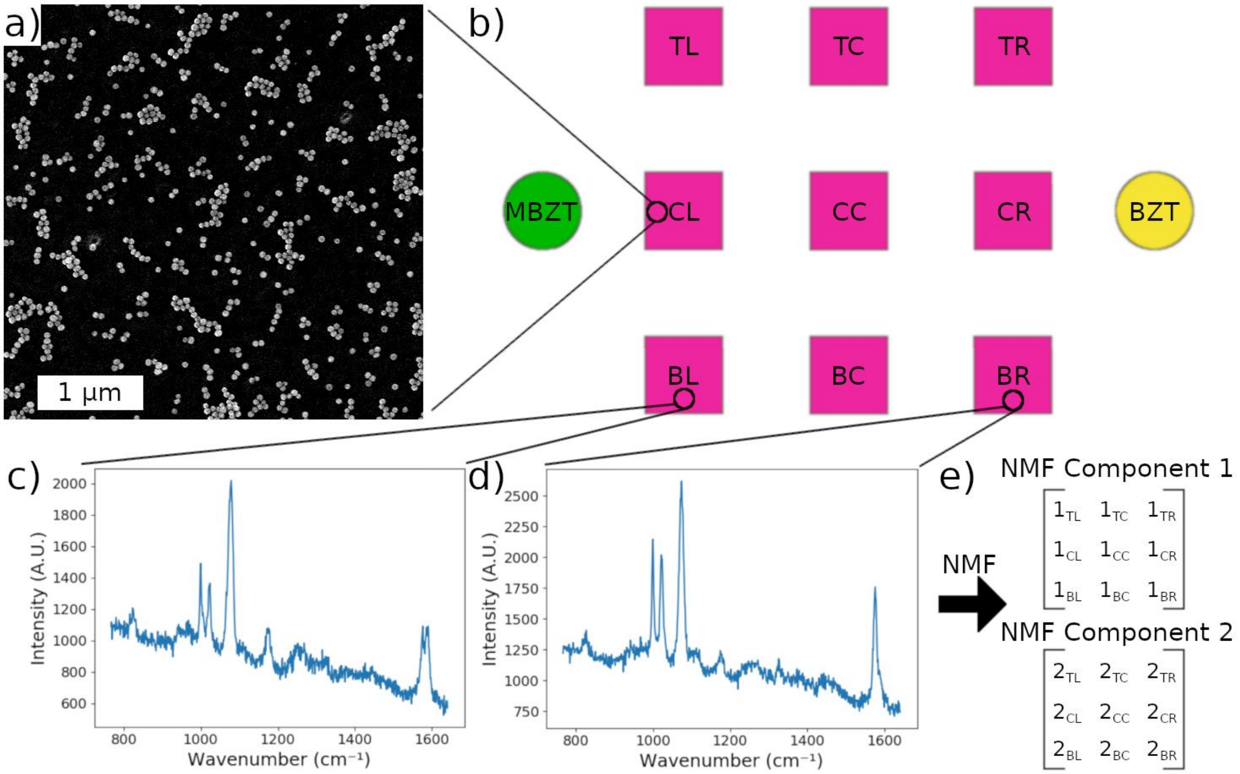

Figure 6.1: a) SEM image of nanosphere assemblies that comprise the SERS sensors. b) Schematic of SERS sensor array and example of multi-analyte placement with respect to array. This schematic depicts the case where MBZT is exposed to the left of the sensor array, and BZT is exposed to the right of the sensor array. Representative SERS spectra of a BZT, MBZT mixture acquired from SERS sensor in c) the bottom left of the array and d) the bottom right of the array. SERS spectra are acquired with a 785 nm diode laser at 760 µW and 0.3 s exposure time. e) Schematic of the resulting input into the model that is constructed from NMF decomposition of the SERS spectra acquired across the sensor array.

6.3 Visualization of Odorant Chemisorption on SERS Surfaces

A intrinsic advantage of SERS over comparable methods, such as an electronic nose, is that the signal is composed of the vibrational spectra of analyte molecules enabling visual matching of signal contributions from individual analyte to the observed SERS spectra. In a complex mixture SERS signals are additive, thus NMF is well suited to to isolate molecular signals in complex spectra because the non-negativity constraint results in a decomposition of the spectra into a parts-based representation with contributions from different analyte primarily emerging as the different parts (components).[355] Figure 6.2 a, c, and e depict plots of the three NMF components extracted from the total training dataset. The training dataset contains spectra of both odor sources, BZT only, and MBZT only. Comparing raw SERS spectra from sensors only exposed to BZT or only to MBZT, plotted in Figure 6.2 b, and c, respectively, — one can easily identify NMF component 1 as mainly corresponding to signal caused by BZT and NMF component 2 mainly corresponding to signal caused by MBZT. NMF component 3 is associated with background signals emerging from the sensor surface chemistry (ligands on the gold nanospheres, polymer template, *etc*.) and molecules in the ambient environment due to the ultralow detection capacity of the sensor.

Figure 6.2 f depicts the values of NMF scores extracted from sensor elements in the array exposed to MBZT and BZT sources emerging from opposite directions. The average values of NMF scores acquired from a sensor element in the same column in the grid, which are expected to have similar scores due to the symmetry of the deposition, are within 10% of one another. Though more variance is observed within a sensor element’s Raman map, with the largest relative standard deviation (RSD) approaching 20% for BZT. This is significantly larger than the 10% RSD reported previously for these Raman sensors soaked overnight in a BZT solution in chapter

3 from identically manufactured SERS sensors.[272] Thus we attribute this variance to two effects: 1) incomplete chemisorption of analyte into the hotspots of the gold nanosphere assemblies and 2) leakage of the NMF components into one another. Leakage may be observed by comparison of neat analyte spectra with their respective components.

From Figure 6.2 f, one observes that a clear trajectory across the SERS sensors in the array emerges in the NMF scores clearly reflecting the location of analytes. The NMF score increases for sensor elements in the array closer to their corresponding analyte. While promising, for the general multiple source problem, the variance in NMF score complicates source direction assignment, and will be further complicated when the odor sources are separated by just 90°. Thus instead of simply comparing average sensor element contributions from a row or column on the sensor array to different NMF components, a more holistic, supervised machine learning method must be used to achieve high prediction accuracy for analyte location.

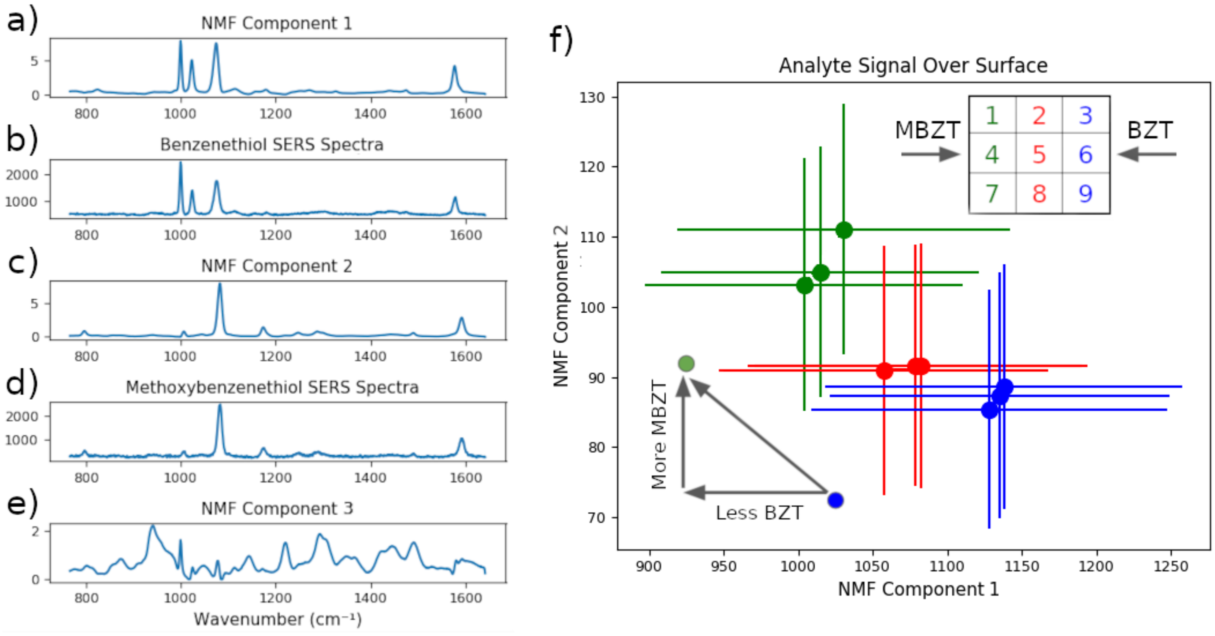

Figure 6.2: Non-Negative Matrix Factorization (NMF) components determined from analyte training datasets. NMF components a) 1, c) 2, e) 3 most strongly corresponds to b) neat BZT SERS spectrum, c) neat methoxybenzenethiol (MBZT) SERS spectrum and e) background spectrum, respectively. f) NMF scores of the BZT and MBZT components in the sensor array. Upper right inset: schematic of SERS sensor array and location of odor source for the plotted data. Crosses correspond to average value of the 800 test spectra acquired in each SERS sensor in the array. The lines represent one standard deviation of NMF values within a sensor element. Lower left inset: schematic of the trajectory of analyte signal across the sensor array.

6.4 Model Performance on Multi-Source Task

While supervised machine learning methods have been applied to the problem of tracking an odor source using a mobile odor compass, comparatively little effort has been put into supervised methods of identifying odor source direction from a stationary measurement. We evaluate 4 machine learning models and 1 heuristic to identify the model best fit for this task, a 2D convolutional neural network (CNN), an artificial neural network (ANN), a support vector machine classifier (SVM), and a k-nearest neighbors classifier (KNN). The label generation for odor location in our supervised approach is depicted in Figure 6.3. Specifically, we consider the 1 or 2 odor source problem with distinct odors with a location without multiple sources of the same odor and a resolution of 90°, resulting in 24 possibilities. The CNN uses the full 3x3 pixel spatial relationships of the input data, while the inputs are flattened for the other classifiers. The architectures of the neural networks are described in the methods section. This results in a multi-class, multi-label problem as one or two labels may be true, thus both neural networks are trained with a binary crossentropy loss function. We simplify the problem for the SVM and KNN, evaluating them with labels transformed from a binary representation of the 4 possible directions per odor to 24 labels total, one for each possible outcome. The heuristic first evaluates the variance of each input matrix to determine whether or not an odor exists based on a user defined threshold, and then the average value from the rows or columns are compared with its opposite and the largest difference is used to identify the odorant direction for each matrix. Each model performance and the heuristic are evaluated using the categorical accuracy and plotted in Figure 6.4 a.

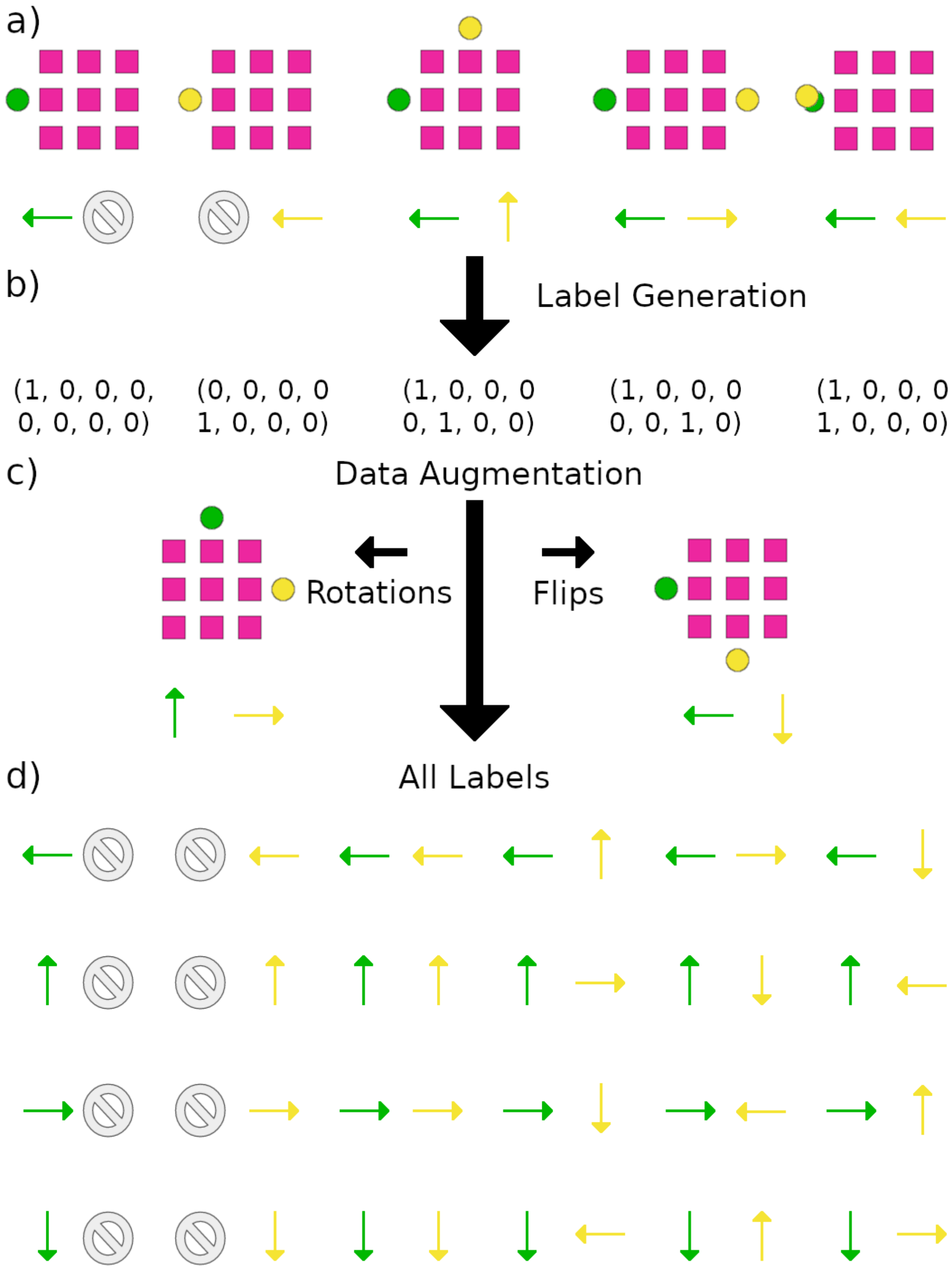

Figure 6.3: a) Schematic of the deposition conditions used in this work, and b) their corresponding labels. c) Schematic of the data augmentation used in this work, where small random numbers are also added to the NMF scores of the training data. d) Schematic of 24 labels possible in this work.

The prediction accuracy can be greatly improved by simultaneously considering multiple input data matrices. Averaging sensor inputs has been an important method for improving sensor precision and is ubiquitous in natural olfaction in the form of sniffing, where it greatly improves the accuracy of predictions made by mammals[356] and spike timing dependent plasticity models of a moth olfactory bulb.[357] We implement an averaging procedure, with each extra input referred to as a sniff, and plot the categorical accuracy over sniffs in Figure 6.4 a. Extra sniffs are added to the inputs as time series in a LSTM for the CNN model, and are added to lengthen the input vectors for the other models. The inputs are averaged elementwise for the heuristic.

All models and the heuristic greatly benefit from the averaging procedure, with the 7+ sniff SVM and 9+ sniff CNN achieving >90% categorical accuracy. The CNN receives a disproportionate boost in model performance from the extra sniffs, especially the 2-4 sniff models due to the integration into a LSTM as a time series, while the KNN actually declines in accuracy beyond 4 sniffs due to the increased sparsity of the input space. Although the sniffing procedure increases the acquisition time, acquiring one SERS matrix only consumes 0.9s of exposure time, so even a 10 sniff input requires only 9 seconds of exposure time, which is significantly less than is necessary for many plume based odor compasses.[322]

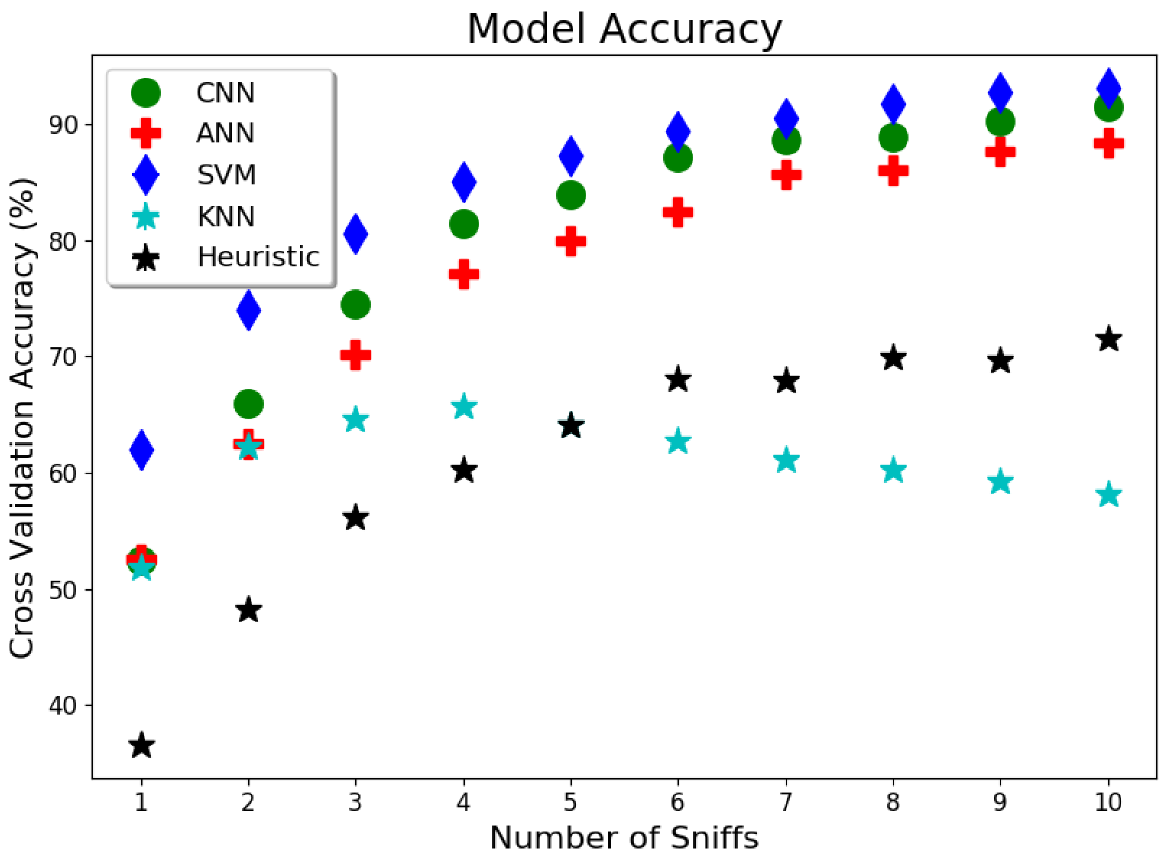

Figure 6.4: Cross validation accuracy of the models used. Error bars are obstructed by the markers, but are less than 0.45% in all cases. For a prediction to be considered correct, the direction of both odors must be correctly identified by the model.

SERS's rich spectral information and low limit of detection for small molecules have made it an important technique in the detection and discrimination of bacteria by sensing the volatile organic compounds (VOCs) that it produces.[358–360] Here, we use *E. coli* as an odor source to test the efficacy of our odor compass with complex mixtures of VOCs. SERS arrays are exposed to *E. coli* VOCs in a desiccator at ambient pressure and for 6 hours. Similar to chemical analytes measured above the E. coli source is placed adjacent to of the array. After exposure, the sensor elements are characterized *ex situ*. The resulting dataset is decomposed into NMF components which were compared to SERS spectra of E. *coli* acquired at 785 nm wavelength excitation. Figure 6.5 a depicts the NMF component that is in best agreement with previous work reporting SERS analysis of E. *coli*.[151] It is also important to note that this NMF component provides the highest accuracy of predictions for the location of the *E. coli* biofilm. The dataset is then reduced to the score of this component and evaluated with a 4 class, 1 sniff SVM model as in this single odor source problem there are just a total of 4 possible outcomes and there is just one possible label. A categorical accuracy of 82.95% is achieved on the test dataset, vs 25% chance. A confusion matrix of the test dataset is plotted in Figure 6.5 b, demonstrating well behaved predictions that are invariant with rotation.

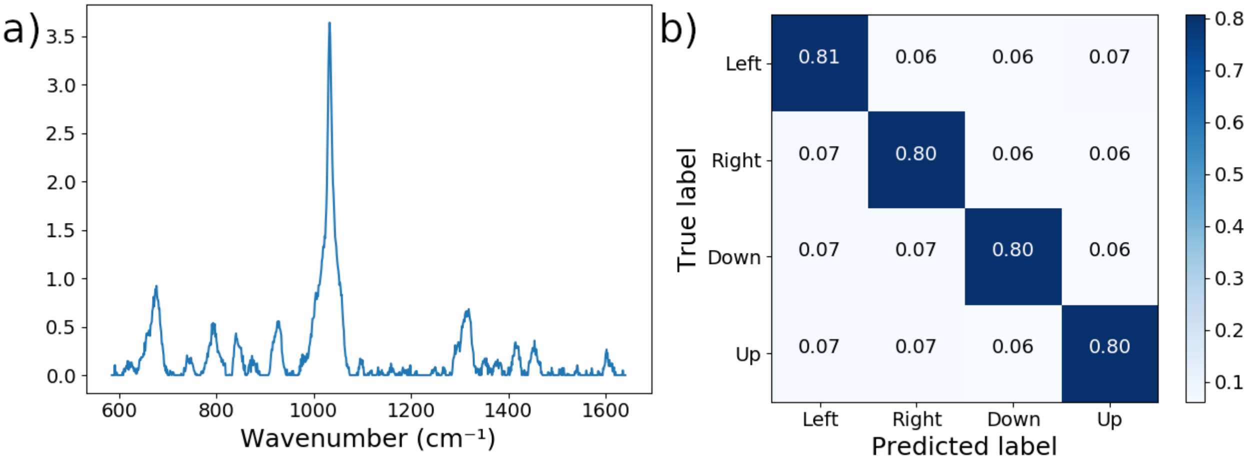

Figure 6.5: a) NMF component acquired from bacterial VOC training dataset used in bacterial odor source localization b) Normalized confusion matrix produced by 1 sniff SVM model applied to bacterial VOC test dataset.

6.5 Conclusion

In this chapter, we have demonstrated the efficacy of a diffusion based odor compass for the multiple odor source problem using surface enhanced Raman scattering (SERS) sensors. Specifically, we have implemented a sensor grid array to identify small spatial variations in analyte chemisorption and used various machine learning models to identify multiple analyte source direction. We have shown that support vector machine classifier and convolutional neural network models can achieve greater than 90% categorical accuracy for this multiple source problem with 90° resolution. We have further shown that the sensor grid array can be used to locate complex odor sources such as *Escherichia coli*. This work uses SERS as the gas sensor in an odor compass and paves the way for its use in passive internet of things devices and mobile odor compasses. We envision this integrating this strategy with newly emerging waveguide excitation of SERS sensors for device miniaturization and continuous monitoring.[249,361] These devices promise to make a material impact on people's health by identifying and locating toxic gases and volatile organic compounds produced by indoor pathogens.

6.6 Methods

6.6.1 Materials

Random copolymer poly(styrene-co-methyl methacrylate)-α-hydroxyl-ω-tempo moiety (PS-r-PMMA) (M_n = 7400, 59.6% PS) and diblock copolymer poly(styrene-b-methyl methacrylate) (PS-b-PMMA) (M_n = 170-b-144 kg mol −1) were purchased from Polymer Source, Inc. (Dorval, Canada). 40 nm diameter gold nanospheres were purchased from Nanocomposix (San Diego, CA, USA). Si(001) wafers with a resistivity of 0.004 ohm-cm were purchased from Virginia Semiconductor (Frederickburg, VA, USA). Hydrofluoric acid (HF) was purchased from Fisher Scientific (Pittsburgh, PA, USA). 2-(N-morpholino)ethanesulfonic acid (MES) 0.1 M buffer, 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide hydrochloride (EDC), and N-hydroxy sulfosuccinimide (s-NHS), dimethyl sulfoxide (DMSO), ethylenediamine, benzenethiol, toluene, ethanol, isopropyl alcohol (IPA), potassium carbonate, and 52-mesh Pt gauze foil were purchased from Sigma-Aldrich (St. Louis, MO, USA). Nanopure deionized water (DI) (18.2 MΩ cm −1) was obtained from a Milli-Q Millipore System.

6.6.2 SERS Sensor Fabrication

Sensor fabrication has been described in depth in previous chapters.[272,273] First, block copolymer templates for nanoparticle attachment are prepared. Random PS-b-PMMA block copolymer is spin-coated onto a HF-cleaned (*the potential of HF to cause severe injury mandates extreme caution during usage*), 0.004 ohm-cm Si wafer and annealed for three days followed by a toluene rinse and spin coating of lamella forming PS-b-PMMA block copolymer and further 3 days of annealing, as described elsewhere.[131,137,141] Then, the PMMA regions are selectively functionalized with amine end groups by immersing the entire substrate in DMSO and then in ethylenediamine/DMSO solution (5% v/v), both for 5 min without rinsing between steps. The functionalized template is then washed with IPA for 1 min and dried under nitrogen, and used immediately.

Au nanosphere assemblies are deposited on the surface as follows: A Au nanosphere solution (0.1 mg/mL, 3 mL) is added to a 10 mL glass beaker. Then s-NHS (20 mM) in a MES (0.1 M) buffer (35 μL) is added to the beaker and swirled. Afterwards, EDC (8 mM) in a MES (0.1 M) buffer (35 μL) is added to the beaker and swirled. This beaker is placed on a hot plate and brought to 60 °C. Next, a 1 cm × 1 cm functionalized template-coated Si substrate is placed into the solution vertically and held in place with alligator clips. 1 mm away from the substrate, a 1 cm × 1 cm Pt mesh is placed into the solution vertically. A DC power supply is used to apply a voltage of 1.2 V across the Pt mesh and substrate for 10 min. The substrate, Pt mesh, and beaker are rinsed with IPA for 1 min and dried under nitrogen. This process is repeated with the same substrate and fresh nanosphere solution as described above, but with 25 μL of EDC and s-NHS solution.

6.6.3 Analyte Deposition

SERS sensors are exposed to analyte with 9 sensors arranged in a grid. Depositions are performed in a sealed desiccator with a 10 cm diameter. A 5 μL drop of neat analyte is placed on a glass slide at the edge of the desiccator, when a second analyte is used it is similarly placed on a glass slide at the edge of the desiccator either 90° offset, 180° offset, or adjacent to the first analyte. The sensor grid array is organized such that each sensor is placed 3 cm away from other sensors to form a square grid, with an analyte - sensor spacing of 2 cm for the closest sensor. The desiccator was then sealed and brought to 0.5 atm for 5 minutes followed by 15 minutes of static deposition.

6.5.4 Characterization

Au nanoparticle assemblies are imaged with a Magellan XHR SEM (FEI). All Raman spectroscopy measurements are conducted using a confocal Renishaw InVia micro Raman system with a 785 nm diode laser, a laser power of 760 μW, an exposure time of 0.3 s, and a 50x air objective with a 0.75 NA. Raman maps are collected with a spacing of 2 μm spacing between points. For the sensor grid array, one 64 x 64 pixel Raman map is acquired per sensor, while for the single sensor measurement, 64 x 64 pixel maps were acquired with 3 mm spacing between each map, arranged on a square grid.

6.6.5 Spectra Preprocessing

Raman spectra preprocessing was performed using Python 3.3 programming language. Savitzky-golay smoothing was implemented with Scikit-learn, using an 11 pixel window and polynomial order 3. Asymmetric least squares background subtraction was implemented in NumPy with $\lambda = 10000$, $p = 0.001$. Non-negative matrix factorization (NMF) was implemented with Scikit-learn, 3 components, and default settings, trained only on the dataset segregated for use in training the models. This trained NMF transformation was used to reduce the dimension of all data to 2. NMF reduced maps from each point on the sensor grids were then rearranged into a 3x3 matrix, with two NMF scores per map per datapoint. Each datapoint was then scaled to a variance of 1 and mean of 0.

Data augmentation was performed by rotating the rank 2 tensors that compose each datapoint 90°, and similarly rotating the labels. A gaussian random number ($\mu = 0$, σ = standard deviation of that component extracted from each deposition condition * 0.1) is added to each component for the training data. These data are then stacked along a third dimension, called “sniffs” producing a rank 4 tensor with shape (sniffs, sensor grid axis 1, sensor grid axis 2, number of analyte). For the SVM, KNN, and ANN models these inputs are flattened into a vector before being fed into the model. For the heuristic, the components are averaged along the sniff axis and then flattened, and each analyte is considered separately.

6.6.6 Compass Models

Support vector machine classifier (SVM) and k-nearest neighbor classifier (KNN) were both implemented in Scikit-learn using default settings. The artificial neural network (ANN) and convolutional neural network (CNN) models were implemented in Keras. The ANN model is composed of an input layer, 3 fully connected layers with succeeding dropout layers, and one fully connected output layer with 8 sigmoid nodes. The CNN model is composed of an input layer, then 4 time distributed 2D convolutional layers performed over the sensor grid axes with succeeding batch normalization layers followed by a time distributed 2D maximum pooling layer with a pool size of 2x2 and strides of 1x1. The pooling layer is followed by 2 more time distributed 2D convolutional layers and a time distributed flattening layer. Next, a long short term memory layer is used across “sniffs” axis. Following that layer are 3 fully connected layers with succeeding batch normalization layers and dropout layers, outputting ultimately to a fully connected output layer with 8 sigmoid nodes. Each output node on the ANN and CNN are associated with a direction and analyte, so binary crossentropy is used as the loss function. Early stopping and reduction of learning rate on a performance plateau are implemented using test loss as the metric. The heuristic identifies odorant direction in two steps, first identifying the odorant, and then the direction. First, the variance of each input matrix to determine whether or not an odor exists based on a user defined threshold. Then the average value from the rows or columns are compared with its opposite and the largest difference is used to identify the odorant direction for each matrix.

The generalizability of each model is performed using k-folds cross validation. First 20% of the data is removed as the validation dataset. The remaining 80% of the data is used to train the NMF transformation which is also applied to the validation dataset. Then k-folds cross validation is implemented on the remaining 80% of the data using Scikit-learn, with the number of folds equal to 5. The training fold is used to train each model. Each test fold is used to train the early stopping for the ANN and CNN models but are discarded for the SVM and KNN models. Finally, each of the trained models are evaluated on the validation dataset, the average and standard deviation across the 5 folds are plotted in Figure 6.6a.

6.6.7 Static Biofilm Preparation and Characterization

Escherichia coli MC4100 strain frozen stock was streaked on Lysogeny broth (LB) agar plates and grown overnight at 37 °C. One colony was harvested, inoculated into 2 mL of liquid LB and grown overnight at 37 °C on an orbital shaker at 200 rpm. The overnight culture was diluted with 10 g/L bacto tryptone (TB) to an optical density (OD) of 0.02. 2 ml of the diluted culture was added to each well of sterile 6-well plate. One 18x18 mm coverslip submerged in each well and plate was kept in room temperature (22 °C) for 24 hours. Three of the coverslips and their inoculum solution were transferred into a sterile petri dish that was placed into a covered 10 cm diameter desiccator. A sensor grid was then placed in the desiccator as described above and the system was kept isolated for 4 hours. These sensors were then removed and SERS measurements were performed as described above.

Chapter 7

Improved Concentration Regressions with Convolutional Neural Networks for Surface Enhanced Raman Scattering Sensing

7.1 Introduction

Vibrational spectroscopies hold enormous promise as chemosensors due to the "fingerprint" spectral region – roughly from 400 cm^{-1} to 1700 cm^{-1} – that can be used to uniquely identify small molecules. Yet traditional vibrational spectroscopies like infrared spectroscopy and Raman spectroscopy have insufficiently large absorption and scattering crossections, respectively, for use in sensing at small concentrations.[14] Surface enhanced Raman spectroscopy (SERS) has emerged as a promising method for enhancing Raman scattering to the point where small concentrations of analyte molecules can be quantified.[362] SERS relies on the near electric field enhancement of light due to scattering with metal nanostructures,[38] which can result in increases in Raman scattering by factors as large as 10^9.[255]

SERS has many attractive qualities for sensing. Large SERS EFs eschew the need for labeled detection schemes as even nonresonant molecules with small Raman cross-sections can be directly sensed.[142] SERS has been used for the detection of explosives,[363] drugs,[364] and toxins.[365] Additionally, the fingerprint spectral window enables highly multiplexed sensing.[366] Finally, SERS surfaces enable real time, in-line measurements for longitudinal monitoring of molecules, making it ideal for monitoring infections.[273]

Fabrication methods compatible with nanomanufacturing, specifically chemical self-assembly, have advanced to the point where the sub-nanometer gaps between nanostructures used in SERS can be reliably produced.[222,224,367,368] In the context of plasmonics these gaps, called hotspots, yield unprecedented enhancements of light matter interactions.[204] We have recently demonstrated a chemical assembly method – 2-dimensional physically activated chemistry (2PAC) – capable of obtaining a SERS EF exceeding 10^9 with just 10% relative standard deviation over a 1 mm^2 area.[272] This advance has enabled the collection of large, uniform datasets that enable a big data, machine learning approach to SERS sensing.[369] Traditional methods of calibrating SERS surfaces' quantitative response only track a single Raman band, an analysis termed univariate linear regression, and discard the remaining rich spectral information.[293,305] We have recently shown that this technique may be greatly improved upon with multilinear regression, due to fluctuations in band enhancements relative to one another.[308]

The complexity in quantitative, predictive SERS has led to the adoption of machine learning methods such as partial least squares,[273] support vector machines,[177] and artificial neural networks (ANNs).[370] Of these ANNs have several properties uniquely advantageous for SERS. As a nonlinear method, ANNs are well suited for deviations from Langmuir adsorption dynamics.[371] ANNs also have the universal approximation property that enables them to approximate any latent random variable, an essential property for multiplexed SERS.[372] Finally, they naturally denoise complexly varying signals.[373]

Yet ANNs are not perfectly suited for predicting analyte concentration from maps of SERS spectra. Underlying spatial variation of Raman enhancement and hotspot volume from the SERS sensing surface leads to variance in predictions.[308] Indeed, consider that at low concentrations, analyte molecules will not be uniformly distributed across the surface.[297,374] In this work, we obviate these challenges by employing convolutional neural networks (CNNs). 1D CNNs, which have been shown to perform well on data having spatial relations in a single dimension, such as stock price,[375] electroencephalograms (EEGs),[376] audio signals.[377]

Here, we bundle multiple Raman spectra into a 1D line with channels representing the wavenumber features of each spectrum. This strategy enables the network to learn the underlying distribution of Raman enhancement and hotspot volume within sensor surfaces. Thus, each input is given as a sparse distribution of these latent variables in addition to the analyte information. In a complex biological fluid, we train models to predict concentrations of pyocyanin, an important biomolecule.[274] We demonstrate that this approach significantly improves prediction accuracy, yielding an r^2 of 0.95, compared to the ANN r^2 of 0.89. Further, we demonstrate that this model uses the important spectral information from the analyte and is capable of making reasonable interpolations between trained concentrations.

7.2 Results and Discussion

2-dimensional physically activated chemistry (2PAC) is used to self-assemble high quality surface enhanced Raman scattering (SERS) sensing surfaces. Figure 7.1a depicts a schematic of the 2PAC assembly, which is a seeded growth of close-packed nanoparticle assemblies. 2PAC is described in more detail elsewhere.[272] Briefly, nanoparticle seeds are driven to a copolymer coated Si electrode by electrophoresis. EDC activates the carboxylic acid functionalized nanospheres, enabling a crosslinking reaction to amine functionalized domains of the copolymer. At the surface, the interaction between the electric double layers of the electrode and the nanoparticles drives electrohydrodynamic (EHD) flow. This lateral, attractive flow entrains nearby particles, forming close-packed nanosphere assemblies. The close proximity of the nanospheres within oligomers gives rise to repeated collisions, enabling a carbodiimide-mediated crosslinking between carboxylic acid groups on adjacent nanoparticles that results anhydride linked gaps of ~0.9 nm.

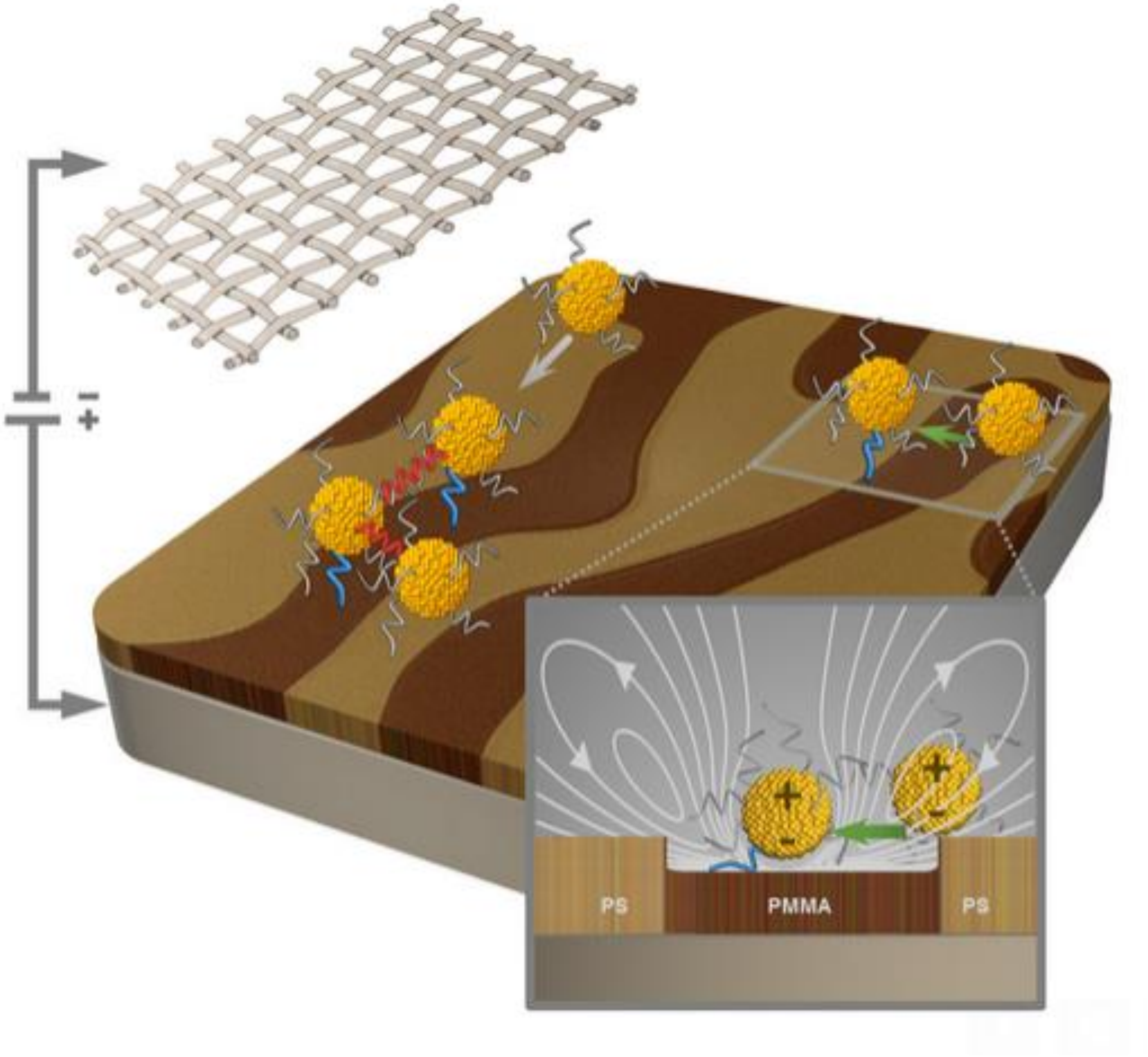

Figure 7.1. a) Schematic of 2PAC self-assembly. Electrophoresis drives carboxylic acid functionalized nanoparticle "seeds" to collide with the amine functionalized block copolymer coated electrode where EDC crosslinks them to the surface. Inset: seeds drive electrohydrodynamic flow that entrains nearby particles enabling a second EDC activated crosslinking between particles yields anhydride linked two-dimensional close-packed assemblies.

Raman spectra used to train and validate our models are acquired from concentrations gradients of the biomolecule pyocyanin. To simulate biosensing environments, pyocyanin is dissolved in LB broth, a bacterial nutrient solution. Figure 7.2 depicts representative Raman spectra of pyocyanin where characteristic Raman bands are observed at 552, 1353, 1602, and 1620 cm^{-1}, among other spectral regions. SERS spectroscopy is performed on the surfaces by acquiring 16 x 24 µm maps with 2 µm spacings between acquisitions. Spectra are acquired for 0.1 s at a laser power of 760 µW at 785 nm with a 1.2 NA water immersion objective.

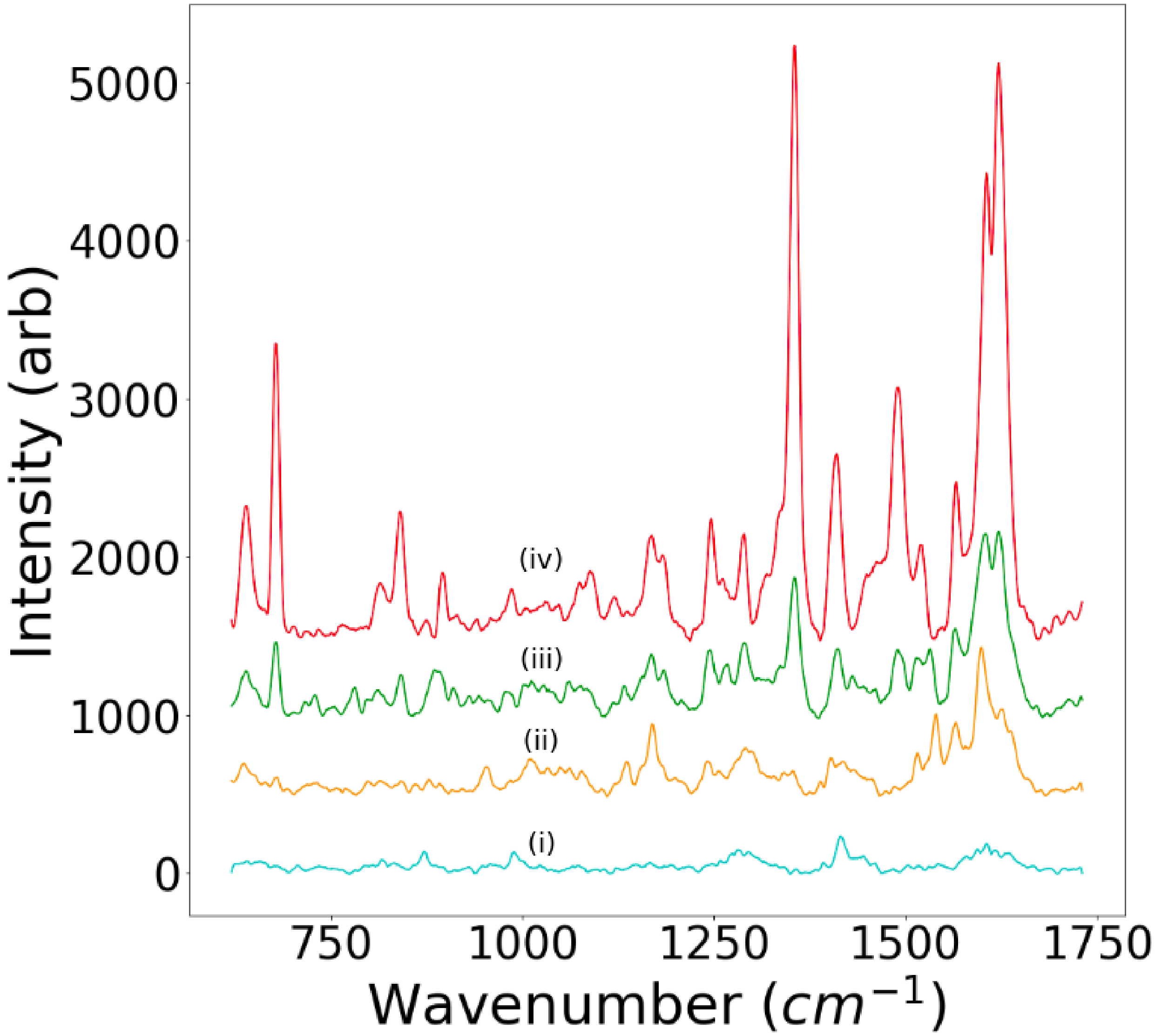

Figure 7.2 depicts representative Raman spectra acquired from 2PAC assembled surfaces with pyocyanin concentrations of i) 0 ng/mL ii) 100 ng/mL iii) 1 µg/mL and iv) 10 µg/mL. Spectra are offset by 500 units for visual clarity.

In this work, we seek to demonstrate that grouping multiple measurements together as inputs into a regression model significantly improves concentration predictions. To this end, we evaluate several methods for quantifying concentration from Raman spectra. A 20/80 test/train split is performed, and the models are evaluated based on its test r^2. Early stopping based on r^2 is used on all neural network models as a regularization method. Principal component analysis (PCA), just fit to the training dataset, is used for all methods to reduce the dimension of the 1011 wavenumber features to 60 features. Analyte hotspot occupation is expected to follow the Langmuir isotherm, which predicts a log-linear relationship between concentration and hotspot occupation. Due to this, all regressions are performed on a log-linear scale with the blank (water) dataset manually set to 10 fold smaller concentration than the next smallest concentration, 0.1 ng/mL.

Figure 7.3 depicts two models that use one spectra to infer one concentration, multilinear regression (MLR) and a multilayer perceptron (MLP). First, we consider MLR, which is used to give a baseline performance for the problem of inferring analyte concentration from Raman spectra. In Figure 7.3 a, one may observe a mean squared error of 0.85 and an r^2 of 0.88, indicating significant bias and variance in the predicts. Bias in the predictions is interpreted as deviation from the Langmuir isotherm, while variance indicates fluctuating Raman enhancement and analyte hotspot occupation from the underlying SERS sensing surface. Using a nonlinear model like MLP, deviations from Langmuir adsorption behavior can be captured upon model training. Further, this model more effectively uses the spectral information to infer the Raman enhancement of each spectra, reducing variance. A mean squared error of 0.74 and r^2 of 0.89 is observed for this model, whose predictions are plotted in Figure 7.3 b.

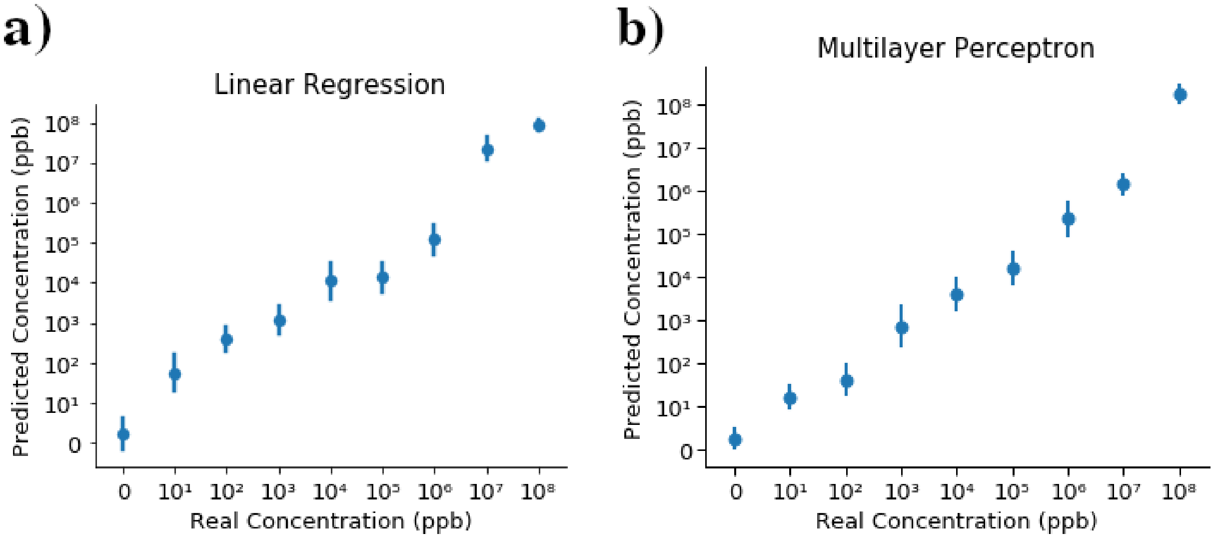

Figure 7.3. a) depicts the concentration predictions of the multilinear regression model for spectra obtained from the test dataset. b) depicts the concentration predictions of the MLP model for spectra obtained from the test dataset.

Due to the use of a sensing surface, SERS necessarily introduces signal fluctuations that emerge from spot to spot variations in Raman enhancement and hotspot volume. While the MLP seems to be able to infer these parameters to some degree from one spectra, significant improvements to predictions could be made by taking into account the distribution of these fluctuations while making predictions about concentration. Convolutional neural networks (CNNs) are well suited for identifying distributions within pixels of an image. We implement a CNN for concentration quantification, depicted in figure 7.4, by encoding five spectra as pixels within a one dimensional image, where each pixel has 60 channels to encode the pca components.

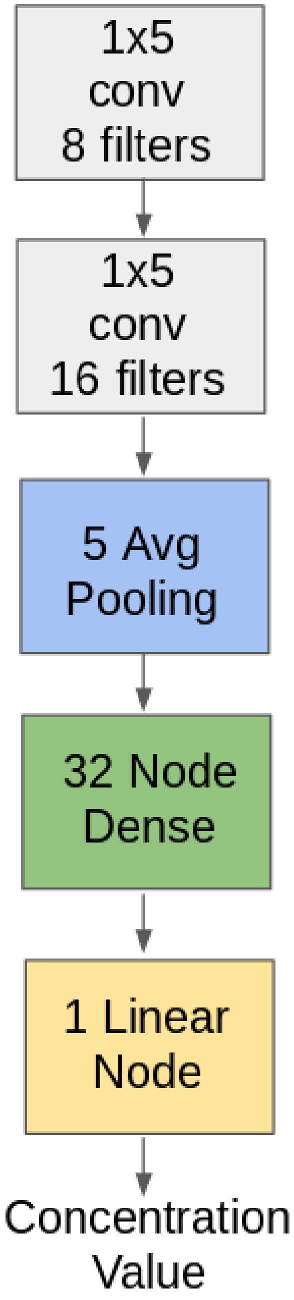

Figure 7.4. A schematic representation of the convolutional neural network model used in this work.

Figure 7.5 a depicts the concentration predictions of the 1D CNN. We find this model achieves a mean squared error of 0.29 and an r^2 of 0.95, significantly better than either the PLSR or MLP models. In the regression problem, interpolation is critical. Often, increased model complexity reduces variance, but at the expense of the ability of the model to interpolate. Yet we find, despite the complexity, the 1D CNN is still capable of making good interpolations. Figure 7.5 b is produced by removing two concentrations from the training dataset, 10^2 ng/mL, and 10^5 ng/mL, and plotting the predictions of the test dataset with the concentrations included. We find the model predicts these concentrations with reasonable accuracy.

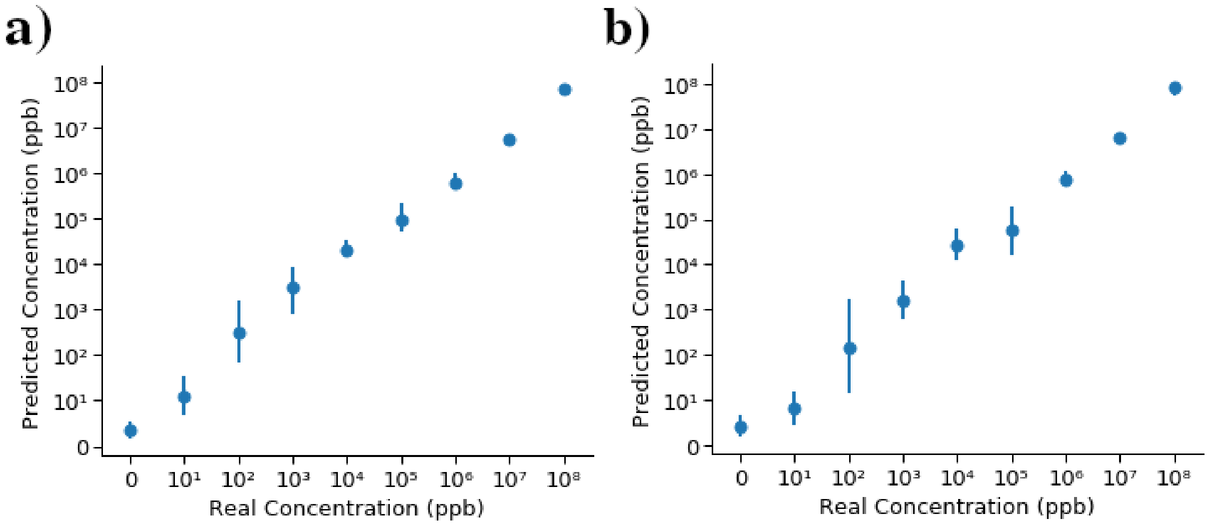

Figure 7.5. a) depicts the concentration predictions of the CNN model for spectra obtained from the test dataset. b) depicts the concentration predictions of the CNN model for spectra obtained from the test dataset with 10^2 and 10^5 ppb removed from the training dataset.

While challenging, interpreting neural network models is critical to ensuring that decisions are being made using features for which there is a causal relationship with the output. In Raman sensing, this amounts to ensuring that the network is using spectral information associated with pyocyanin vibrations. Here, we implement gradient assisted class activation mapping, gradCAM,[378] to interpret what spectral information is being used. Gradcam identifies network attention by determining how much the output changes with respect to the derivative of the input into the final fully connected layer. We implement this by omitting PCA dimensionality reduction and averaging the gradCAM values in each wavenumber channel over all of the spectra. Thus, we visualize the attention on a given wavenumber value. We find the network uses the 1353, 1602, and 1620 cm^{-1} to make its predictions, corresponding to prominent pyocyanin vibrational energies.

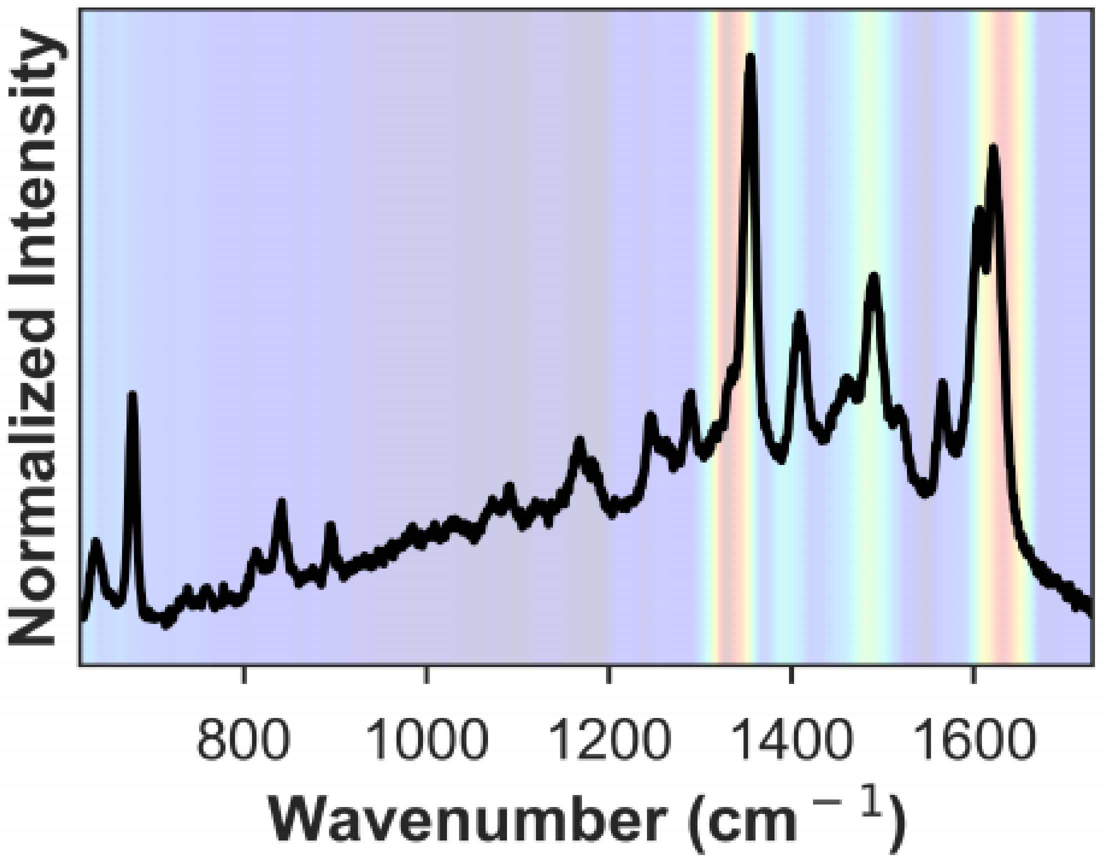

Figure 7.6 depicts a heat map of the spectral averaged gradCAM score of the wavenumber features. A Raman spectrum is plotted for reference.

7.3 Conclusion

In this work, we have demonstrated the use of 1D convolutional neural networks (CNNs) to learn the distribution of underlying, sensor based signal fluctuations to reduce variance in concentration predictions. We have shown this method yields an improvement in r^2 from 0.89 to 0.95, for the multilayer perceptron and 1D CNN, respectively. We have shown that the 1D CNN model uses analyte spectral information for its predictions, and the ability of the model to interpolate between concentration values. This method is broadly applicable to any sensing platform where it is unnecessary to predict concentration from each measurement, and where variance arises from latent distributions of intensity enhancement that emerge from spatial variations on the sensor surface.

7.4 Methods

7.4.1 Materials

Random copolymer poly(styrene-co-methyl methacrylate)-α-hydroxyl-ω-tempo moiety (PS-r-PMMA) (Mn = 7400, 59.6% PS) and diblock copolymer poly(styrene-b-methyl methacrylate) (PSb-PMMA) (Mn = 170-b-144 kg mol^{-1}) were purchased from Polymer Source, Inc. (Dorval, Canada). Si(001) wafers with a resistivity of 0.004 ohm-cm were purchased from Virginia Semiconductor (Frederickburg, VA, USA). Hydrofluoric acid (HF) was purchased from Fisher Scientific (Pittsburgh, PA, USA). 2-(N-morpholino)ethanesulfonic acid (MES) 0.1 M buffer, 1-ethyl-3-[3- (dimethylamino)propyl]carbodiimide hydrochloride (EDC), and N-hydroxy sulfosuccinimide (s-NHS) were purchased from Pierce (Rockford, IL, USA). Rhodamine 800, Methylene Blue, dimethyl sulfoxide (DMSO), ethylenediamine, toluene, ethanol, isopropyl alcohol (IPA), potassium carbonate, and 52-mesh Pt gauze foil were all purchased from Sigma-Aldrich (St. Louis, MO, USA). Nanopure deionized water (DI) (18.2 MΩ cm^{-1}) was obtained from a Milli-Q Millipore System.

7.4.2 2-Dimensional Physically Activated Chemical (2PAC) Assembly

Assembly is performed as in previous work. Random PS-b-PMMA block copolymer is spin-coated onto a HF-cleaned (the potential of HF to cause severe injury mandates extreme caution during usage), heavily doped Si wafer and annealed at 198 ^{0}C for three days. The sample is rinsed with toluene and lamella-forming PS-b-PMMA block copolymer is spin-coated onto the sample. PMMA regions are selectively functionalized with amine end groups by immersing the entire substrate in DMSO and then in ethylenediamine/DMSO solution (5% v/v), both for 5 min without rinsing between steps. The surface is rinsed with IPA and dried under nitrogen. Au nanosphere solution (0.1 mg/mL, 3 mL) is added to a 10 mL glass beaker. Freshly prepared s-NHS (20 mM) in a MES (0.1 M) buffer (35 μL) is added to the beaker and swirled. Next, freshly prepared EDC (8 mM) in a MES (0.1 M) buffer (35 μL) is added to the beaker and swirled. The solution then is brought to 60 °C. A 1 cm × 1 cm functionalized copolymer-coated Si substrate is placed into the solution vertically. One millimeter away from the substrate, a 1 cm × 1 cm Pt mesh is placed into the solution vertically. A dc regulated power supply is used to apply a voltage of 1.2 V for 10 min. This process is then repeated with the same substrate and fresh nanosphere solution as described above, but with 25 μL of EDC and s-NHS solution.

7.4.3 Characterization

SEM Images are collected with a Magellan XHR SEM (FEI). Raman spectroscopy measurements are conducted using a confocal Renishaw InVia micro Raman system with a laser excitation wavelength of 785 nm. All measurements are taken at 7.3 μW with exposure of 0.5 s and use a 60× water immersion objective with a 1.2 NA.

Chapter 8

Quantification of Analyte Concentrations in the Single Molecule Regime Using Convolutional Neural Networks

8.1 Introduction

Optical detection and spectroscopy of single molecules in the condensed phase has uncovered a wealth of understanding of molecular dynamics in physical, chemical, and biological systems since the first measurement in 1989 using frequency modulated laser spectroscopy by Moerner and Kador.[379] This led to an explosion of demonstrations of single molecule (SM) imaging and spectroscopy including SM imaging using near-field scanning optical microscopy in 1993,[380] confocal microscopy of a dye molecule at an air/polymer interface,[381] and diffusion of a rhodamine tagged phospholipid in membranes[382] in 1996. In 1997 SM surface enhanced Raman scattering (SERS) of dye molecules was first reported.[121,122] Unlike many imaging techniques, SM-SERS can directly sense non-emitters and non-resonant molecules with small Raman cross-sections,[142] i.e., when the enhancement to the signal is sufficiently high. SERS enhancement factors needed for SM SERS are estimated to be about 10^8 non-resonant molecules,[383] and as low as about 10^6 for resonant molecules.[374] Furthermore, SM-SERS has benefited from techniques used for SM imaging to examine the location of the analyte with respect to the nanoantenna and the resulting effect on signal using fitting of the point spread function for super localization.[384] SERS in general acquires rich spectral information from the "fingerprint" region in vibrational spectroscopy which can be used to uniquely identify molecules.[14] For this reason, SERS sensing may be highly multiplexed[366] and enable real time, in-line measurements for longitudinal monitoring of molecules.[143,385] One of the most alluring applications of SM-SERS is in sensing for medical diagnostics[293] where concentrations of biomarkers are often ultralow[386] and present in the background requiring quantification.

The following challenges are associated with SM-SERS sensing and quantification in aqueous media: 1) generation of hot spots providing sufficient signal enhancements to enable SM detection. 2) SM signal intensity must compete with background signal from ligands and other molecules adsorbed to hotspots which may greatly outnumber analyte molecules.[272] Additionally, with in-line measurements, water interacts with analyte molecules and the surface.[14] 3) Spectral positions and intensities of vibrations are not as reliable as in bulk SERS due to molecular orientation,[387] vibronic coupling,[388] and other factors[389] which are averaged out in bulk SERS. 4) Inconsistent signal intensities due to analyte diffusing in and out of hotspots which have rapid spatial variations of field enhancement.[390] These challenges are compounded when one's goal is to use SM-SERS measurements to quantify analyte concentration below nM values needed for biological and water quality applications.

The complexity in quantitative, predictive SERS sensing even at □M concentration, well above the single molecule limit, has led to the widespread adoption of nonlinear machine learning methods such as support vector machines,[177] artificial neural networks.[370] These techniques greatly improve concentration predictions because they are robust to non-analyte signal and can handle deviations from the Langmuir isotherm, but they map one spectrum to one concentration value. SM SERS is characterized by discrete jumps in analyte number observed in any given measurement, so concentration will not be correctly inferred from one measurement. This observation has led quantification methods based on mapping the distribution of analyte observations to concentration including the Brule methods[143] and the digital method.[349] Yet these methods are not robust to complex backgrounds from non-analyte molecules which complicate sensing at ultralow concentrations in solution needed for real time measurements of, e.g., biomarkers in biological systems or contaminants in water.

Here we approach this problem by recasting the analyte signal variations and non-analyte fluctuations as distributions of latent variables that must be learned for correct concentration predictions. A robust CNN model is used to address the challenge of mapping analyte signal distributions to concentrations and separating from fluctuating non-analyte signals. In recent years, convolutional neural networks (CNN) have been demonstrated to be exceptional at determining complex distributions, they are widely used for the problem of image recognition where the distribution of pixel values that correspond to various objects,[391] and have recently been used in SERS.[188]

Yet, machine learning necessitates big, high quality datasets. This is hindered by one of the longstanding challenges in SERS, lack of point to point reproducibility in SERS intensity and variability from sample to sample. While nanofabrication – specifically chemical self-assembly – has advanced to the point where sub-nanometer gaps can be reliably produced,[222,368] it is difficult to obtain high coverage of discrete clusters needed for acquiring large data sets of SM sensing events. While top down methods are able to achieve high density over large areas, nanogap spacing is typically larger than 1 nm.[235,367] To overcome these challenges, we use our recently demonstrated chemical self-assembly method – 2-dimensional physically activated chemical self-assembly (2PAC) – capable of obtaining a SERS EF exceeding 10^9 with just 10% relative standard deviation over a 1 mm^2 area.[272] With these sensors, spectra are acquired across the SM concentration regime and are bundled into 8x8 pixel maps (with each pixel being a SERS spectra), and used to train a CNN model. The model's predicted concentrations have an average r^2 value of 0.958 over 6 orders of magnitude as determined by k-folds cross validation. Further, the demonstrated limit of blank (LOB) is 1 fM for Rhodamine 800, with a limit of quantification (LOQ) of 10 fM. The LOQ observed here is comparable to state-of-the-art methods involving recognition elements,[392] and is significantly lower than has been achieved by the same sensor without the use of CNNs.[385] Generalization of the model to other molecules using transfer learning is demonstrated with methylene blue. Transfer learning also significantly reduces the amount of training data necessary, with good results being observed with as few as 50 8x8 pixel maps, requiring just 5.3 minutes of total laser exposure time per concentration.

8.2 Discussion

In order to understand how CNN enables quantification in the SM regime let's first examine SERS quantification in a typical concentration regime. The central problem in SERS is the inverse problem of determining analyte concentration, C, from SERS intensity of vibrational modes, denoted by I_P, i.e., the p-dimensional spectral feature space acquired in a Raman measurement. Two classes of parameters, concentration independent and concentration dependent, determine I_P. First: those that do not vary with concentration include Raman cross section, location and orientation of molecule in hotspot, hotspot volume, and electric field enhancement. Second: clearly the number of molecules adsorbed in the hotspots, N_H, is concentration dependent and is often related to C by some molecular adsorption isotherm such as the Langmuir adsorption isotherm or the Freundlich isotherm. In SERS analysis, one often assumes N_H is large such that it will vary continuously with concentration. The adsorption isotherms are log-linear, so it is common practice to log transform the concentration values and absorb the adsorption rate constant into the other concentration-independent factors that determine I_P, and fit the data with a multilinear regression from I_P to C.

The situation is very different for SM SERS sensing. C, and consequently N_H are extremely low, so at any moment N_H can be 0 in the laser spot size. One may alternately lengthen the acquisition time sufficiently to enable N_H to converge to its expected value, yet, I_P will average all the signals in the acquisition time and the analyte signal may not be discernible above the background. Thus it is necessary to acquire short measurements where I_P from molecules of interest can be differentiated from I_P of the blank signal. It is clear then that in the SM regime it is not possible to map individual SERS measurements to concentration as most acquisitions will have N_H of zero. Instead, the sensing problem for SM SERS amounts to using multiple I_P measurements to determine the distribution of I_P that is reflective of N_H having finite or zero value in a given measurement and mapping this onto C.

Our approach to determining the distribution of I_P begins with bundling Raman spectra into 8×8 pixel maps, where each pixel is a Raman spectrum. These maps are randomly split into test and train groups for k-fold cross validation. Each spectrum is reduced to 32 independent signals using non-negative matrix factorization (NMF) that is trained on the training dataset. During the training period, each map is given a log transformed C label. At these low concentrations, these 64 spectra exhibit a sparse distribution of N_H which the regression model uses to learn the mapping of I_p to C.

8.3 Experimental System

Integral to any machine learning approach are access to large and high quality datasets. Using 2-dimensional physically activated chemical (2PAC) self-assembly, we have shown we can achieve enhancement factors of 10^9 with less than 10% RSD over 1 cm^2.[272] Since 2PAC is described in detail elsewhere,[272] we only briefly describe here. Gold nanoparticle seeds are assembled on a poly(styrene-b-poly(methyl methacrylate) copolymer coated Si electrode by electrophoresis. 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) activates a crosslinking reaction between the carboxylic acid functionalized nanoparticles and amine functionalized poly(methyl methacrylate) domains. The nanoparticle seed perturbs the electrical double layers of the electrode and leading to electrohydrodynamic (EHD) flow. This lateral, attractive flow entrains nearby particles, forming close-packed nanoparticle assemblies. The close proximity of the nanoparticles during 2PAC assembly enables a carbodiimide-mediated crosslinking between carboxylic acid groups on adjacent nanoparticles forming an anhydride bond with a gap spacings of ~0.9 nm. Figure 8.1a is a schematic of a resulting dimer formed from 2PAC assembly with an anhydride bond linking the Au nanoparticles. A scanning electron microscopy (SEM) image of a 2PAC assembled SERS surface is depicted in Figure 8.1b.

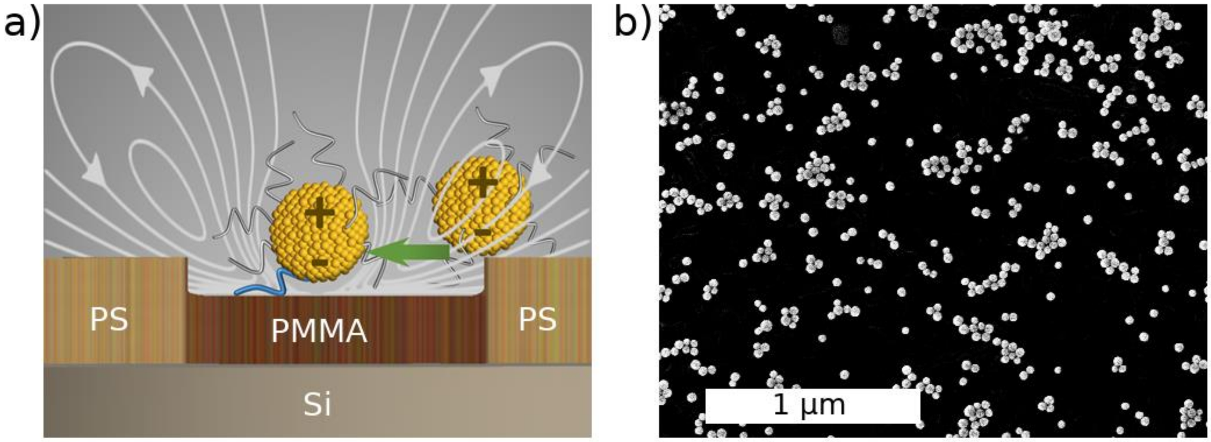

Figure 8.1. (a) depicts a schematic of 2PAC self-assembly, where EHD flow enables a chemical crosslinking reaction to occur between gold nanoparticles and results in the formation of 2-dimentional close-packed assemblies. (b) SEM micrograph of as-assembled gold nanoparticle assemblies.

Large area SERS maps are acquired from these 2PAC assembled surfaces. Unless otherwise stated, SERS spectroscopy is performed on the surfaces by acquiring maps across an area of 100 µm □ 120 µm with 2 µm spacing between acquisition spots. Spectra are acquired for 0.1 s at a laser power (wavelength) of 760 µW (785 nm) with a 1.2 NA water immersion objective.

Here, nonnegative matrix factorization (NMF) is used to visualize areas with spectral features associated with Rhodamine 800 (R800) scattering events on sample surfaces. NMF is chosen over PCA because SERS spectra satisfy NMF's assumption of all positive signals, resulting in improved differentiation of spectral features that emerge from analyte and non-analyte sources,[355] which is necessary in SM SERS where non-analyte spectra will have more variation than analyte spectra. NMF obviates the need to perform coordinate transformations after applying PCA that are common in SM SERS.[166] An extracted NMF component vector associated with R800 is acquired from a 100 nM R800 dataset and depicted in Figure 8.2 a. Here, in the many molecule regime, the R800 NMF signal is free of contamination from non-analyte signal. Thus, we use this extracted loading vector to assess R800 score of spectra acquired at trace concentrations (1 nM and below). In this work, we seek to map groups of 64 Raman spectra to concentration values. We visualize the average R800 score of these 64 spectra groups as a function of concentration in Figure 8.2 b. The violin plots depict an upward trend of these average scores across the trace concentration regime. Yet the complex background arising from non-analyte signals causes large variance in the R800 loading. These results are similar to those obtained by Albequerque et al.[349] This complexity motivates the deep learning approach taken below.

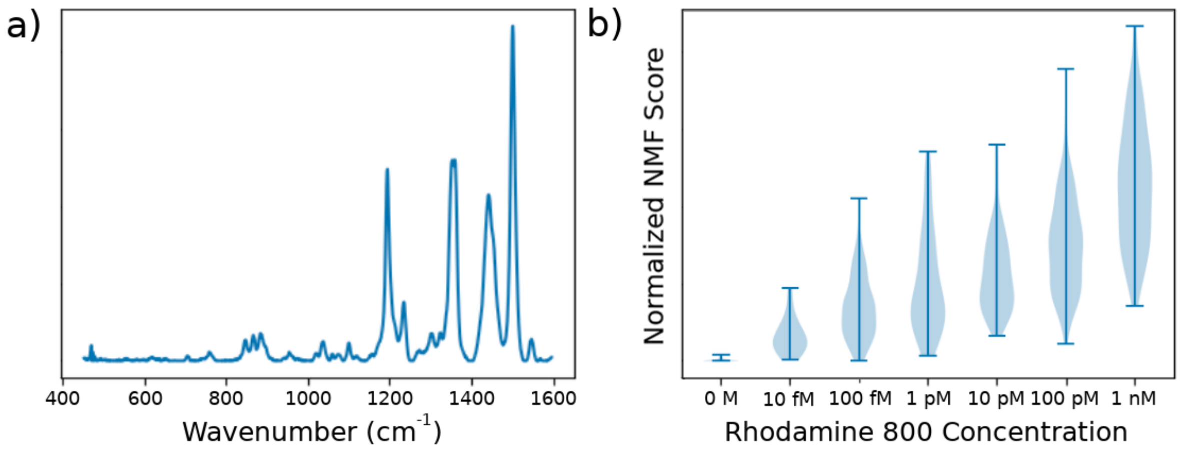

Figure 8.2. (a) NMF component of Rhodamine 800 acquired from 100 nM SERS measurements. (b) Normalized NMF scores averaged from 64 spectra bins across the SM concentration regime.

8.4 Convolutional Neural Network (CNN) for Single Molecule SERS Quantification

The CNN architecture used in this work is depicted schematically in Figure 8.3 a. The model's predictions of the CNN on a test dataset are shown in Figure 8.3 b, where we observe an r^2 of 0.95 over 5 orders of magnitude. We perform a 5-fold k-folds cross validation to validate the model, observing a cross validation mean squared error (MSE) loss of 0.111+/-0.029, and r^2 of 0.958+/-0.012. The limit of blank (LOB) is determined to be 1 +/- 1.1 fM and the limit of quantification (LOQ) is approximately 10 fM. These results demonstrate that the CNN model is robust to complex background, which resulted in difficult to discriminate NMF scores. Figure 8.3 c depicts the MSE loss as a function of training epoch. One may observe the model briefly finding local minima corresponding to the degenerate solution of predicting all inputs as having identical concentrations, depicted in the inset of Figure 8.3 c.

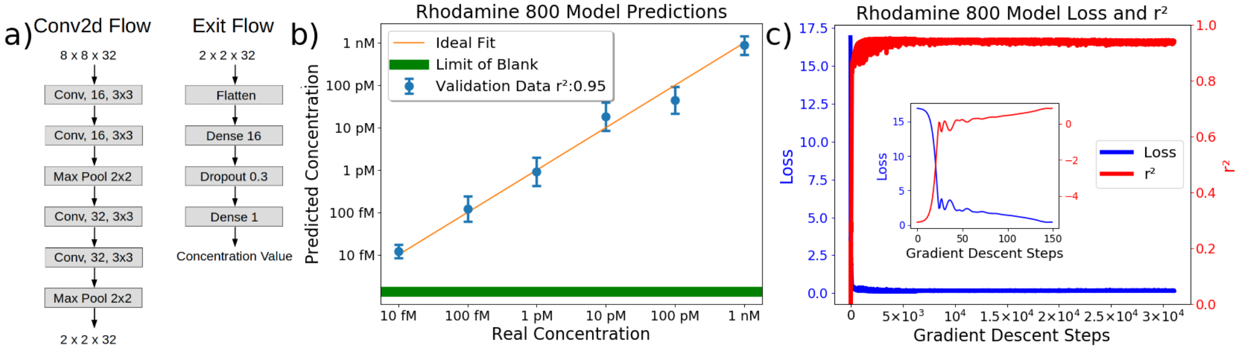

Figure 8.3. (a) Schematic of the CNN model used to determine concentration of analyte molecules. (b) Representative Rhodamine 800 concentration predicted by the CNN model on a test dataset acquired during k-fold cross validation. (c) Model test loss and r^2 over gradient descent steps of the model depicted in Figure 8.3 b. Inset: test loss and r^2 depicting local minima of loss when the model predicts similar concentrations for all inputs.

8.5 Transfer Learning for New Analyte Molecule

Here, we seek to test the generalizability of the R800 CNN model to another analyte, Methylene Blue (MB). We do this with a transfer learning approach, that is, pretrained weights from the R800 CNN model are used in the training of a MB CNN model. The weights of the third and fourth convolutional layers are frozen throughout training. New NMF components are found using the MB training dataset. First, we evaluate a MB CNN model trained from scratch with the same architecture as the R800 CNN model. The regression results are depicted in Figure 8.4 a, where a cross validation MSE loss of 0.288 +/- 0.085, r^2 of 0.932 +/- 0.02, and LOB of 1.2 fM +/- 1 fM. Here we observe increased variance of the predictions compared to those in R800. This is due to greater overlap between MB Raman vibrational modes and non-analyte modes arising from the SERS surface. The loss, LOB and r^2 of this model, trained on a full MB dataset, is used as the lower (upper) limit of transfer learned model performance depicted in Figure 9.4 b as horizontal bars. In the transfer learning experiments we greatly restrict the number of training examples observed by the model to be between 1 and 100. One may observe near optimal model performance with as few as 100 training examples. Models trained without transfer learning from the R800 model with 100 or less examples are unable to escape the degenerate local minimum described previously. This effect also leads to large standard deviations in the cross validation values of metrics at 1 and 3 examples. These results demonstrate that the R800 CNN model is broadly applicable to SM concentration regression.

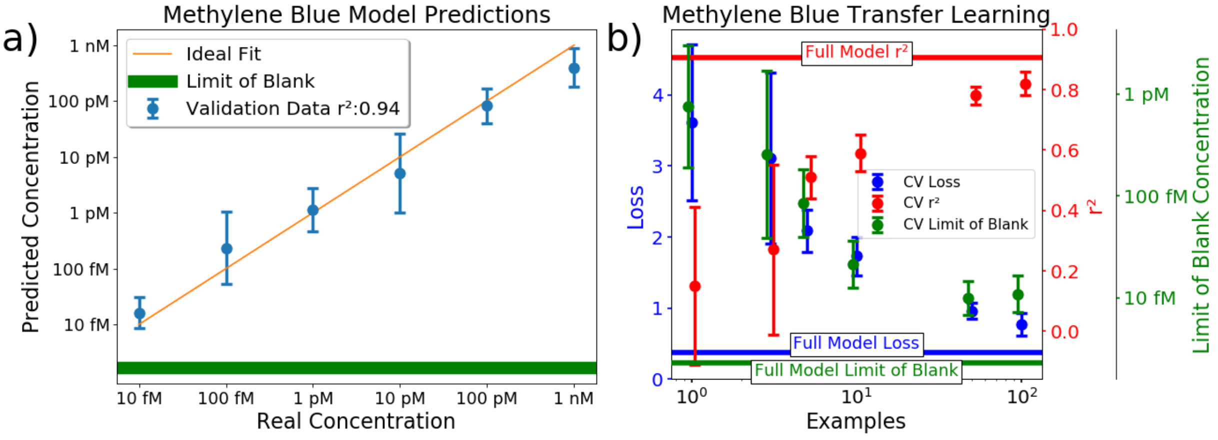

Figure 8.4. (a) Representative Methylene Blue concentration predicted by the CNN model on a test dataset acquired during k-fold cross validation. (b) Cross validation loss, r^2, and LOB of transfer learned Methylene Blue model as a function of number of unique input maps. The model is identical to Figure 8.3 a, but with weights acquired from the Rhodamine 800 model depicted in Figure 8.3 b, with the third and fourth convolutional layers frozen. Horizontal bars represent the metric performance depicted in Figure 8.4 a. Number of examples for r^2 and LOB are offset by 5% for visual clarity.

8.6 Conclusion

In this work, we have proposed a convolutional neural network (CNN) model to perform concentration regressions on 8x8 maps of surface enhanced Raman scattering (SERS) spectra. This proof of concept shows that the CNN model dramatically simplifies the implementation of concentration regressions in the single molecule (SM) concentration regime. Using Rhodamine 800 , we have demonstrated the the CNN model makes excellent concentration predictions, even with complex background signals. Specifically, a k-fold cross validated r^2 of 0.958 is achieved with quantification down to 10 fM. Further, we have shown that the model is generalizable to new analyte (methylene blue) by the implementation of transfer learning. This strategy results in good predictions with as few as 50 training examples. We believe that CNN models applied to maps of spectra can be generalized both to other SERS sensor platforms as well as other chemical analysis techniques capable of single molecule detection.

8.7 Methods

8.7.1 Materials

Random copolymer poly(styrene-co-methyl methacrylate)-α-hydroxyl-ω-tempo moiety (PS-r-PMMA) (Mn = 7400, 59.6% PS) and diblock copolymer poly(styrene-b-methyl methacrylate) (PSb-PMMA) (Mn = 170-b-144 kg mol^{-1}) were purchased from Polymer Source, Inc. (Dorval, Canada). Si(001) wafers with a resistivity of 0.004 ohm-cm were purchased from Virginia Semiconductor (Frederickburg, VA, USA). Hydrofluoric acid (HF) was purchased from Fisher Scientific (Pittsburgh, PA, USA). 2-(N-morpholino)ethanesulfonic acid (MES) 0.1 M buffer, 1-ethyl-3-[3- (dimethylamino)propyl]carbodiimide hydrochloride (EDC), and N-hydroxy sulfosuccinimide (s-NHS) were purchased from Pierce (Rockford, IL, USA). Rhodamine 800, Methylene Blue, dimethyl sulfoxide (DMSO), ethylenediamine, toluene, ethanol, isopropyl alcohol (IPA), potassium carbonate, and 52-mesh Pt gauze foil were all purchased from Sigma-Aldrich (St. Louis, MO, USA). Nanopure deionized water (DI) (18.2 MΩ cm^{-1}) was obtained from a Milli-Q Millipore System.

8.7.2 2-Dimensional Physically Activated Chemical (2PAC) Assembly

Assembly is performed as in previous work.[272,385] Random PS-b-PMMA block copolymer is spin-coated onto a HF-cleaned heavily doped Si wafer (the potential of HF to cause severe injury mandates extreme caution during usage), and annealed at 170^0 C for three days. The sample is rinsed with toluene and lamella-forming PS-b-PMMA block copolymer is spin-coated onto the sample and annealed. PMMA regions are selectively functionalized with amine end groups by immersing the entire substrate in DMSO and then in ethylenediamine/DMSO solution (5% v/v), both for 5 min without rinsing between steps. The surface is rinsed with IPA and dried under nitrogen. Au nanosphere solution (0.1 mg/mL, 3 mL) is added to a 10 mL glass beaker. Freshly prepared s-NHS (20 mM) in a MES (0.1 M) buffer (35 µL) is added to the beaker and swirled. Next, freshly prepared EDC (8 mM) in a MES (0.1 M) buffer (35 µL) is added to the beaker and swirled. The solution then is brought to 60 °C. A 1 cm × 1 cm functionalized copolymer-coated Si substrate is placed into the solution vertically. One millimeter away from the substrate, a 1 cm × 1 cm Pt mesh is placed into the solution vertically. A DC regulated power supply is used to apply a voltage of 1.2 V for 10 min. This process is then repeated with the same substrate and fresh nanosphere solution as described above, but with 25 µL of EDC and s-NHS solution.

8.7.3 Characterization

SEM Images are collected with a Magellan XHR SEM (FEI). Raman spectroscopy measurements are conducted using a confocal Renishaw InVia micro Raman system with a laser excitation wavelength of 785 nm. All measurements are taken at 760 µW with exposure of 0.1 s and use a 60× water immersion objective with a 1.2 NA.

8.7.4 Data Analysis and Models

Each spectrum acquired is Savitsky-Golay smoothed (implemented in python with scikit-learn), background subtracted (with asymetric least squares), and min-max scaled to have a minimum value of 0 and a maximum value of 1 (implemented in python with scikit-learn). At each value of concentration, the preprocessed spectra from are randomly shuffled together and reshaped into 8 x 8 maps of 1011 wavenumber spectra. Next, data augmentation[393] is performed to increase the amount of training data fed into the model. In general, data augmentation applies a transformation to the data that is invariant. Here, we reshuffle data and produce new 8×8×1011 maps, in this way the data is expanded by a factor of 32.

Before input into the model, the dimension of spectra are reduced from 1011 to 32 using non-negative matrix factorization (NMF), implemented in python with scikit-learn. The convolutional neural network (CNN) architecture is depicted in Figure 8.3 a, and implemented in python with keras. The model begins with an entry flow of convolutional layers: two 2×2 convolutional layers with 16 filters and relu activations followed by a 2×2 max pooling layer, yielding 4×4×16 maps. This is followed by two 2×2 convolutional layers with 32 filters and relu activations and a followed by a 2x2 max pooling layer, reducing the dimension to 2x2x32. This is followed by an exit flow where the output is flattened yielding 128 dimensional vectors that are fed into a fully connected layer with 16 nodes with a relu activation and dropout regularization set to 0.3. Finally this is sent to a single linear node that executes the regression.

Transfer learning, depicted in Figure 8.4 b, is implemented in python with keras. A Rhodamine 800 model's pretrained weights are used with the third and fourth convolutional layers frozen. New NMF components are used for dimensional reduction that are identified from the new analyte's training dataset. Each unique input, or example, used to train the model in Figure 8.4 b is augmented 32 fold with the strategy described above.

Chapter 9

Semi-Supervised and Unsupervised Rapid Antimicrobial Susceptibility Testing with SERS

9.1 Introduction

Every year 700,000 people die of antimicrobial resistant (AMR) bacteria infections.[394] Today a full third of antibiotics prescribed are to treat bacteria that are resistant to those drugs.[395] The problem is especially bad for the treatment of urinary tract infections – one of the most prevalent diseases, accounting for 17% of all hospital acquired infections[157] – where more than half of bacteria tested were AMR.[396] AMR is a growing problem, and by 2050 it is expected to cause 10 million deaths and a 3.8% reduction of the total world gross domestic product.[397] While AMR is a multifaceted problem that will require many systemic changes to healthcare, the 2016 Review on Antimicrobial Resistance lists rapid diagnostics to reduce the unnecessary use of antimicrobials as a key intervention for the reduction of AMR. [398]

To this end, antimicrobial susceptibility testing (AST) is an indispensable tool. Traditionally, AST has involved acquiring bacteria samples from patients, culturing the samples, and using disk-diffusion or broth microdilution studies (among other techniques) to identify a suitable antibiotic for treatment.[399] Despite challenges associated with long culture times, AST is widely used and has been shown to reduce hospital stays and save money.[400] AST has also been proposed to aid the development of new antibiotics,[401,402] and reduce the costs of clinical trials.[403] Point-of-care genomic AST has been proposed to obviate the need for culturing and has shown promising results, but phenotypic (growth based) AST is considered the gold standard for AST and genomic AST still requires phenotypic validation.[404–406]

Rapid phenotypic AST is possible through metabolomic profiling of bacterial response due to antibiotic treatment through methods such as surface enhanced Raman scattering (SERS) spectroscopy.[407] Rapid AST seeks to bridge the gap between traditional AST, which takes between 24 and 72 hours,[408] and AST times that are necessary for optimal patient outcomes, 30 minutes to 1 hour.[409] SERS is an obvious choice for rapid AST due the rich vibrational information that Raman scattering produces and can be used to identify small molecules or differentiate complex mixtures of small molecules with a label-free approach.[14] Additionally, with sensor surfaces capable of single molecule detection and proper statistical analysis tools, SERS can detect molecules in concentrations as low as 1 part per trillion,[410] enabling its application for a host of metabolomics problems. For these reasons SERS has been extensively used in bacteria detection,[358,411,412] and the metabolites identified with SERS[151] have been validated with mass spectroscopy.[413] SERS can also be used for in situ bacteria detection,[414] and can monitor bacteria biofilm formation longitudinally to detect infection before some AMR mechanisms are generated.[385] This work culminated (for our purposes) in 2016 when liu et al[407] convincingly demonstrated AST and identified the minimum inhibitory concentration (MIC) of antibiotics with SERS.

We believe that for SERS to enter general use in healthcare, a machine learning (ML) based approach will be necessary. ML, especially deep learning, is emerging as an important force to revolutionize healthcare, with ML approaches now surpassing the performance of doctors in computer vision tasks like diagnosing skin cancer[415] and breast cancer.[416] Deep ML approaches are not limited to computer vision and have applications in surgery,[417,418] predicting diagnoses from medical records,[419,420] and in interpreting electroencephalograms to predict seizures[421] and parkinson's disease.[422] The fates of SERS and ML have been entangled since 2006 when principal component analysis was used to convincingly demonstrate single molecule SERS for the first time by Le Ru et al.[423] Since then great progress has been made in applying sophisticated ML techniques to solve SERS problems. These include: the fully-connected artificial neural network for analyte concentration regression,[370] DNA classification,[424] and cancer detection,[425] the convolutional neural network (CNN) for classification of metabolite signals from cells,[188] the support vector machine for classification of drug use from urine,[177] and the genetic algorithm for cancer diagnosis.[426]

From this discussion it seems like an ML based SERS approach to AST would be easily implemented due to the relative ease of collecting large SERS datasets of bacterial metabolites. Yet one key challenge faces this approach that was alluded to earlier: rapid AST approaches must be validated with traditional AST approaches to be accepted into practical use.[427,428] From an ML viewpoint this means that every AMR status label that corresponds to a SERS spectra will require a 24-72 hour culturing process. Consider that deep ML algorithms that tackle healthcare problems require tens of thousands of labeled examples.[429] This challenge places an enormous barrier for the development of SERS AST in a clinical setting.

In this work, we seek to overcome the challenges of label generation through the use of few shot and semi-supervised ML, which have already been used in healthcare.[430–433] Here, we use SERS to examine cell lysate of *Psuedomonas aeruginosa* (PA) that has been treated with antibiotics that it is resistant or susceptible to. We demonstrate the use of a deep generative model, the variational autoencoder[434] (VAE) – which has already revolutionized the field of drug discovery[435] – as an ideal method for SERS based semi-supervised ML. We show that the VAE produces good clustering behavior of SERS spectra from different antibiotic treatments into its latent space. Further we show that the ability to continuously reconstruct spectra across the latent space enables facile interpretation of the decision making process of the deep learning algorithm. Encoding spectra into the latent space of a VAE enables greatly improved performance of AST classification algorithms, with as SVM trained on encoded spectra achieving over 80% accuracy on a test dataset with just 10 example spectra. Finally, we expand the latent space of a VAE by including spectra acquired from 63 different mixtures of bacteria metabolites. This results in superior clustering of SERS AST spectra, enabling an unsupervised Bayesian Gaussian Mixture model to achieve 99.3% accuracy on a test AST dataset, even higher than a deep CNN transfer learning based approach for fewer than 10 example spectra. This work is an important proof of concept, demonstrating that SERS analysis of AST data is amenable to a semi-supervised and even an unsupervised approach, making SERS AST a practical method for clinical AST.

9.2 Experimental System

Above, we discussed how growth based (phenotypic) antimicrobial susceptibility testing (AST) is considered the gold standard. In order to reduce the time required for phenotypic AST we, and many other groups, seek to use a metabolomics approach over the direct measurement of cell viability. We pay an important price to reduce AST time in this way: metabolomics approaches introduce an enormous parameter space. For example, the e-coli metabolome contains over 2600 unique small molecules (metabolites).[101] AST analysis is even more complex in a clinical setting where samples extracted from patients are typically mixtures of many bacterial species. Therefore machine learning based approach is essential in order to build complex enough models to capture these variations.

We then need a tool capable of measuring the metabolomic profile of a bacterial system rapidly enough to be compatible with a machine learning approach. To this end, surface enhanced Raman scattering (SERS) spectroscopy is an ideal method. Sample preparation is minimal, where cellular supernatant or lysate can be directly placed on a SERS sensor with no separation processes, and individual measurements can be acquired in under one second with sufficiently powerful SERS sensors. The primary challenge with SERS is just that one must manufacture SERS sensors with sufficiently large and uniform signal enhancements. To this end, our group has recently proposed a 2-dimensional physically activated self-assembly method[272] (2PAC) that produces SERS sensors with excellent signal enhancements and uniformity. The sensor fabrication procedure is described in the methods section. We have used this technique to drive forward SERS-metabolomics-machine learning methods,[273] and we are now turning to AST.

SERS AST has already been demonstrated,[407] so the focus of this work is to demonstrate that SERS spectra are compatible with semi-supervised and unsupervised deep machine learning techniques that will greatly reduce the costly overnight cell culture that is necessary to acquire labels (eg antibiotic resistant or susceptible). Our proof of concept experiment is to differentiate cell lysate of *Psuedomonas aeruginosa* (PA) acquired from cultures exposed to three antibiotic conditions: 1) untreated with antibiotics (-), 2) treated with carbenicillin (+C), which PA is susceptible to, and 3) treated with rifampicin (+R), which PA is resistant to. We perform these treatments at, 50 μg/mL and 400 μg/mL of carbenicillin and rifampicin, respectively, concentrations at the minimum inhibitory concentration (MIC), where PA growth is reduced, but not so large that the colonies are unable to flourish. MIC curves depicting cell density after 24 hours carbenicillin and rifampicin are depicted in Figure 9.8.

Averages of 400 SERS spectra acquired from PA lysate at 30 minutes and 2 hours for the -, +C, and +R conditions are plotted in Figure 9.1, a, b, and c, respectively. These spectra are averaged after preprocessing which involves smoothing, background subtraction, and minimum-maximum scaling. SERS collection and spectra preprocessing are elaborated in the methods section. Cell growth, as measured by the optical density at 600 nm (OD600) from these experiments are depicted in Figure 9.1 d, e, and f for -, +C, and +R, respectively. Cell cultures are prepared to have 0.5 OD before treatment. One may observe increased cell density even with the +C treatment, which PA is susceptible to, due to the fact that we use relatively low antibiotic concentrations. Significant variations in the rate of cell growth are observed, which is the primary motivation of using metabolomics based AST over cell viability AST, as this is expected at such short time periods.

Prior work in SERS AST has typically involved directly measuring bacterial cell cultures and monitoring the 730 cm^{-1} vibrational band associated with cell walls.[407,436,437] This method has been useful because it enables facile determination of cell proliferation, but SERS measurements that rely on individual Raman bands train models with much higher bias and variance than models that use the full spectral information. As we use a deep machine learning approach, we use cell lysate to maximize the diversity in metabolites and consider full spectra. In order to holistically assess entire spectra we use unsupervised learning approaches, which are not given informed of the class (the antibiotic treatment condition) or each spectra in order to avoid overfitting to noise, as supervised learning techniques using a parameter space as large as the SERS fingerprint vibrational window are susceptible to do.

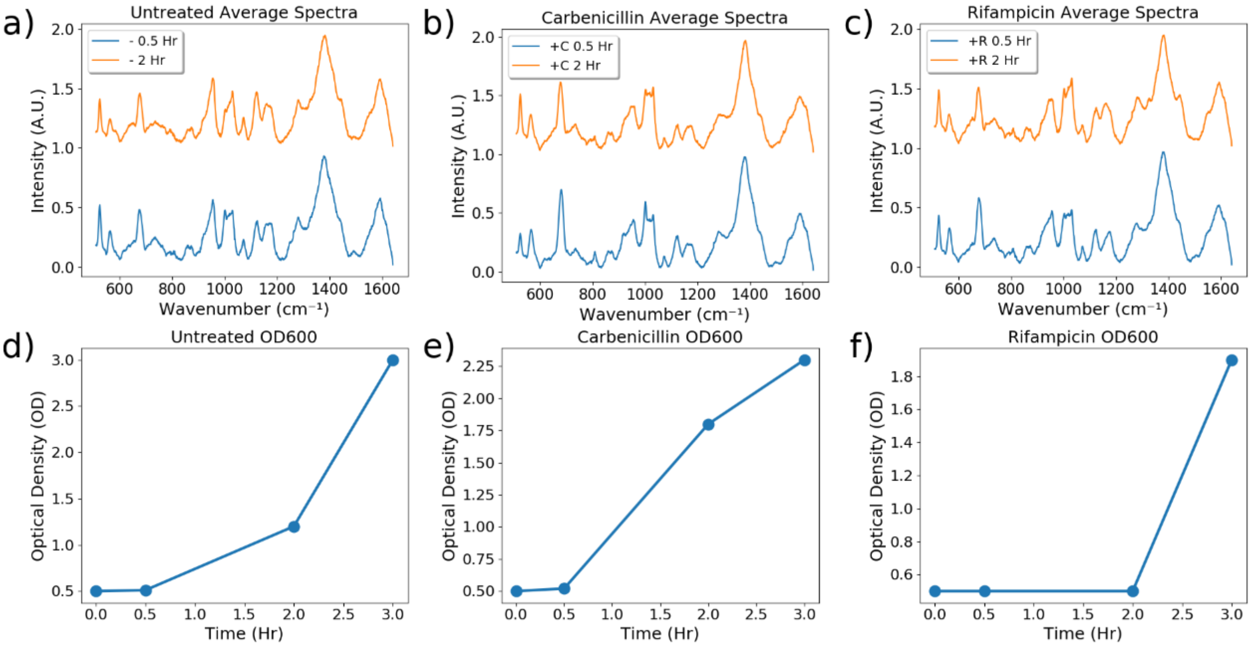

Figure 9.1: a, b, c average SERS spectra of 30 minutes and 2 hrs post treatment of control, carbenicillin treatment, and rifampicin treatment, respectively. Spectra are scaled between values of 0 and 1 and 2 hour SERS spectra are offset by 1 for visual clarity. d, e, f the optical density acquired at 600 nm of the *Psuedomonas aeruginosa* cultures at various times for the untreated, carbenicillin, and rifampacin treated cases, respectively.

For a first pass visualization of SERS spectra we use t-distributed stochastic neighbor embedding (t-SNE).[438] Figure 9.2 a depicts the t-SNE of the SERS AST dataset described above. t-SNE is a relatively new unsupervised data visualization technique that has ubiquitous use within the machine learning community, due to the clarity of the clustering that it achieves. t-SNE can be thought of as attempting to preserve the distances between spectra in their native 1011-dimensional space in the 2-dimensional representation shown in Figure 9.2 a. One may observe excellent clustering of the different classes of antibiotic treatments, indicating significantly different metabolomic profiles. Some of these differences can be observed in Figure 9.1 a, b, c, where vibrations between 1000 cm^{-1} and 1200 cm^{-1} are particularly different between the classes. Yet, despite the utility of t-SNE as a visualization tool, it is probabilistic in nature and its embeddings cannot be used to develop a predictive model, and does not easily lend itself to interpretation. For this purpose we use non-negative matrix factorization (NMF).[355]

Figure 9.2 b depicts the first NMF component extracted from the SERS AST dataset with a total of 3 components, and Figure 9.2 c depicts the score of SERS spectra in the different treatment classes with respect to this component. NMF is a powerful dimensional reduction tool for SERS due to the fact that all signals are positive in SERS. The non-negativity constraint enables superior isolation of independent signals compared to the traditional principal component analysis (PCA). Like PCA, we can interpret the NMF components as the isolated signal due to our experimental condition. From Figure 9.2 c one can observe a clear correlation between PA health and NMF score, with score rising from the 30 minute untreated culture (- 0.5 Hr) to the 2 hour untreated culture (- 2 Hr), score declining from the 30 minute carbenicllin treatment (0.5 Hr +C) to the 2 hour carbenicillin treatment (2 Hr +C), and little difference between the score of the 30 minute rifampicin treatment (0.5 Hr +R) to the 2 hour rifampicin treatment (2 Hr +R). While powerful, NMF clearly loses much of the separation between treatment classes observed in the t-SNE visualization, and we lose the ability to distinguish the 0.5 Hr +R from the 2 Hr +R.

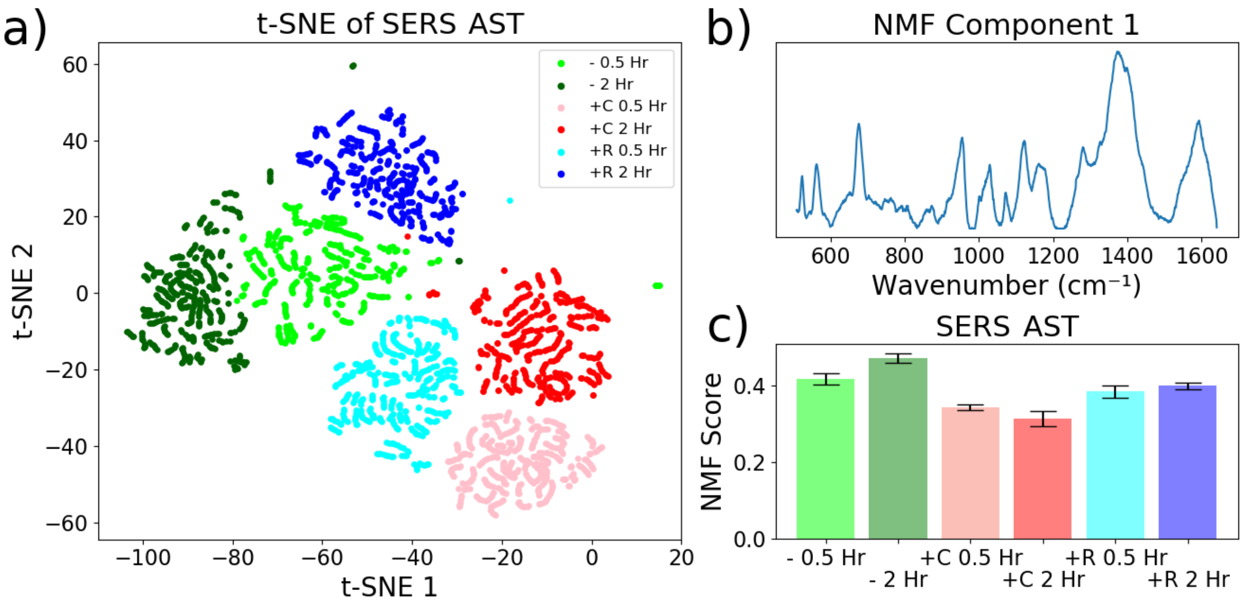

Figure 9.2: a) t-SNE of the various antibiotic treatments, untreated for 30 minutes (- 0.5 Hr), and 2 hours (- 0.5 Hr), treated with carbenicillin for 30 minutes (+C 0.5 Hr), and 2 hours (+C 2 Hr), and treated with rifampicin for 30 min (+R 0.5 Hr), and 2 hours (+R 2 Hr). b) First NMF component extracted from training dataset c) First NMF score of antibiotic treatment test dataset.

9.3 Semi-Supervised Learning for SERS AST

While deep machine learning models are valuable for capturing the complexity of SERS spectra of bacterial metabolites, a trade-off exists between the interpretability of these models and their performance. When dealing with the health of patients, which is likely in dire conditions if AST is being considered, interpretable models are essential so that experts can validate the model predictions. While interpretability-performance trade-offs are fundamental to machine learning, deep generative models – like the variational autoencoder (VAE) – give the user insight into the model's decision making.

Figure 9.3 depicts a scheme for using a VAE for SERS analysis. As is typical of unsupervised learning methods, the VAE works by encoding a high dimensional data point (here a SERS spectra) into a low dimensional latent space to capture an essential representation of the data. We construct our VAE using deep convolutional neural networks whose architecture and training parameters is described at length in the methods section. The VAE is composed of an encoder network that encodes spectra as a Gaussian probability distribution in the 2-dimensional latent space (schematically depicted as μ and Σ), and a decoder network that takes points from the latent space and decodes them back into the original spectra.

By encoding spectra as probability distributions in a low dimensional latent space, the VAE provides us with 3 useful features: 1) Clustering. As all spectra are encoded as distributions, they will overlap with one-another. If overlapping distributions are not from similar spectra the model will be heavily penalized during training. This results in a well structured latent space that enables the use of simple models to make predictions from encoded data. 2) Denoising. The low dimensional latent space does not contain enough information to encode for noise. This improves predictions made from models trained on encoded data, especially for small amounts of labeled data. 3) Interpretation. Encoding spectra as distributions ensures that the latent space will be a continuous representation of the different classes of antibiotic treatments. This allows us to decode spectra across the latent space to ensure that the variation of SERS intensity is consistent with our expectations.

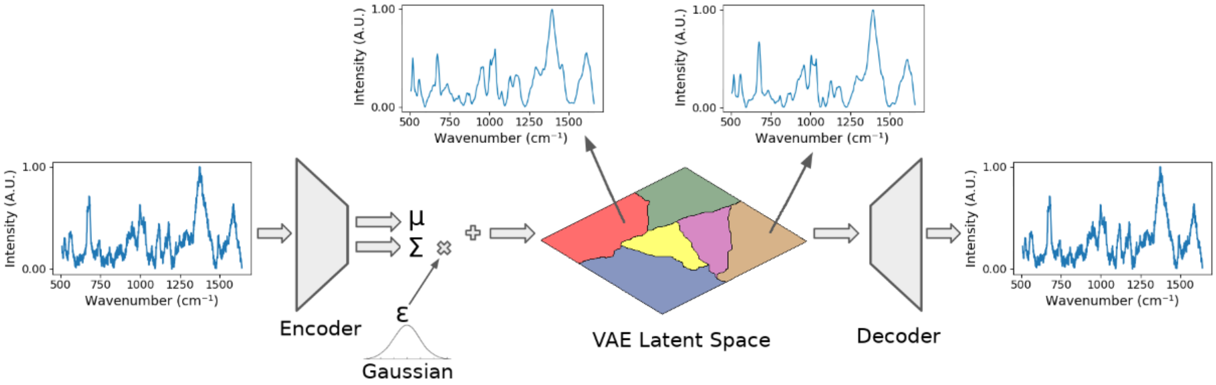

Figure 9.3: depiction of the VAE model with training data plotted on the left and right, and VAE generated spectra plotted on the top left and top right. Spectra are encoded into the latent space as Gaussian distributions with mean and variance sigma. The encoder and decoder are deep convolutional neural networks.

Figure 9.4 a depicts the VAE space obtained using the AST dataset described above. The clustering of the different treatment classes is obvious upon inspection, although large variances within sample classes and the clusters are strongly anisotropic between the first and second VAE dimension, VAE 1 and VAE 2, respectively. From Figure 9.4 a one may observe a clear trend of cell health across VAE 1, with small VAE 1 values corresponding to treatment with antibiotics PA is susceptible to, intermediate VAE 1 values corresponding to treatment with antibiotics PA is resistant to, and large VAE 1 values corresponding to the untreated PA. We also observe lower VAE 1 values with shorter treatment times for all treatment classes, which is expected for these antibiotic treatments at the MIC as the bacteria mostly recovers from treatment.

We can now use the generative aspect of the VAE to validate the clustering behaviour observed. Figure 9.4 b depicts VAE generated (decoded) spectra from the latent space that corresponds to the mean value of each treatment class. Comparison of these spectra to the average spectra plotted in Figure 9.1 a, b, c reveals a striking similarity. The denoising property of the VAE can be seen by considering that Figure 9.1 spectra are averages of many spectra, and Figure 9.4 spectra are individual spectra. Comparison of the preprocessed spectra in Schematic 1 and generated spectra is also revealing. From this, one can observe continuous deformation across treatment classes. It is obvious that good cell health is associated with large vibrations in the 1100 cm^{-1} - 1200 cm^{-1} bands, and larger 1030.3 cm^{-1} band than 1000.7 cm^{-1} band, among many other intensity changes.

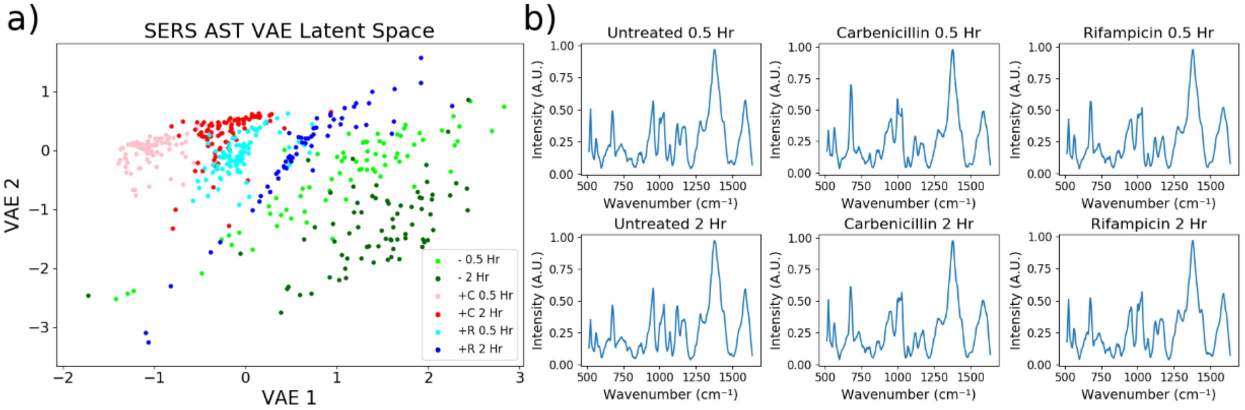

Figure 9.4: a) VAE space of AST spectra, with test AST spectra depicted. b) VAE generated spectra from the center of the class centroid.

As discussed above, the primary challenge to the acceptance of SERS based AST is the need to validate the SERS analysis with 24 hour cell culture analysis. From a machine learning perspective we can think of the cell culture as our method of generating labels, and the SERS analysis as the input data. The huge mismatch between label generation time and data generation time motivates a semi-supervised learning approach. The correspondence between the VAE encoded SERS spectra and the antibiotic treatment classes suggests that a semi-supervised approach using the VAE latent space. In this approach, large quantities of unlabeled SERS spectra are used to train the VAE, generating a meaningful representation of antibiotic treatment response in the VAE latent space. We can then train models with the few labeled examples from this latent space.

Figure 9.5 depicts model performance on the AST dataset of k-nearest neighbor (KNN) and support vector machine (SVM) models on preprocessed spectra and VAE encoded spectra as a function of the number of examples. Random examples are removed from the AST dataset without replacement and used to train the models. This process is repeated 50 times and the mean and standard deviation of the model accuracy are reported. From Figure 9.5, one may observe a significant improvement of both accuracy and variance from models trained using VAE encoded spectra compared to preprocessed spectra. The largest accuracy achieved is 80.6 +/- 0.2 % and is achieved by the 10 example SVM on VAE encoded data. This relatively high accuracy achieved with so few labeled examples is an important achievement as it shows that very little 24 hour cell culture is necessary to build a SERS AST model.

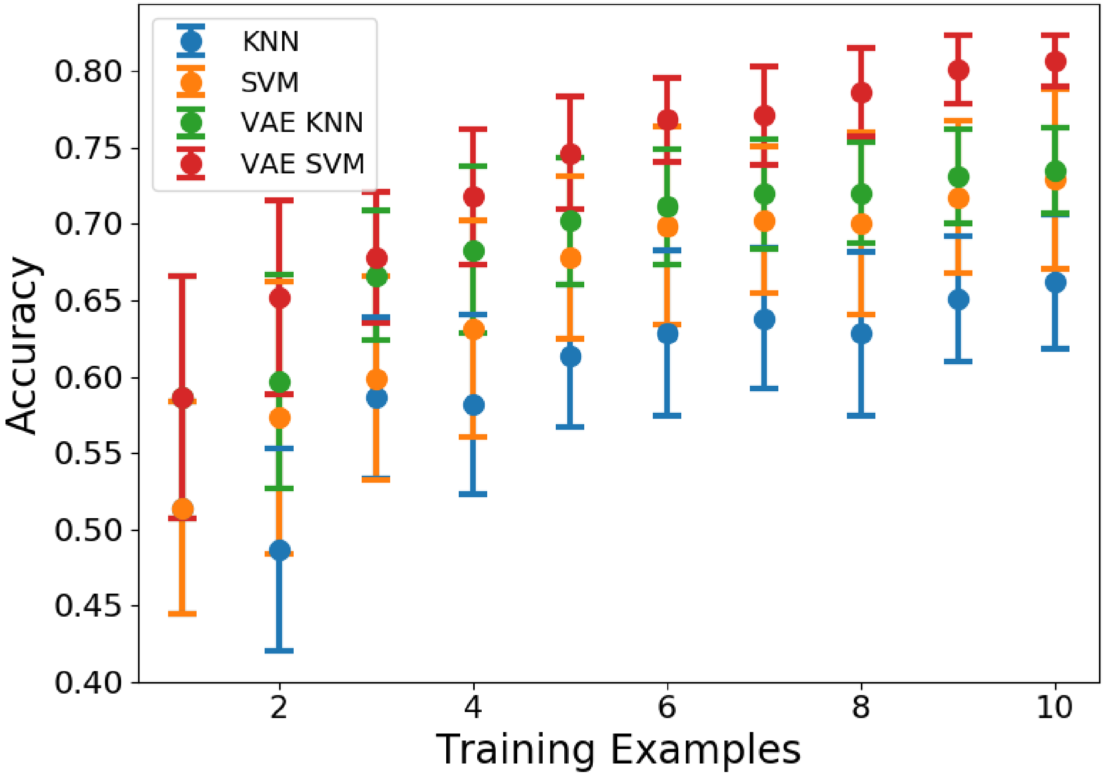

Figure 9.5: depicts various model performance versus the number of training examples from the AST dataset used. KNN and SVM models are evaluated on preprocessed spectra (KNN), (SVM), and on VAE encoded spectra (VAE KNN), (VAE SVM). The best accuracy is achieved after 10 examples by the VAE SVM at 80.6 +/- 0.2 %.

9.4 Transfer Learning for SERS AST

Transfer learning, the use of a large dataset to improve model predictions on a related but smaller dataset, has emerged as one of the most important tools in deep machine learning.[439] While large SERS datasets aren't currently publicly available, they are easy to generate quickly using 2PAC fabricated SERS sensors as measurements are performed in just 0.1 s. Based on the importance of 1000 cm^{-1} - 1200 cm^{-1} in differentiating the antibiotic treatment classes, we have chosen 6 aromatic volatile bacterial metabolites: 2-methyl napthalene, o-cresol, 2-amino acetophenone, pyrrole, 2-pentyl furan, and indole.[440] We construct a "metabolite mixture" dataset by preparing all 63 combinations of these 6 metabolites at a total concentration of 1 part per million in water and acquiring SERS spectra from each of them. Solution preparation details are described in the methods section. Figure 9.6 a depicts a t-SNE visualization of this dataset, where good differentiation between the mixtures is apparent. While t-SNE is known not to be an ideal algorithm for finding global similarities within datasets, larger clusters of similar mixtures are still observed.

Traditionally, transfer learning with deep neural networks is done by training a model, like a convolutional neural network (CNN) with the large dataset, and then fine tuning the model's parameters with the smaller dataset. This method has been shown to produce outstanding results and will be explored below, but the VAE method enables us to utilize transfer learning in a totally unsupervised way, which we will discuss here. Our approach is to combine the large, metabolite mixture dataset with the small AST dataset and train a VAE. This combined VAE space is depicted in Figure 9.6 b. While this complex dataset, the "combined" dataset, greatly benefits from a higher dimensional latent space (see Figure 9.9 for a t-SNE of the metabolite mixture VAE space in 32 dimensions), we constrain the latent space to 2 dimensions so that it can be easily examined. This constraint increases the loss of the fully trained model by nearly 50% and results in anisotropic and mixed clusters of much of the data. Yet from the inset of FIgure 9.6 b, one can observe that the original AST dataset is well separated in this space with extremely low density regions between the different conditions.

The exceptional clustering of the AST dataset observed in the inset of Figure 9.6 b is the most important result of this work. The clustering (in just 2 dimensions) is significantly better than that observed in Figure 9.4 a, and is similar to t-SNE, but (unlike t-SNE) can be used to develop predictive models. This demonstrates that a simple, and easily acquired toy dataset of bacterial metabolites can be leveraged to improve models of complex bacterial lysate.

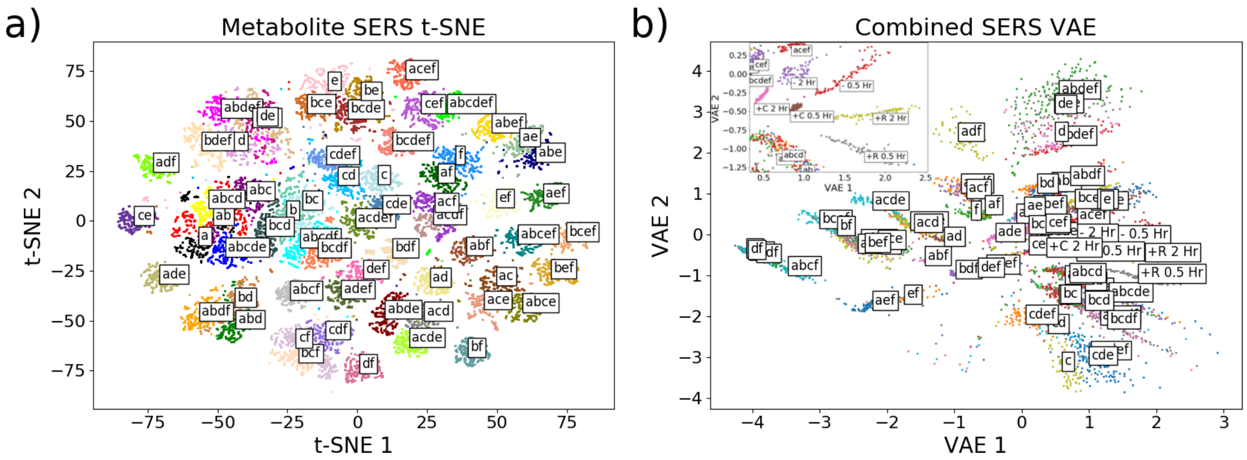

Figure 9.6: a) t-SNE of metabolite mixture dataset. Labels are defined as follows: 2-methyl napthalene (A), o-cresol (B), 2-amino acetophenone (C), Pyrrole (D), 2-pentyl furan (E), and Indole (F). All 63 possible mixture combinations are plotted. Labels are set to the mean value of each condition cluster. b) VAE latent space from the test portion of both datasets combined, inset: zoom in of the region encoding the AST dataset.

From here, it is straightforward to build a predictive model. The VAE encoding makes it obvious that some of the spectra are outliers that are unduly influenced by the signal of the underlying SERS surface. These spectra are easily removed with an isolation forest set to remove 5% of spectra. This process is detailed in the methods section and the isolation forest identification is depicted in Figure 9.10. In Figure 9.7 a, we depict a simple, unsupervised Bayesian Gaussian Mixture model of the outlier removed AST dataset encoded with the combined VAE. Nearly perfect identification of the different classes is achieved, with an accuracy of 99.3%. From Figure 9.7 a, one can also observe that the combined VAE encoding nicely groups the different antibiotic conditions together, which is important for the application of this strategy to clinical datasets.

Finally, we turn to examine modern neural network model performance on preprocessed spectra and combined VAE encoded spectra. These results are plotted in Figure 9.7 b, where as in Figure 9.5, random examples are taken to train a model and the performance is evaluated on the rest of the model. Here we repeat this procedure 10 times and depict that mean and standard deviation of the model accuracy. We compare two models, a multilayer perceptron (MLP, a fully connected artificial neural network with 1 hidden layer) which is used for inference from the combined VAE latent space, and a convolutional neural network which is used for inference from preprocessed spectra. The architecture and training details are described in the methods section. Transfer learning is performed by taking the 6 unmixed metabolite datasets (eg 2-methyl napthalene, o-cresol, 2-amino acetophenone, pyrrole, 2-pentyl furan, and indole dissolved in water by themselves), and training a model with that data.

From Figure 9.7 b it is obvious that CNNs are an incredibly powerful tool for inference from SERS spectra. Even without transfer learning a CNN achieves good results with just 4 examples. Transfer learning with the CNN improves results regardless of the number of examples but the difference is most pronounced with very few examples. On the other hand, an MLP does not yield good predictions from the VAE space, likely due to underfitting as there are only 2 features. However, a transfer learned MLP is determined to produce the best predictions, especially for 1 shot learning, but also at higher number of examples. Interestingly, none of these models outperforms the simple Bayesian Gaussian Mixture model discussed above.

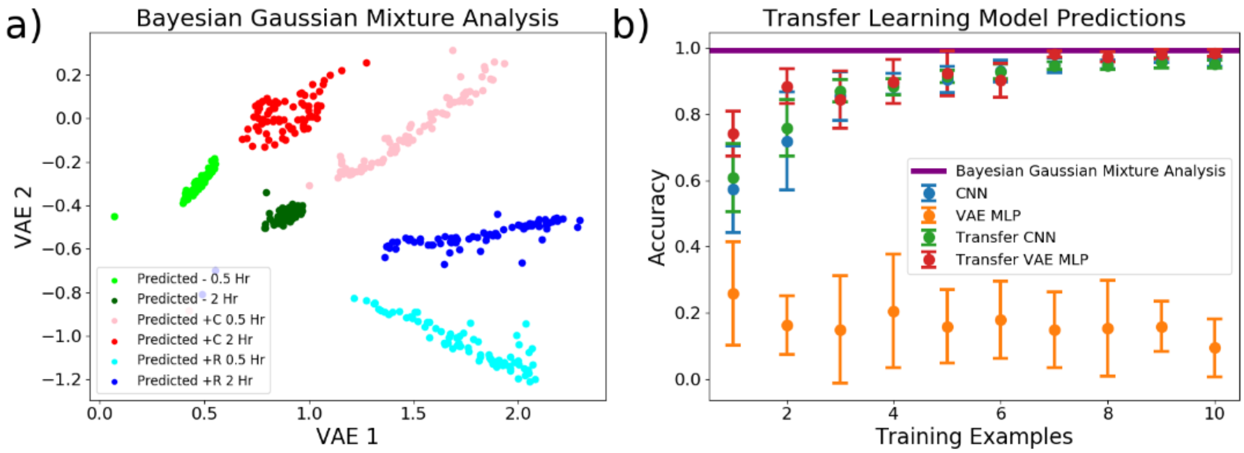

Figure 9.7: a) Bayesian Gaussian Mixture analysis of combined VAE encoded AST test data b) Comparison of transfer learning model performance. b) Transferred model weights are trained using the unmixed preprocessed spectra from the metabolite dataset (Transfer CNN) and these spectra encoded using the combination VAE (Transfer MLP). These models are fine tuned with example from the AST dataset and evaluated on test data. Model performance with various numbers of training examples using models with the same architecture but with standard Xavier weight initialization are also depicted for preprocessed spectra (CNN) and VAE encoded spectra (MLP). The unsupervised Bayesian Gaussian mixture method achieves the highest accuracy at 99.3%.

9.5 Conclusion

In conclusion, we have proposed a semi-supervised and unsupervised deep machine learning approach for surface enhanced Raman scattering (SERS) based antimicrobial susceptibility testing (AST) using the variational autoencoder (VAE). For our proof of concept, we have collected SERS spectra from bacterial lysate of cultures that were exposed to various antibiotic conditions. We have shown that the VAE's latent space enables facile interpretation of the important features within the SERS spectra. The VAE latent space is well behaved and produces clusters of SERS spectra that represent healthy cells at one end of the space and unhealthy cells at the other end of the space. We take advantage of this clustering, as well as the denoising property of the VAE latent space to construct a semi-supervised support vector machine that surpasses 80% accuracy with just 10 labeled SERS AST examples. We show that a toy dataset of various mixtures of metabolites greatly improves the clustering of the SERS AST data. The VAE encoding using the metabolite dataset enabled us to use unsupervised Bayesian Gaussian Mixture analysis to achieve 99.3% classification accuracy of this dataset, which is shown to be better than that achieved using a transfer learned convolutional neural network. These results show that the time consuming 24 hour cell culture which is necessary to generate labels for SERS AST can be minimized, which will greatly reduce the cost of SERS AST and enable its use in a clinical setting. Future work will analyze samples large numbers of complex bacteria from drug resistant clinical urinary tract infection patients.

9.6 Methods

9.6.1 Materials

Random copolymer poly(styrene-co-methyl methacrylate)-α-hydroxyl-ω-tempo moiety (PS-r-PMMA) (M_n = 7400, 59.6% PS) and diblock copolymer poly(styrene-b-methyl methacrylate) (PS-b-PMMA) (M_n = 170-b-144 kg mol $^{-1}$) were purchased from Polymer Source, Inc. (Dorval, Canada). 40 nm diameter lipoic acid functionalized gold nanospheres were purchased from Nanocomposix (San Diego, CA, USA). Si(001) wafers with a resistivity of 0.004 ohm-cm were purchased from Virginia Semiconductor (Frederickburg, VA, USA). Hydrofluoric acid (HF) was purchased from Fisher Scientific (Pittsburgh, PA, USA). 2-(N-morpholino)ethanesulfonic acid (MES) 0.1 M buffer, 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide hydrochloride (EDC), and N-hydroxy sulfosuccinimide (s-NHS), dimethyl sulfoxide (DMSO), ethylenediamine, benzenethiol, toluene, ethanol, isopropyl alcohol (IPA), potassium carbonate, and 52-mesh Pt gauze foil were purchased from Sigma-Aldrich (St. Louis, MO, USA). Nanopure deionized water (DI) (18.2 MΩ cm^{-1}) was obtained from a Milli-Q Millipore System.

9.6.2 Surface enhanced Raman scattering (SERS) Sensor Fabrication

Sensor fabrication has been described in depth in previous work.[272] First we prepare block copolymer templates for Au nanosphere assembly attachment. Random PS-b-PMMA block copolymer is spin-coated onto a HF-cleaned Si wafer (*HF has serious potential to cause severe injury which mandates extreme care during treatment*) and annealed for 72 hours. The wafer is rinsed with toluene rinse and lamella forming PS-b-PMMA block copolymer is spin coated onto the wafer, which is annealed for another 72 hours. This process is described elsewhere.[137] Next, PMMA regions within the block copolymer are selectively functionalized with amine terminated end groups by immersing a 1 cm x 1 cm piece of the wafer in DMSO. This substrate is then transferred into an ethylenediamine/DMSO solution (5% v/v). Both immersions are performed for 5 min without rinsing between steps. The functionalized template is then rinsed with IPA for 1 min and dried under nitrogen for immediate use.

An electrohydrodynamic flow driven assembly of Au nanospheres is used to generate assemblies with the following method: Au nanosphere solution (0.1 mg/mL, 3 mL) is added to a clean 10 mL glass beaker. s-NHS (20 mM) in MES (0.1 M) buffer (35 μL) is added to the nanosphere solution and swirled. Next, EDC (8 mM) in MES (0.1 M) buffer (35 μL) is added to this solution and swirled. The solution brought to, and maintained at 60 °C with a hot plate. The functionalized block copolymer-coated Si substrate is placed vertically into the solution and held in place with alligator clips, taking care to avoid any contact of the alligator clips with the solution. A 1 cm × 1 cm Pt mesh is placed parallelly 1 mm away from the substrate. 1.2 V is applied across the mesh and substrate using a DC power supply for 10 min. Everything is then rinsed with IPA for 1 min and dried under nitrogen. The process is repeated with the same substrate and a fresh nanosphere solution, but with 25 μL of s-NHS and EDC solution.

9.6.3 Bacterial Cell Culture and Antibiotic Treatments

Growth Conditions: *Pseudomonas aeruginosa* strain PA14 are revived from frozen stock by streaking on Luria Bertani (LB) agar (1.5% agar) and then incubated in 37 ^{o}C incubator. A single colony from the agar plate is then inoculated in 5 mL of LB broth, and then grown overnight with shaking in 37 ^{o}C.

Minimum inhibition curves: Overnight culture of PA14 are centrifuged, and then re-suspended in enough fresh LB media to reach a final optical density at 600 nm (OD600) of 1. Carbenicillin and Rifamipicin are serially diluted 2-fold in fresh LB broth to reach concentrations double our final desired concentrations. Finally, the re-suspended PA14 culture is mixed one to one (1 mL bacteria + 1 mL of LB antibiotic mixture) with media containing antibiotics in culture tubes, so that the final OD was 0.5 and the antibiotics are at the desired concentrations. The untreated control sample is prepared by mixing the re-suspended PA14 culture one to one with antibiotic-free LB broth. The culture tubes are shaken for 24 hours at 37 ^oC, and then bacterial proliferation was assessed by measuring the optical density of treated bacterial culture. From these MIC curves, a dose of either Carbenicillin or Rifampicin is chosen that affects PA14 proliferation in 24 hours without completely killing off the culture for SERS study.

Antibiotic treatment: Overnight PA14 cultures are centrifuged and re-suspended to reach a volume of 20 mL at OD600 of 0.5. Cabenicillin is added so that the final concentration was 50 μg/mL. The final concentration of Rifampicin is 400 μg/mL. At the 0.5, 2, and 3 hours timepoints, the OD600 of the treated PA14 culture is measured, and 10^9 CFUs are drawn out for lysate collection.

Lysate collection: 10^9 CFUs of PA14 cells treated or untreated with antibiotic are spun down and re-suspended in 500 μL of sterile PBS three times to wash off the LB growth media. Then cells are re-suspended 50 μL of sterile double deionized water, and then heated at 95 ^oC for 20 minutes using heated aluminum block. Cell debris is centrifuged and discarded, and the 50 μL supernatant is collected for SERS study.

9.6.4 Metabolite Mixture Preparation

Metabolite mixtures are prepared as follows: 2-methyl napthalene (A), o-cresol (B), 2-amino acetophenone (C), Pyrrole (D), 2-pentyl furan (E), and Indole (F) are dissolved in ethanol at a concentration of 100 ppm. Then 1 ppm solutions are prepared in water from these ethanol stock solutions. The 63 combinations of metabolites are prepared by mixing the water stock solutions to maintain a total metabolite concentration of 1 ppm.

9.6.5 SERS Spectroscopy

All SERS spectroscopy measurements are conducted using a confocal Renishaw InVia micro Raman system with a 785 nm diode laser, a laser power of 14 μW, an exposure time of 0.5

s, and a 60x water immersion objective with a 1.2 numerical aperture. Bacteria cell lysate or metabolite mixture solutions are used as the immersion media. After soaking the SERS substrate in the sample for 15 minutes, Raman maps are collected with a spacing of 4 μm spacing between points. For each sample one 20 x 20 pixel Raman map is acquired.

9.6.6 Spectra Preprocessing

Spectra are preprocessed in three steps: 1) smoothing 2) background subtraction and 3) scaling, all done using the Python 3.3 programming language. Smoothing was done with the Savitzky-Golay method[441] as implemented in Scikit-Learn using an 11 pixel window and polynomial order 3. Background subtraction was done with the asymmetric least squares method[442] and was implemented in NumPy with $\lambda = 10000$, $p = 0.001$. Spectra were scaled to have a minimum value of 0 and maximum value of 1 with Scikit-learn's minmaxscaler.

9.6.7 Unsupervised Data Visualization

t-distributed stochastic neighbor embedding (t-SNE)[438] and non-negative matrix factorization (NMF)[355] are used for unsupervised visualization of the SERS datasets. Both techniques are performed on preprocessed spectra using Scikit-Learn with default parameter settings. NMF is performed with 3 components, and the spectra scores are evaluated using the first component. Data acquired with deionized water is included in training these algorithms.

9.6.8 Variational Autoencoder Implementation

All artificial neural network models are implemented in keras and use the adam optimizer.[443]

Prior to use in the variational autoencoder[434] (VAE), spectra are preprocessed as described above. These 1011 dimensional spectra are padded with zeros to 1024 dimensions and reshaped to a dimension of (examples, 1024, 1) for use in 1 dimensional convolutional neural network (1D

CNN) layers. All 1D CNN layers have a kernel window of 8 pixels, a stride of 2, are regularized with a maximum kernel norm of 3, have parametric relu activations, are batch normalized, and followed with a 30% dropout layer. Early stopping is implemented with test loss, and the batch size used is 32.

The VAE is implemented differently for the antimicrobial susceptibility testing (AST) dataset and the AST and metabolite mixture combined dataset. For the smaller AST dataset, the encoder network is composed of 4 1D CNN layers with 32, 32, 64, and 64 filters. This output is flattened and sent to a 128 node fully connected layer with parametric relu activation, batch normalization, and 30% dropout and sent to a 32 node fully connected layer with parametric relu activation, and finally to fully connected layers with 2 nodes that represent the mean and standard deviation of the encoded input. The decoder is similar with a 1344 node fully connected layer, reshaped and sent to 4 1D transposed CNN layers with 64, 64, 32, and 32 filters. This is output to a 1D transposed CNN with stride 1, sigmoid activation, and a stride of 1. The model was optimized with the adam optimizer with a loss function defined as KL divergence + mean absolute error * 10.

For the larger combined dataset, the encoder network is composed of 6 1D CNN layers with 32, 32, 64, 64, 128, and 128 filters, with 40% dropout. This output is flattened and sent to a 256 node fully connected layer with parametric relu activation, batch normalization, and 50% dropout and sent to a 64 node fully connected layer with parametric relu activation, and finally to fully connected layers with 2 nodes that represent the mean and standard deviation of the encoded input. The decoder is similar with a 2048 node fully connected layer, reshaped and sent to 8 1D transposed CNN layers with 256, 256 (stride 1), 256, 256 (stride 1), 128, 128, 64, and 64 filters. This is output to a 1D CNN with 1 filter, sigmoid activation, and a stride of 1. The model was optimized with the adam optimizer with a loss function defined as KL divergence + mean absolute error * 80.

9.6.9 Semi-supervised Learning

The models in Figure 9.5 are evaluated as follows. Examples are pulled from the test dataset used in training the VAE described above. These are used to train K-nearest neighbors (KNN) – with the number of neighbors being the number of examples – and support vector machine (SVM) models with Scikit-learn using default settings. The KNN and SVM models are trained using preprocessed spectra with dimension 1011 and the accuracy is evaluated using the rest of the test dataset. The VAE KNN and VAE SVM models were evaluated with the same examples projected into the latent space of the trained AST VAE and evaluated with the same dataset. This process is done 50 times and the mean and standard deviation are of the model accuracy on the remaining data are depicted.

9.6.10 Transfer Learning

This section describes the methods used to produce Figure 9.7. The dataset used is the AST dataset encoded into a 2 dimensional latent space using the combined VAE. Outliers are removed by training an isolation forest[444] on the training dataset and applying it to the training and test dataset. Isolation forest is implemented in Scikit-learn with the default settings and an outlier fraction of 5%. The outlier removed training dataset is then used to train a Bayesian Gaussian Mixture Model, which is implemented in Scikit-learn with the default settings and 6 components,[445] and evaluated on the test dataset, which is plotted in Figure 9.7a.

The neural network models evaluated in Figure 9.7b are evaluated with binary cross entropy loss and have the following architectures. First two models are trained for the transfer learning with the six unmixed metabolite full datasets. The first network is a deep CNN trained on the full dimensional preprocessed data and is composed of 4 1D CNN layers with parameters as above and filters of 16, 16, 32, and 32 that are followed by 50% dropout layers and batch normalized. This output is flattened and set to a 6 node fully connected layer with softmax activation. The second network is a multilayer perceptron trained on the VAE encoded, outlier removed data and is composed of 2 fully connected layers with 8 and 16 nodes and are batch normalized and with relu activation. This output is sent to a fully connected layer with 6 nodes and softmax activation. The weights of these trained networks are then fine tuned with the few example AST dataset. Additionally, these same models are evaluated with the same AST examples with standard Xavier initialization of the weights. The accuracy of these models are evaluated, and repeated 10 times to obtain a mean and standard deviation of the model accuracy and plotted.

9.7 Supplemental Information

Figure 9.8 a,b depicts the data used to identify the minimum inhibitory concentration (MIC) of carbenicillin and rifampicin applied to *Pseudomonas aeruginosa* (PA), respectively. The optical density (OD) of a PA culture is measured 24 hours after exposure of antibiotic. Prior to treatment the PA culture is diluted to an OD of 0.5. The MIC is determined to be 50 μg/mL for carbenicillin and 400 μg/mL for rifampicin.

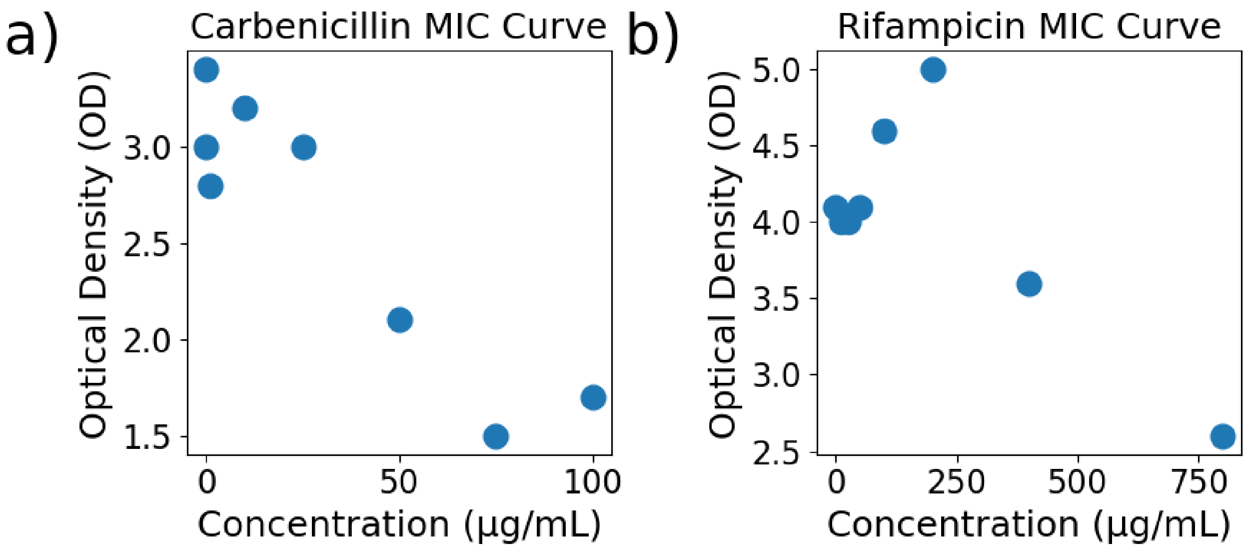

Figure 9.8 a: Curve used to identify the minimum inhibitory concentration of carbenicillin for *Pseudomonas aeruginosa*. B: Curve used to identify the minimum inhibitory concentration of rifampicin for *Pseudomonas aeruginosa*.

Figure 9.9 depicts a t-sne visualization of a 32-dimensional VAE latent space trained on the metabolite combination dataset. Increasing the latent space from 2-dimensions to 32-dimensions greatly improves the clustering of the different metabolite conditions, and reduces the test loss of the VAE model by nearly 50%.

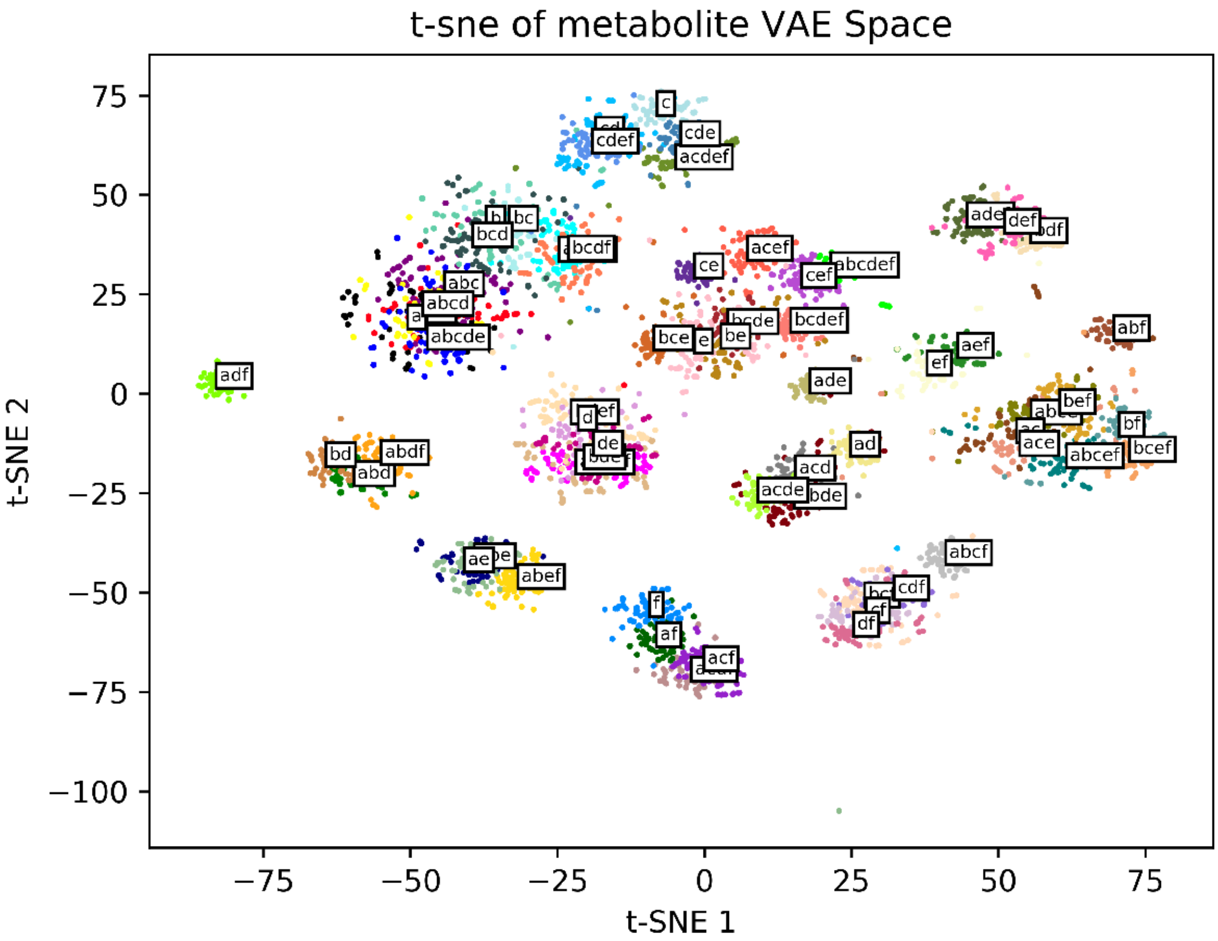

Figure 9.9: depicts a t-SNE visualization of a 32-dimensional metabolite combination VAE space.

In addition to the superior clustering that the combination VAE latent space provides, it also enhances the difference between outliers and inliers. Figure 9.10 depicts an isolation forest outlier detection applied to the AST spectra that has been encoded into the combination VAE model. Outliers are easily identified by isolation forest and removed for processing in the predictive models.

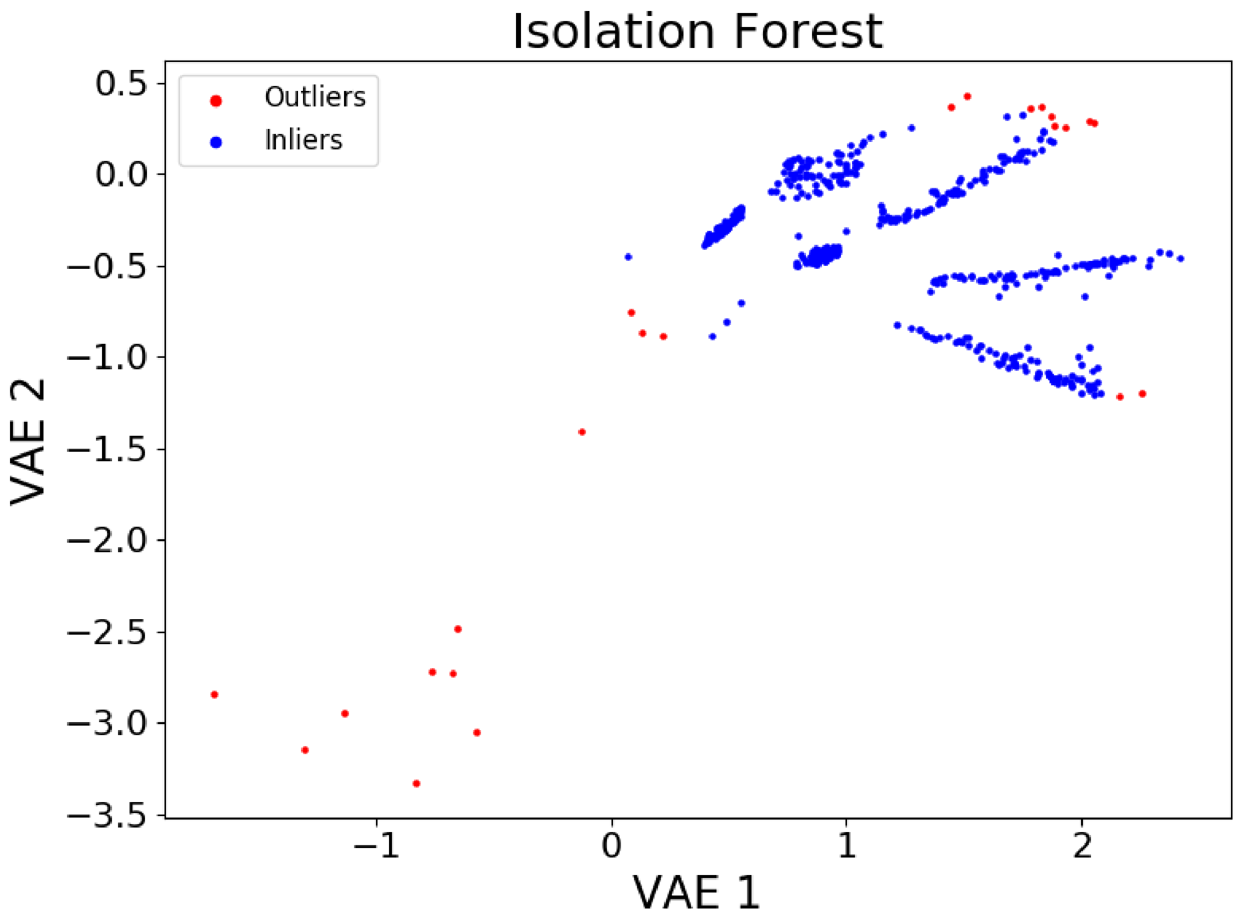

Figure 9.10: Isolation forest predictions of AST spectra that have been encoded with the combination VAE model. Outliers are shown as red and inliers are shown as blue.

Chapter 10

Conclusion

This work first demonstrated a novel nanomanufacturing method where electrohydrodynamic (EHD) flow was used as a long-range driving force to enable carbodiimide cross-linking between nanospheres. This produced nanosphere oligomers exhibiting sub-nanometer gap spacing over mm^2 areas. I investigated the anhydride linkers between nanospheres are observed via surface-enhanced Raman scattering (SERS) spectroscopy. These anhydride linkers were found to be cleavable via nucleophilic substitution and enabled the placement of nucleophilic molecules in electromagnetic hotspots. Atomistic simulations elucidated the transient attractive force provided by EHD flow, which was found to be needed to provide a sufficient residence time for anhydride cross-linking to overcome slow reaction kinetics. This synergistic analysis showed that assembly involves an interplay between long-range and short range driving forces. The long-range driving forces increased nanoparticle − nanoparticle interactions and the probability that ligands are in proximity to overcome activation energy barriers associated with short-range chemical reactions. Absorption spectroscopy and electromagnetic full-wave simulations showed that variations in nanogap spacing had a greater influence on optical response than variations in close-packed oligomer geometry. The EHD flow−anhydride cross-linking assembly method (referred to as a 2-dimensional physically activated chemical self-assembly, or 2PAC) enable close-packed oligomers with uniform gap spacings that produce uniform SERS enhancement factors. These results demonstrated the efficacy of colloidal driving forces to selectively enable chemical reactions leading to future assembly platforms for large-area nanodevices.

Then I investigated the use of 2PAC to develop a directed chemical assembly of nanoparticles on electron beam lithography defined Au nanopillars. This method addressed the challenges in obtaining uniformity in nanostructure geometry and nanometer scale gap spacings in structures. EHD flows yielded robust driving forces between the template and nanoparticles as well as between nanoparticles on the surface which promoted the assembly of close-packed metamolecules. Chemical crosslinking between Au surfaces enabled molecular control over gap spacings between nanoparticles and Au pillars. An as-fabricated structure was analyzed via full wave electromagnetic simulations and shown to produce large magnetic field enhancements on the order of 3.5 at optical frequencies. This novel method for directed self-assembly demonstrated 2PAC's utility in the fabrication of plasmonic metamolecules with unique electromagnetic properties.

The thesis then turned to applications of these 2PAC manufactured chemical sensors. First Cuong Nguyen and I developed a method to detect pyocyanin a secondary metabolite of *Pseudomonas aeruginosa*, in aqueous media at concentration of 100 $pg \cdot mL^{-1}$. We used a partial least squares regression to analyze SERS data of conditioned medium from a bacterial culture, which had a more complex background than aqueous media and we achieve 1 $ng \cdot mL^{-1}$ limit of detection. The model that found pyocyanin concentration from SERS spectra was robust and concentration quantification spanned five-orders of magnitude. 2PAC manufactured sensors were also incorporated in an in-line microfluidic device, enabling real-time pyocyanin detection in medium effluent as early as 3 hours after inoculation and quantification in under 9 hours during *P. aeruginosa* biofilm formation. Surface-attached bacteria exposed to a bactericidal antibiotic were differentially less susceptible after 10 h of growth, indicating that these devices may be useful for early intervention of bacterial infections.

Next I demonstrated the use of a SERS based odor compass. An odor compass is a sophisticated example of an artificial olfaction device which uses chemical sensor data to identify odor source direction. Before this work passive, diffusion based odor compasses remained elusive as detection of low analyte concentrations and quantification of small concentration gradients from within the sensor platform are a necessity and challenging to achieve. Additionally, simultaneously identifying multiple odor sources using an odor compass was an ongoing challenge, especially for similar analytes. In that work, I showed that 2PAC sensors overcame these challenges and presented the first SERS odor compass. With a grid array of SERS sensors and machine learning analysis I showed reliable identification of multiple odor sources arising from diffusion of analytes from one or two localized sources. Convolutional neural network and support vector machine classifier models achieved over 90% accuracy for a multiple odor source problem. I also showed that this system identifies the location of an *Escherichia coli* biofilm via its complex signature of volatile organic compounds. Solving the multiple odor source problem with a passive platform opens a path toward an internet of things approach to monitor toxic gases and indoor pathogens.

I then turned improve analyte concentration regressions using machine learning. Underpinning SERS sensing are the sensor surfaces that are composed of vast quantities of metal nanostructures which confine light into small gaps called “hotspots”, that enhance Raman scattering. While these surfaces are essential for increasing Raman scattering intensity so that analyte signal may be observed in small concentrations, they introduce signal variations due to spatial distributions of Raman enhancement and hotspot volume. I introduced a convolutional neural network model that improves concentration regressions in SERS sensors by learning the distributions of sensor surface dependent latent variables. I demonstrated that this model significantly improves predictions compared to a traditional multilayer perceptron approach, and that the model uses analyte spectral information and is capable of reasonable interpolations.

With the methods that I had developed I sought to solve a longstanding challenge within the SERS community, identifying concentration in the single molecule regime (SM). Single molecule (SM) detection represents the ultimate limit of chemical detection. Yet concentration quantification in the SM regime remained an ongoing challenge, with the few existing methods requiring carefully developed calibration curves that must be completely redeveloped for each new analyte molecule. I demonstrated that a convolutional neural network (CNN) model applied to small images of SERS spectra yields is a robust, facile method for concentration quantification in the SM regime. I further demonstrated that transfer learning, the process of reusing the weights of a trained CNN model, greatly reduced the amount of data required to train new CNN models on new analyte molecules. These results point the way for the use of SERS in important ultralow concentration chemical detection applications such as metabolomic profiling, water quality evaluation, and fundamental research.

Finally, I demonstrated a semi-supervised machine learning method to reduce the amount of data necessary for antimicrobial susceptibility testing (AST) using SERS. My approach used a variational autoencoder (VAE) to learn a representation of the important vibrations observed in bacterial cell lysate that had been treated with different antibiotic conditions. This work showed that VAEs provide spectroscopists with highly interpretable, yet highly powerful models to analyze spectra. I demonstrated that VAEs naturally cluster spectra of bacterial cell lysate into their antibiotic treatment conditions, and that this clustering greatly improved predictions of models using the VAE latent space. I demonstrated that a SVM model achieved achieved over 80% accuracy on a test dataset with just 10 example spectra. I then showed that a toy dataset of bacterial metabolites dissolved in water could greatly improve the VAE latent space. This enabled an unsupervised Bayesian Gaussian Mixture model to achieve 99.3% accuracy on a test AST dataset, even higher than a deep CNN transfer learning based approach for fewer than 10 example spectra. This work was an important proof of concept, demonstrating that SERS analysis of AST data is amenable to a semi-supervised and even an unsupervised approach, making SERS AST a practical method for clinical AST.

In conclusion, my thesis has worked toward understanding the state of bacterial systems with machine learning-enabled interpretation of SERS spectra that were produced by a novel nanomanufacturing method. In the beginning, I used nanomanufacturing to produce state of the art sensors with chemical self-assembly. While this enabled the detection of even single molecules, the strategy also introduces an enormous amount of complexity compared to traditional Raman spectroscopy. The ligand molecules within the hot spots produce a complex background signal, and the signal enhancement is not completely uniform. For this reason, I developed machine learning approaches for SERS. These approaches also ended up being tremendously valuable for handling extremely complex bacterial fluids. This work points the way for using SERS sensors in clinical analysis, and could revolutionize the way that we as a society diagnose diseases.

References

(1) Oddy, A. Assaying in Antiquity. *Gold Bull.* **1983**, *16* (2), 52–59. https://doi.org/10.1007/BF03214624.

(2) History of Analytical Chemistry - 1st Edition https://www.elsevier.com/books/history-of-analytical-chemistry/belcher/978-0-08-010980-0 (accessed Jul 16, 2019).

(3) Boyle, R. *Experiments and Observations upon the Saltiness of the Sea*; 1674.

(4) Griffin, J. *A Practical Treatise on the Use of the Blowpipe*; 1827.

(5) Yount, L. *Antoine Lavoisier Founder of Modern Chemistry*; 2008.

(6) Wisniak, J. Jöns Jacob Berzelius A Guide to the Perplexed Chemist. *Chem. Educ.* **2000**, *5* (6), 343–350. https://doi.org/10.1007/s00897000430a.

(7) Gay-Lussac. Assays of Commercial Potash. *Ann. Chim. Phys.* **1828**.

(8) Pfaff. *Handbuch Der Analytischen Chemie Für Chemiker, Staatsarzte, Apotheker, Oekenomen Und Bergwerks*; 1821.

(9) Kirchhoff, G.; Bunsen, R. Chemische Analyse Durch Spectralbeobachtungen. *Ann. Phys.* **1860**, *186* (6), 161–189. https://doi.org/10.1002/andp.18601860602.

(10) BEER, A. Bestimmung Der Absorption Des Rothen Lichts in Farbigen Flussigkeiten. *Ann Phys Chem* **1852**, *86* (2), 78–90.

(11) Thomas, N. The Early History of Spectroscopy. *J. Chem. Educ.* **1991**, *68* (8), 631.

(12) Langley. The Bolometer. *Proc. Am. Metrol. Soc.* **1880**, *2*, 184–190.

(13) Julius, W. H. *Bolometrisch onderzoek van absorptiespectra*; J. Müller, 1893.

(14) Larkin, P. *Infrared and Raman Spectroscopy; Principles and Spectral Interpretation*; Elsevier, 2011.

(15) Raman, C. V. A New Radiation. **1927**.

(16) Boyd, R. W. *Nonlinear Optics*; Academic Press, 2003.

(17) Maiman, T. H. Stimulated Optical Radiation in Ruby. *Nature* **1960**, *187* (4736), 493. https://doi.org/10.1038/187493a0.

(18) Plucker, J. Über Die Einwirkung Des Magneten Auf Die Elektrischen Entladungen in Verdünnten Gasen. *Ann. Phys. Chem.* **1858**, *103*, 88–106.

(19) Thomson, J. J. *Rays of Positive Electricity and Their Application to Chemical Analyses*; Longmans, Green and Company, 1921.

(20) Gohlke, R. S. Time-of-Flight Mass Spectrometry and Gas-Liquid Partition Chromatography. *Anal. Chem.* **1959**, *31* (4), 535–541. https://doi.org/10.1021/ac50164a024.

(21) Munson, M. S. B.; Field, F. H. Chemical Ionization Mass Spectrometry. I. General Introduction. *J. Am. Chem. Soc.* **1966**, *88* (12), 2621–2630. https://doi.org/10.1021/ja00964a001.

(22) Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. Molecular Beams of Macroions. *J. Chem. Phys.* **1968**, *49* (5), 2240–2249. https://doi.org/10.1063/1.1670391.

(23) Karas, Michael.; Bachmann, Doris.; Hillenkamp, Franz. Influence of the Wavelength in High-Irradiance Ultraviolet Laser Desorption Mass Spectrometry of Organic Molecules. *Anal. Chem.* **1985**, *57* (14), 2935–2939. https://doi.org/10.1021/ac00291a042.

(24) Griffiths, W. J. *Metabolomics, Metabonomics and Metabolite Profiling*; Royal Society of Chemistry, 2007.

(25) Pauli, W. Zur Frage der theoretischen Deutung der Satelliten einiger Spektrallinien und ihrer Beeinflussung durch magnetische Felder. *Naturwissenschaften* **1924**, *12* (37), 741–743. https://doi.org/10.1007/BF01504828.

(26) Rabi, I. I.; Zacharias, J. R.; Millman, S.; Kusch, P. A New Method of Measuring Nuclear Magnetic Moment. *Phys. Rev.* **1938**, *53* (4), 318–318. https://doi.org/10.1103/PhysRev.53.318.

(27) Bloch, F.; Hansen, W. W.; Packard, M. The Nuclear Induction Experiment. *Phys. Rev.* **1946**, *70* (7–8), 474–485. https://doi.org/10.1103/PhysRev.70.474.

(28) Purcell, E. M.; Torrey, H. C.; Pound, R. V. Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Phys. Rev.* **1946**, *69* (1–2), 37–38. https://doi.org/10.1103/PhysRev.69.37.

(29) Proctor, W. G.; Yu, F. C. The Dependence of a Nuclear Magnetic Resonance Frequency upon Chemical Compound. *Phys. Rev.* **1950**, *77* (5), 717–717. https://doi.org/10.1103/PhysRev.77.717.

(30) Kalcher, K. Renaissances and Current Trends with Electrochemical Sensors and Biosensors. *Int. J. Biosens. Bioelectron.* **2017**, *3* (5).

(31) Zuman, P. Electrolysis with a Dropping Mercury Electrode: J. Heyrovsky's Contribution to Electrochemistry. *Crit. Rev. Anal. Chem.* **2001**, *31* (4), 281–289. https://doi.org/10.1080/20014091076767.
(32) Scholz, F. *Electroanalytical Methods: Guide to Experiments and Applications*; Springer, 2013.
(33) Brainina, K.; Neyman, E. *Electroanalytical Stripping Methods*; John Wiley & Sons, 1994.
(34) Guilbault, G. G.; Lubrano, G. J. An Enzyme Electrode for the Amperometric Determination of Glucose. *Anal. Chim. Acta* **1973**, *64* (3), 439–455. https://doi.org/10.1016/S0003-2670(01)82476-4.
(35) Enzyme-Linked Immunosorbent Assay, Elisa | The Journal of Immunology https://www.jimmunol.org/content/109/1/129 (accessed Jul 25, 2019).
(36) Single-molecule enzyme-linked immunosorbent assay detects serum proteins at subfemtomolar concentrations | Nature Biotechnology https://www.nature.com/articles/nbt.1641 (accessed Jul 25, 2019).
(37) Liedberg, B.; Nylander, C.; Lunström, I. Surface Plasmon Resonance for Gas Detection and Biosensing. *Sens. Actuators* **1983**, *4*, 299–304. https://doi.org/10.1016/0250-6874(83)85036-7.
(38) Maier, S. A. *Plasmonics: Fundamentals and Applications*; Springer Science & Business Media, 2007.
(39) Wicher, D. Functional and Evolutionary Aspects of Chemoreceptors. *Front. Cell. Neurosci.* **2012**, *6*. https://doi.org/10.3389/fncel.2012.00048.
(40) Mao, H.; Cremer, P. S.; Manson, M. D. A Sensitive, Versatile Microfluidic Assay for Bacterial Chemotaxis. *Proc. Natl. Acad. Sci.* **2003**, *100* (9), 5449–5454. https://doi.org/10.1073/pnas.0931258100.
(41) Fuller, D.; Chen, W.; Adler, M.; Groisman, A.; Levine, H.; Rappel, W.-J.; Loomis, W. F. External and Internal Constraints on Eukaryotic Chemotaxis. *Proc. Natl. Acad. Sci.* **2010**, *107* (21), 9656–9659. https://doi.org/10.1073/pnas.0911178107.
(42) Ahmer, B. M. M. Cell-to-Cell Signalling in Escherichia Coli and Salmonella Enterica. *Mol. Microbiol.* **2004**, *52* (4), 933–945. https://doi.org/10.1111/j.1365-2958.2004.04054.x.
(43) Miller, M. B.; Bassler, B. L. Quorum Sensing in Bacteria. *Annu. Rev. Microbiol.* **2001**, *55* (1), 165–199. https://doi.org/10.1146/annurev.micro.55.1.165.
(44) Lau, G. W.; Hassett, D. J.; Ran, H.; Kong, F. The Role of Pyocyanin in Pseudomonas Aeruginosa Infection. *Trends Mol. Med.* **2004**, *10* (12), 599–606. https://doi.org/10.1016/j.molmed.2004.10.002.
(45) Michael, B.; Smith, J. N.; Swift, S.; Heffron, F.; Ahmer, B. M. M. SdiA of Salmonella Enterica Is a LuxR Homolog That Detects Mixed Microbial Communities. *J. Bacteriol.* **2001**, *183* (19), 5733–5742. https://doi.org/10.1128/JB.183.19.5733-5742.2001.
(46) Silpe, J. E.; Bassler, B. L. A Host-Produced Quorum-Sensing Autoinducer Controls a Phage Lysis-Lysogeny Decision. *Cell* **2019**, *176* (1), 268-280.e13. https://doi.org/10.1016/j.cell.2018.10.059.
(47) Staymates, M. E.; MacCrehan, W. A.; Staymates, J. L.; Kunz, R. R.; Mendum, T.; Ong, T.-H.; Geurtsen, G.; Gillen, G. J.; Craven, B. A. Biomimetic Sniffing Improves the Detection Performance of a 3D Printed Nose of a Dog and a Commercial Trace Vapor Detector. *Sci. Rep.* **2016**, *6*, 36876. https://doi.org/10.1038/srep36876.
(48) Buck, L.; Axel, R. A Novel Multigene Family May Encode Odorant Receptors: A Molecular Basis for Odor Recognition. *Cell* **1991**, *65* (1), 175–187. https://doi.org/10.1016/0092-8674(91)90418-X.
(49) The Good Scents Company - Aromatic/Hydrocarbon/Inorganic Ingredients Catalog information http://www.thegoodscentscompany.com/data/rw1008641.html (accessed Jul 8, 2019).
(50) Herz, R. *The Scent of Desire: Discovering Our Enigmatic Sense of Smell*; William Morrow: New York, 2007.
(51) Sankaran, S.; Khot, L. R.; Panigrahi, S. Biology and Applications of Olfactory Sensing System: A Review. *Sens. Actuators B Chem.* **2012**, *171–172*, 1–17. https://doi.org/10.1016/j.snb.2012.03.029.
(52) Chen, D.; Haviland-Jones, J. Human Olfactory Communication of Emotion. *Percept. Mot. Skills* **2000**, *91* (3), 771–781. https://doi.org/10.2466/pms.2000.91.3.771.
(53) Bushdid, C.; Magnasco, M. O.; Vosshall, L. B.; Keller, A. Humans Can Discriminate More than 1 Trillion Olfactory Stimuli. *Science* **2014**, *343* (6177), 1370–1372. https://doi.org/10.1126/science.1249168.
(54) Mitro, S.; Gordon, A. R.; Olsson, M. J.; Lundström, J. N. The Smell of Age: Perception and Discrimination of Body Odors of Different Ages. *PLOS ONE* **2012**, *7* (5), e38110. https://doi.org/10.1371/journal.pone.0038110.
(55) Wedekind Claus; Seebeck Thomas; Bettens Florence; Paepke Alexander J. MHC-Dependent Mate Preferences in Humans. *Proc. R. Soc. Lond. B Biol. Sci.* **1995**, *260* (1359), 245–249. https://doi.org/10.1098/rspb.1995.0087.
(56) L Bijland; Bomers, M.; Smulders, Y. Smelling the Diagnosis: A Review on the Use of Scent in Diagnosing. *Neth. J. Med.* **2013**.
(57) Bomers, M. K.; Agtmael, M. A. van; Luik, H.; Veen, M. C. van; Vandenbroucke-Grauls, C. M. J. E.;

Smulders, Y. M. Using a Dog's Superior Olfactory Sensitivity to Identify Clostridium Difficile in Stools and Patients: Proof of Principle Study. *BMJ* **2012**, *345*, e7396. https://doi.org/10.1136/bmj.e7396.

(58) Edwards, T. L.; Browne, C. M.; Schoon, A.; Cox, C.; Poling, A. Animal Olfactory Detection of Human Diseases: Guidelines and Systematic Review. *J. Vet. Behav.* **2017**, *20*, 59–73. https://doi.org/10.1016/j.jveb.2017.05.002.

(59) Maurer, M.; McCulloch, M.; Willey, A. M.; Hirsch, W.; Dewey, D. Detection of Bacteriuria by Canine Olfaction. *Open Forum Infect. Dis.* **2016**, *3* (2). https://doi.org/10.1093/ofid/ofw051.

(60) Sonoda, H.; Kohnoe, S.; Yamazato, T.; Satoh, Y.; Morizono, G.; Shikata, K.; Morita, M.; Watanabe, A.; Morita, M.; Kakeji, Y.; et al. Colorectal Cancer Screening with Odour Material by Canine Scent Detection. *Gut* **2011**, *60* (6), 814–819. https://doi.org/10.1136/gut.2010.218305.

(61) Brooks, S. W.; Moore, D. R.; Marzouk, E. B.; Glenn, F. R.; Hallock, R. M. Canine Olfaction and Electronic Nose Detection of Volatile Organic Compounds in the Detection of Cancer: A Review. *Cancer Invest.* **2015**, *33* (9), 411–419. https://doi.org/10.3109/07357907.2015.1047510.

(62) Willis, C. M.; Church, S. M.; Guest, C. M.; Cook, W. A.; McCarthy, N.; Bransbury, A. J.; Church, M. R. T.; Church, J. C. T. Olfactory Detection of Human Bladder Cancer by Dogs: Proof of Principle Study. *BMJ* **2004**, *329* (7468), 712. https://doi.org/10.1136/bmj.329.7468.712.

(63) Moser, E.; McCulloch, M. Canine Scent Detection of Human Cancers: A Review of Methods and Accuracy. *J. Vet. Behav.* **2010**, *5* (3), 145–152. https://doi.org/10.1016/j.jveb.2010.01.002.

(64) Dehlinger, K.; Tarnowski, K.; House, J. L.; Los, E.; Hanavan, K.; Bustamante, B.; Ahmann, A. J.; Ward, W. K. Can Trained Dogs Detect a Hypoglycemic Scent in Patients With Type 1 Diabetes? *Diabetes Care* **2013**, *36* (7), e98–e99. https://doi.org/10.2337/dc12-2342.

(65) Rooney, N. J.; Guest, C. M.; Swanson, L. C. M.; Morant, S. V. How Effective Are Trained Dogs at Alerting Their Owners to Changes in Blood Glycaemic Levels?: Variations in Performance of Glycaemia Alert Dogs. *PLOS ONE* **2019**, *14* (1), e0210092. https://doi.org/10.1371/journal.pone.0210092.

(66) Catala, A.; Grandgeorge, M.; Schaff, J.-L.; Cousillas, H.; Hausberger, M.; Cattet, J. Dogs Demonstrate the Existence of an Epileptic Seizure Odour in Humans. *Sci. Rep.* **2019**, *9* (1), 4103. https://doi.org/10.1038/s41598-019-40721-4.

(67) Crick, F. On Protein Synthesis. *Symp Soc Exp Biol* **1958**, *12* (138–63), 8.

(68) Kitano, H. Systems Biology: A Brief Overview. *Science* **2002**, *295* (5560), 1662–1664. https://doi.org/10.1126/science.1069492.

(69) Woo, S. L. C.; Lidsky, A. S.; Güttler, F.; Chandra, T.; Robson, K. J. H. Cloned Human Phenylalanine Hydroxylase Gene Allows Prenatal Diagnosis and Carrier Detection of Classical Phenylketonuria. *Nature* **1983**, *306* (5939), 151. https://doi.org/10.1038/306151a0.

(70) Hansen, T. v. O.; Jønson, L.; Albrechtsen, A.; Andersen, M. K.; Ejlertsen, B.; Nielsen, F. C. Large BRCA1 and BRCA2 Genomic Rearrangements in Danish High Risk Breast-Ovarian Cancer Families. *Breast Cancer Res. Treat.* **2009**, *115* (2), 315–323. https://doi.org/10.1007/s10549-008-0088-0.

(71) Yeager, M.; Orr, N.; Hayes, R. B.; Jacobs, K. B.; Kraft, P.; Wacholder, S.; Minichiello, M. J.; Fearnhead, P.; Yu, K.; Chatterjee, N.; et al. Genome-Wide Association Study of Prostate Cancer Identifies a Second Risk Locus at 8q24. *Nat. Genet.* **2007**, *39* (5), 645. https://doi.org/10.1038/ng2022.

(72) Williams, R. *Biochemical Individuality; the Basis for the Genetotrophic Concept*; Wiley: Oxford, England, 1956.

(73) Martin, A. J. P. Partition Chromatography. *Annu. Rev. Biochem.* **1950**, *19* (1), 517–542. https://doi.org/10.1146/annurev.bi.19.070150.002505.

(74) Horning, E. C.; Horning, M. G. Metabolic Profiles: Gas-Phase Methods for Analysis of Metabolites. *Clin. Chem.* **1971**, *17* (8), 802–809.

(75) Hoult, D. I.; Busby, S. J. W.; Gadian, D. G.; Radda, G. K.; Richards, R. E.; Seeley, P. J. Observation of Tissue Metabolites Using 31 P Nuclear Magnetic Resonance. *Nature* **1974**, *252* (5481), 285. https://doi.org/10.1038/252285a0.

(76) Cravatt, B. F.; Prospero-Garcia, O.; Siuzdak, G.; Gilula, N. B.; Henriksen, S. J.; Boger, D. L.; Lerner, R. A. Chemical Characterization of a Family of Brain Lipids That Induce Sleep. *Science* **1995**, *268* (5216), 1506–1509. https://doi.org/10.1126/science.7770779.

(77) Smith, C.; Maille, G.; Want, E.; Qin, C.; Trauger, S.; Brandon, T.; Custodio, D.; Abagyan, R.; Siuzdak, G. METLIN: A Metabolite Mass Spectral Database. *Ther. Drug Monit.* **2005**, *27* (6), 747–751. https://doi.org/10.1097/01.ftd.0000179845.53213.39.

(78) Zhang, F.; Zhang, Y.; Zhao, W.; Deng, K.; Wang, Z.; Yang, C.; Ma, L.; Openkova, M. S.; Hou, Y.; Li, K. Metabolomics for Biomarker Discovery in the Diagnosis, Prognosis, Survival and Recurrence of Colorectal

Cancer: A Systematic Review. *Oncotarget* **2017**, *8* (21), 35460–35472. https://doi.org/10.18632/oncotarget.16727.

(79) Yu, L.; Li, K.; Zhang, X. Next-Generation Metabolomics in Lung Cancer Diagnosis, Treatment and Precision Medicine: Mini Review. *Oncotarget* **2017**, *8* (70), 115774–115786. https://doi.org/10.18632/oncotarget.22404.

(80) Mikkonen, J. J. W.; Singh, S. P.; Herrala, M.; Lappalainen, R.; Myllymaa, S.; Kullaa, A. M. Salivary Metabolomics in the Diagnosis of Oral Cancer and Periodontal Diseases. *J. Periodontal Res.* **2016**, *51* (4), 431–437. https://doi.org/10.1111/jre.12327.

(81) Shin, J. M.; Kamarajan, P.; Fenno, J. C.; Rickard, A. H.; Kapila, Y. L. Metabolomics of Head and Neck Cancer: A Mini-Review. *Front. Physiol.* **2016**, *7*. https://doi.org/10.3389/fphys.2016.00526.

(82) Kelly, R. S.; Heiden, M. G. V.; Giovannucci, E.; Mucci, L. A. Metabolomic Biomarkers of Prostate Cancer: Prediction, Diagnosis, Progression, Prognosis, and Recurrence. *Cancer Epidemiol. Prev. Biomark.* **2016**, *25* (6), 887–906. https://doi.org/10.1158/1055-9965.EPI-15-1223.

(83) Cheng, Y.; Yang, X.; Deng, X.; Zhang, X.; Li, P.; Tao, J.; Qin, C.; Wei, J.; Lu, Q. Metabolomics in Bladder Cancer: A Systematic Review. *Int. J. Clin. Exp. Med.* **2015**, *8* (7), 11052–11063.

(84) Long, N. P.; Yoon, S. J.; Anh, N. H.; Nghi, T. D.; Lim, D. K.; Hong, Y. J.; Hong, S.-S.; Kwon, S. W. A Systematic Review on Metabolomics-Based Diagnostic Biomarker Discovery and Validation in Pancreatic Cancer. *Metabolomics* **2018**, *14* (8), 109. https://doi.org/10.1007/s11306-018-1404-2.

(85) Hadi, N. I.; Jamal, Q.; Iqbal, A.; Shaikh, F.; Somroo, S.; Musharraf, S. G. Serum Metabolomic Profiles for Breast Cancer Diagnosis, Grading and Staging by Gas Chromatography-Mass Spectrometry. *Sci. Rep.* **2017**, *7* (1), 1715. https://doi.org/10.1038/s41598-017-01924-9.

(86) Xie, J.; Zhang, A.; Wang, X. Metabolomic Applications in Hepatocellular Carcinoma: Toward the Exploration of Therapeutics and Diagnosis through Small Molecules. *RSC Adv.* **2017**, *7* (28), 17217–17226. https://doi.org/10.1039/C7RA00698E.

(87) Khatami, F.; Payab, M.; Sarvari, M.; Gilany, K.; Larijani, B.; Arjmand, B.; Tavangar, S. M. Oncometabolites as Biomarkers in Thyroid Cancer: A Systematic Review. *Cancer Manag. Res.* **2019**, *11*, 1829–1841. https://doi.org/10.2147/CMAR.S188661.

(88) Pallares-Méndez, R.; Aguilar-Salinas, C. A.; Cruz-Bautista, I.; Bosque-Plata, L. del. Metabolomics in Diabetes, a Review. *Ann. Med.* **2016**, *48* (1–2), 89–102. https://doi.org/10.3109/07853890.2015.1137630.

(89) de Sousa, E. B.; dos Santos Junior, G. C.; Duarte, M. E. L.; Moura Neto, V.; Aguiar, D. P.; de Sousa, E. B.; dos Santos Junior, G. C.; Duarte, M. E. L.; Moura Neto, V.; Aguiar, D. P. Metabolomics as a Promising Tool for Early Osteoarthritis Diagnosis. *Braz. J. Med. Biol. Res.* **2017**, *50* (11). https://doi.org/10.1590/1414-431x20176485.

(90) Checkley, W.; Deza, M. P.; Klawitter, J.; Romero, K. M.; Klawitter, J.; Pollard, S. L.; Wise, R. A.; Christians, U.; Hansel, N. N. Identifying Biomarkers for Asthma Diagnosis Using Targeted Metabolomics Approaches. *Respir. Med.* **2016**, *121*, 59–66. https://doi.org/10.1016/j.rmed.2016.10.011.

(91) Cavaleiro Rufo, J.; Madureira, J.; Oliveira Fernandes, E.; Moreira, A. Volatile Organic Compounds in Asthma Diagnosis: A Systematic Review and Meta-Analysis. *Allergy* **2016**, *71* (2), 175–188. https://doi.org/10.1111/all.12793.

(92) Lian, J.-S.; Liu, W.; Hao, S.-R.; Chen, D.-Y.; Wang, Y.-Y.; Yang, J.-L.; Jia, H.-Y.; Huang, J.-R. A Serum Metabolomic Analysis for Diagnosis and Biomarker Discovery of Primary Biliary Cirrhosis and Autoimmune Hepatitis. *Hepatobiliary Pancreat. Dis. Int.* **2015**, *14* (4), 413–421. https://doi.org/10.1016/S1499-3872(15)60393-9.

(93) Preter, V. D. Metabolomics in the Clinical Diagnosis of Inflammatory Bowel Disease. *Dig. Dis.* **2015**, *33* (Suppl. 1), 2–10. https://doi.org/10.1159/000437033.

(94) Ryan, D.; Newnham, E. D.; Prenzler, P. D.; Gibson, P. R. Metabolomics as a Tool for Diagnosis and Monitoring in Coeliac Disease. *Metabolomics* **2015**, *11* (4), 980–990. https://doi.org/10.1007/s11306-014-0752-9.

(95) Kao, D.; Ismond, K. P.; Tso, V.; Millan, B.; Hotte, N.; Fedorak, R. N. Urine-Based Metabolomic Analysis of Patients with Clostridium Difficile Infection: A Pilot Study. *Metabolomics* **2016**, *12* (8), 135. https://doi.org/10.1007/s11306-016-1080-z.

(96) Zetola, N. M.; Modongo, C.; Matsiri, O.; Tamuhla, T.; Mbongwe, B.; Matlhagela, K.; Sepako, E.; Catini, A.; Sirugo, G.; Martinelli, E.; et al. Diagnosis of Pulmonary Tuberculosis and Assessment of Treatment Response through Analyses of Volatile Compound Patterns in Exhaled Breath Samples. *J. Infect.* **2017**, *74* (4), 367–376. https://doi.org/10.1016/j.jinf.2016.12.006.

(97) Rabis, T.; Sommerwerck, U.; Anhenn, O.; Darwiche, K.; Freitag, L.; Teschler, H.; Bödeker, B.; Maddula,

S.; Baumbach, J. I. Detection of Infectious Agents in the Airways by Ion Mobility Spectrometry of Exhaled Breath. *Int. J. Ion Mobil. Spectrom.* **2011**, *14* (4), 187–195. https://doi.org/10.1007/s12127-011-0077-6.

(98) Ulanowska, A.; Kowalkowski, T.; Hrynkiewicz, K.; Jackowski, M.; Buszewski, B. Determination of Volatile Organic Compounds in Human Breath for Helicobacter Pylori Detection by SPME-GC/MS. *Biomed. Chromatogr.* **2011**, *25* (3), 391–397. https://doi.org/10.1002/bmc.1460.

(99) Lam, C.-W.; Law, C.-Y.; Sze, K.-H.; To, K. K.-W. Quantitative Metabolomics of Urine for Rapid Etiological Diagnosis of Urinary Tract Infection: Evaluation of a Microbial–Mammalian Co-Metabolite as a Diagnostic Biomarker. *Clin. Chim. Acta* **2015**, *438*, 24–28. https://doi.org/10.1016/j.cca.2014.07.038.

(100) Ogren, P. J.; Meetze, A.; Duer, W. C. The Limit of Detection in Generalized Least-Squares Calibrations: An Example Using Alprazolam Liquid Chromatography-Tandem Mass Spectrometry Data. *J. Anal. Toxicol.* **2009**, *33* (3), 129–142. https://doi.org/10.1093/jat/33.3.129.

(101) Guo, A. C.; Jewison, T.; Wilson, M.; Liu, Y.; Knox, C.; Djoumbou, Y.; Lo, P.; Mandal, R.; Krishnamurthy, R.; Wishart, D. S. ECMDB: The E. Coli Metabolome Database. *Nucleic Acids Res.* **2013**, *41* (D1), D625–D630. https://doi.org/10.1093/nar/gks992.

(102) Basuri, P.; Baidya, A.; Pradeep, T. Sub-Parts-per-Trillion Level Detection of Analytes by Superhydrophobic Preconcentration Paper Spray Ionization Mass Spectrometry (SHPPSI MS). *Anal. Chem.* **2019**, *91* (11), 7118–7124. https://doi.org/10.1021/acs.analchem.9b00144.

(103) Fan, M.; Andrade, G. F. S.; Brolo, A. G. A Review on the Fabrication of Substrates for Surface Enhanced Raman Spectroscopy and Their Applications in Analytical Chemistry. *Anal. Chim. Acta* **2011**, *693* (1), 7–25. https://doi.org/10.1016/j.aca.2011.03.002.

(104) Fleischmann, M.; Hendra, P. J.; McQuillan, A. J. Raman Spectra from Electrode Surfaces. *J. Chem. Soc. Chem. Commun.* **1973**, No. 3, 80–81. https://doi.org/10.1039/C39730000080.

(105) Fleischmann, M.; Hendra, P. J.; McQuillan, A. J. Raman Spectra of Pyridine Adsorbed at a Silver Electrode. *Chem. Phys. Lett.* **1974**, *26* (2), 163–166. https://doi.org/10.1016/0009-2614(74)85388-1.

(106) Jeanmaire, D. L.; Van Duyne, R. P. Surface Raman Spectroelectrochemistry: Part I. Heterocyclic, Aromatic, and Aliphatic Amines Adsorbed on the Anodized Silver Electrode. *J. Electroanal. Chem. Interfacial Electrochem.* **1977**, *84* (1), 1–20. https://doi.org/10.1016/S0022-0728(77)80224-6.

(107) Moskovits, M. Surface Roughness and the Enhanced Intensity of Raman Scattering by Molecules Adsorbed on Metals. *J. Chem. Phys.* **1978**, *69* (9), 4159–4161. https://doi.org/10.1063/1.437095.

(108) Kerker, M.; Siiman, O.; Bumm, L. A.; Wang, D.-S. Surface Enhanced Raman Scattering (SERS) of Citrate Ion Adsorbed on Colloidal Silver. *Appl. Opt.* **1980**, *19* (19), 3253. https://doi.org/10.1364/AO.19.003253.

(109) Koglin, E.; Séquaris, J.-M.; Fritz, J.-C.; Valenta, P. Surface Enhanced Raman Scattering (SERS) of Nucleic Acid Bases Adsorbed on Silver Colloids. *J. Mol. Struct.* **1984**, *114*, 219–223. https://doi.org/10.1016/0022-2860(84)87131-8.

(110) Kneipp, K.; Flemming, J. Surface Enhanced Raman Scattering (SERS) of Nucleic Acids Adsrobed on Colloidal Silver Particles. *J. Mol. Struct.* **1986**, *145* (1), 173–179. https://doi.org/10.1016/0022-2860(86)87041-7.

(111) Cotton, T. M.; Kim, J.-H.; Chumanov, G. D. Application of Surface-Enhanced Raman Spectroscopy to Biological Systems. *J. Raman Spectrosc.* **1991**, *22* (12), 729–742. https://doi.org/10.1002/jrs.1250221203.

(112) Suh, J. S.; Moskovits, M. Surface-Enhanced Raman Spectroscopy of Amino Acids and Nucleotide Bases Adsorbed on Silver. *J. Am. Chem. Soc.* **1986**, *108* (16), 4711–4718. https://doi.org/10.1021/ja00276a005.

(113) Herne, T. M.; Ahern, A.; Garrell, R. L. Surface-Enhanced Raman Spectroscopy of Peptides: Preferential N-Terminal Adsorption on Colloidal Silver. *J. Am. Chem. Soc.* **1991**, *113* (3), 846–854. https://doi.org/10.1021/ja00003a018.

(114) Ahern, A. M.; Garrell, R. L. Protein-Metal Interactions in Protein-Colloid Conjugates Probed by Surface-Enhanced Raman Spectroscopy. *Langmuir* **1991**, *7* (2), 254–261. https://doi.org/10.1021/la00050a009.

(115) Kneipp, K.; Wang, Y.; Dasari, R. R.; Feld, M. S.; Gilbert, B. D.; Janni, J.; Steinfeld, J. I. Near-Infrared Surface-Enhanced Raman Scattering of Trinitrotoluene on Colloidal Gold and Silver. *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* **1995**, *51* (12), 2171–2175. https://doi.org/10.1016/0584-8539(95)01474-7.

(116) Torres, E. L.; Winefordner, J. D. Trace Determination of Nitrogen-Containing Drugs by Surface Enhanced Raman Scattering Spectrometry on Silver Colloids. *Anal. Chem.* **1987**, *59* (13), 1626–1632. https://doi.org/10.1021/ac00140a010.

(117) Alak, A. M.; Vo-Dinh, Tuan. Surface-Enhanced Raman Spectrometry of Organo Phosphorus Chemical Agents. *Anal. Chem.* **1987**, *59* (17), 2149–2153. https://doi.org/10.1021/ac00144a030.

(118) Kneipp, K.; Wang, Y.; Dasari, R. R.; Feld, M. S. Near-Infrared Surface-Enhanced Raman Scattering (NIR-SERS) of Neurotransmitters in Colloidal Silver Solutions. *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.*

1995, *51* (3), 481–487. https://doi.org/10.1016/0584-8539(94)00235-4.

(119) Fang, Y.; Seong, N.-H.; Dlott, D. D. Measurement of the Distribution of Site Enhancements in Surface-Enhanced Raman Scattering. *Science* **2008**, *321* (5887), 388–392. https://doi.org/10.1126/science.1159499.

(120) Surface Enhanced Raman Spectroscopy of Individual Rhodamine 6G Molecules on Large Ag Nanocrystals | Journal of the American Chemical Society https://pubs.acs.org/doi/abs/10.1021/ja992128q (accessed Jul 30, 2019).

(121) Kneipp, K.; Wang, Y.; Kneipp, H.; Perelman, L. T.; Itzkan, I.; Dasari, R. R.; Feld, M. S. Single Molecule Detection Using Surface-Enhanced Raman Scattering (SERS). *Phys. Rev. Lett.* **1997**, *78* (9), 1667–1670. https://doi.org/10.1103/PhysRevLett.78.1667.

(122) Nie, S.; Emory, S. R. Probing Single Molecules and Single Nanoparticles by Surface-Enhanced Raman Scattering. *Science* **1997**, *275* (5303), 1102–1106. https://doi.org/10.1126/science.275.5303.1102.

(123) Ebbesen, T. W.; Lezec, H. J.; Ghaemi, H. F.; Thio, T.; Wolff, P. A. Extraordinary Optical Transmission through Sub-Wavelength Hole Arrays. *Nature* **1998**, *391* (6668), 667. https://doi.org/10.1038/35570.

(124) Maier, S. A.; Brongersma, M. L.; Kik, P. G.; Meltzer, S.; Requicha, A. a. G.; Atwater, H. A. Plasmonics—A Route to Nanoscale Optical Devices. *Adv. Mater.* **2001**, *13* (19), 1501–1505. https://doi.org/10.1002/1521-4095(200110)13:19<1501::AID-ADMA1501>3.0.CO;2-Z.

(125) Ozbay, E. Plasmonics: Merging Photonics and Electronics at Nanoscale Dimensions. *Science* **2006**, *311* (5758), 189–193. https://doi.org/10.1126/science.1114849.

(126) Prime, K. L.; Whitesides, G. M. Self-Assembled Organic Monolayers: Model Systems for Studying Adsorption of Proteins at Surfaces. *Science* **1991**, *252* (5009), 1164–1167.

(127) Freeman, R. G.; Grabar, K. C.; Allison, K. J.; Bright, R. M.; Davis, J. A.; Guthrie, A. P.; Hommer, M. B.; Jackson, M. A.; Smith, P. C.; Walter, D. G.; et al. Self-Assembled Metal Colloid Monolayers: An Approach to SERS Substrates. *Science* **1995**, *267* (5204), 1629–1632. https://doi.org/10.1126/science.267.5204.1629.

(128) Baker, G. A.; Moore, D. S. Progress in Plasmonic Engineering of Surface-Enhanced Raman-Scattering Substrates toward Ultra-Trace Analysis. *Anal. Bioanal. Chem.* **2005**, *382* (8), 1751–1770. https://doi.org/10.1007/s00216-005-3353-7.

(129) Driskell, J. D.; Lipert, R. J.; Porter, M. D. Labeled Gold Nanoparticles Immobilized at Smooth Metallic Substrates: Systematic Investigation of Surface Plasmon Resonance and Surface-Enhanced Raman Scattering. *J. Phys. Chem. B* **2006**, *110* (35), 17444–17451. https://doi.org/10.1021/jp0636930.

(130) Prodan, E.; Radloff, C.; Halas, N. J.; Nordlander, P. A Hybridization Model for the Plasmon Response of Complex Nanostructures. *Science* **2003**, *302* (5644), 419–422. https://doi.org/10.1126/science.1089171.

(131) Choi, J. H.; Adams, S. M.; Ragan, R. Design of a Versatile Chemical Assembly Method for Patterning Colloidal Nanoparticles. *Nanotechnology* **2009**, *20* (6), 065301. https://doi.org/10.1088/0957-4484/20/6/065301.

(132) Brongersma, M. L.; Hartman, J. W.; Atwater, H. A. Electromagnetic Energy Transfer and Switching in Nanoparticle Chain Arrays below the Diffraction Limit. *Phys. Rev. B* **2000**, *62* (24), R16356–R16359. https://doi.org/10.1103/PhysRevB.62.R16356.

(133) Kraus, T.; Malaquin, L.; Delamarche, E.; Schmid, H.; Spencer, N. D.; Wolf, H. Closing the Gap Between Self-Assembly and Microsystems Using Self-Assembly, Transfer, and Integration of Particles. *Adv. Mater.* **2005**, *17* (20), 2438–2442. https://doi.org/10.1002/adma.200501171.

(134) Henzie, J.; Lee, M. H.; Odom, T. W. Multiscale Patterning of Plasmonic Metamaterials. *Nat. Nanotechnol.* **2007**, *2* (9), 549–554. https://doi.org/10.1038/nnano.2007.252.

(135) Chiu, J. J.; Kim, B. J.; Kramer, E. J.; Pine, D. J. Control of Nanoparticle Location in Block Copolymers. *J. Am. Chem. Soc.* **2005**, *127* (14), 5036–5037. https://doi.org/10.1021/ja050376i.

(136) Zhang, Q.; Xu, T.; Butterfield, D.; Misner, M. J.; Ryu, D. Y.; Emrick, T.; Russell, T. P. Controlled Placement of CdSe Nanoparticles in Diblock Copolymer Templates by Electrophoretic Deposition. *Nano Lett.* **2005**, *5* (2), 357–361. https://doi.org/10.1021/nl048103t.

(137) Adams, S. M.; Campione, S.; Caldwell, J. D.; Bezares, F. J.; Culbertson, J. C.; Capolino, F.; Ragan, R. Non-Lithographic SERS Substrates: Tailoring Surface Chemistry for Au Nanoparticle Cluster Assembly. *Small* **2012**, *8* (14), 2239–2249. https://doi.org/10.1002/smll.201102708.

(138) Choi, W. M.; Park, O. O. The Fabrication of Micropatterns of a 2D Colloidal Assembly by Electrophoretic Deposition. *Nanotechnology* **2005**, *17* (1), 325–329. https://doi.org/10.1088/0957-4484/17/1/056.

(139) Patel, M. N.; Williams, R. D.; May, R. A.; Uchida, H.; Stevenson, K. J.; Johnston, K. P. Electrophoretic Deposition of Au Nanocrystals inside Perpendicular Mesochannels of TiO2. *Chem. Mater.* **2008**, *20* (19), 6029–6040. https://doi.org/10.1021/cm8012705.

(140) Grzelczak, M.; Vermant, J.; Furst, E. M.; Liz-Marzán, L. M. Directed Self-Assembly of Nanoparticles. *ACS*

Nano **2010**, *4* (7), 3591–3605. https://doi.org/10.1021/nn100869j.

(141) Adams, S. M.; Campione, S.; Capolino, F.; Ragan, R. Directing Cluster Formation of Au Nanoparticles from Colloidal Solution. *Langmuir* **2013**, *29* (13), 4242–4251. https://doi.org/10.1021/la3051719.

(142) Blackie, E. J.; Le Ru, E. C.; Etchegoin, P. G. Single-Molecule Surface-Enhanced Raman Spectroscopy of Nonresonant Molecules. *J. Am. Chem. Soc.* **2009**, *131* (40), 14466–14472. https://doi.org/10.1021/ja905319w.

(143) Brulé, T.; Bouhelier, A.; Yockell-Lelièvre, H.; Clément, J.-E.; Leray, A.; Dereux, A.; Finot, E. Statistical and Fourier Analysis for In-Line Concentration Sensitivity in Single Molecule Dynamic-SERS. *ACS Photonics* **2015**, *2* (9), 1266–1271. https://doi.org/10.1021/acsphotonics.5b00134.

(144) Durucan, O.; Rindzevicius, T.; Schmidt, M. S.; Matteucci, M.; Boisen, A. Nanopillar Filters for Surface-Enhanced Raman Spectroscopy. *ACS Sens.* **2017**, *2* (10), 1400–1404. https://doi.org/10.1021/acssensors.7b00499.

(145) Efrima, S.; Bronk, B. V. Silver Colloids Impregnating or Coating Bacteria. *J. Phys. Chem. B* **1998**, *102* (31), 5947–5950. https://doi.org/10.1021/jp9813903.

(146) Jarvis, R. M.; Goodacre, R. Discrimination of Bacteria Using Surface-Enhanced Raman Spectroscopy. *Anal. Chem.* **2004**, *76* (1), 40–47. https://doi.org/10.1021/ac034689c.

(147) Jarvis, R. M.; Law, N.; Shadi, I. T.; O'Brien, P.; Lloyd, J. R.; Goodacre, R. Surface-Enhanced Raman Scattering from Intracellular and Extracellular Bacterial Locations. *Anal. Chem.* **2008**, *80* (17), 6741–6746. https://doi.org/10.1021/ac800838v.

(148) Kahraman, M.; Yazıcı, M. M.; Şahin, F.; Çulha, M. Convective Assembly of Bacteria for Surface-Enhanced Raman Scattering. *Langmuir* **2008**, *24* (3), 894–901. https://doi.org/10.1021/la702240q.

(149) Mircescu, N. E.; Zhou, H.; Leopold, N.; Chiş, V.; Ivleva, N. P.; Niessner, R.; Wieser, A.; Haisch, C. Towards a Receptor-Free Immobilization and SERS Detection of Urinary Tract Infections Causative Pathogens. *Anal. Bioanal. Chem.* **2014**, *406* (13), 3051–3058. https://doi.org/10.1007/s00216-014-7761-4.

(150) Dina, N. E.; Zhou, H.; Colniţă, A.; Leopold, N.; Szoke-Nagy, T.; Coman, C.; Haisch, C. Rapid Single-Cell Detection and Identification of Pathogens by Using Surface-Enhanced Raman Spectroscopy. *Analyst* **2017**, *142* (10), 1782–1789. https://doi.org/10.1039/C7AN00106A.

(151) Premasiri, W. R.; Moir, D. T.; Klempner, M. S.; Krieger, N.; Jones, G.; Ziegler, L. D. Characterization of the Surface Enhanced Raman Scattering (SERS) of Bacteria. *J. Phys. Chem. B* **2005**, *109* (1), 312–320. https://doi.org/10.1021/jp040442n.

(152) Chu, H.; Huang, Y.; Zhao, Y. Silver Nanorod Arrays as a Surface-Enhanced Raman Scattering Substrate for Foodborne Pathogenic Bacteria Detection. *Appl. Spectrosc.* **2008**, *62* (8), 922–931. https://doi.org/10.1366/000370208785284330.

(153) Jia, H.; Wang, W.; Qiu, L.; Zhang, N.; Ge, H.; Wang, J. Fabricating a Long-Range Ordered 3D Bimetallic Nanoassembly with Edge-On Substrate for Highly Sensitive SERS Sensing of Escherichia Coli Bacteria. *Plasmonics* **2015**, *10* (6), 1889–1894. https://doi.org/10.1007/s11468-015-0012-5.

(154) Qiu, L.; Wang, W.; Zhang, A.; Zhang, N.; Lemma, T.; Ge, H.; Toppari, J. J.; Hytönen, V. P.; Wang, J. Core–Shell Nanorod Columnar Array Combined with Gold Nanoplate–Nanosphere Assemblies Enable Powerful In Situ SERS Detection of Bacteria. *ACS Appl. Mater. Interfaces* **2016**, *8* (37), 24394–24403. https://doi.org/10.1021/acsami.6b06674.

(155) Khan, S. A.; Singh, A. K.; Senapati, D.; Fan, Z.; Ray, P. C. Targeted Highly Sensitive Detection of Multi-Drug Resistant Salmonella DT104 Using Gold Nanoparticles. *Chem. Commun.* **2011**, *47* (33), 9444–9446. https://doi.org/10.1039/C1CC13199K.

(156) Zhang, H.; Ma, X.; Liu, Y.; Duan, N.; Wu, S.; Wang, Z.; Xu, B. Gold Nanoparticles Enhanced SERS Aptasensor for the Simultaneous Detection of Salmonella Typhimurium and Staphylococcus Aureus. *Biosens. Bioelectron.* **2015**, *74*, 872–877. https://doi.org/10.1016/j.bios.2015.07.033.

(157) Mitchell, B. G.; Ferguson, J. K.; Anderson, M.; Sear, J.; Barnett, A. Length of Stay and Mortality Associated with Healthcare-Associated Urinary Tract Infections: A Multi-State Model. *J. Hosp. Infect.* **2016**, *93* (1), 92–99. https://doi.org/10.1016/j.jhin.2016.01.012.

(158) Blango, M. G.; Mulvey, M. A. Persistence of Uropathogenic Escherichia Coli in the Face of Multiple Antibiotics. *Antimicrob. Agents Chemother.* **2010**, *54* (5), 1855–1863. https://doi.org/10.1128/AAC.00014-10.

(159) Avci, E.; Kaya, N. S.; Ucankus, G.; Culha, M. Discrimination of Urinary Tract Infection Pathogens by Means of Their Growth Profiles Using Surface Enhanced Raman Scattering. *Anal. Bioanal. Chem.* **2015**, *407* (27), 8233–8241. https://doi.org/10.1007/s00216-015-8950-5.

(160) Cheong, Y.; Kim, Y. J.; Kang, H.; Choi, S.; Lee, H. J. Rapid Label-Free Identification of Klebsiella

Pneumoniae Antibiotic Resistant Strains by the Drop-Coating Deposition Surface-Enhanced Raman Scattering Method. *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* **2017**, *183*, 53–59. https://doi.org/10.1016/j.saa.2017.04.044.

(161) Premasiri, W. R.; Lemler, P.; Chen, Y.; Gebregziabher, Y.; Ziegler, L. D. SERS Analysis of Bacteria, Human Blood, and Cancer Cells: A Metabolomic and Diagnostic Tool. In *Frontiers of Surface-Enhanced Raman Scattering*; John Wiley & Sons, Ltd, 2014; pp 257–283. https://doi.org/10.1002/9781118703601.ch12.

(162) Premasiri, W. R.; Lee, J. C.; Sauer-Budge, A.; Théberge, R.; Costello, C. E.; Ziegler, L. D. The Biochemical Origins of the Surface-Enhanced Raman Spectra of Bacteria: A Metabolomics Profiling by SERS. *Anal. Bioanal. Chem.* **2016**, *408* (17), 4631–4647. https://doi.org/10.1007/s00216-016-9540-x.

(163) Chen, Y.; Zhang, Y.; Pan, F.; Liu, J.; Wang, K.; Zhang, C.; Cheng, S.; Lu, L.; Zhang, W.; Zhang, Z.; et al. Breath Analysis Based on Surface-Enhanced Raman Scattering Sensors Distinguishes Early and Advanced Gastric Cancer Patients from Healthy Persons. *ACS Nano* **2016**, *10* (9), 8169–8179. https://doi.org/10.1021/acsnano.6b01441.

(164) Ali, M. R. K.; Wu, Y.; Han, T.; Zang, X.; Xiao, H.; Tang, Y.; Wu, R.; Fernández, F. M.; El-Sayed, M. A. Simultaneous Time-Dependent Surface-Enhanced Raman Spectroscopy, Metabolomics, and Proteomics Reveal Cancer Cell Death Mechanisms Associated with Gold Nanorod Photothermal Therapy. *J. Am. Chem. Soc.* **2016**, *138* (47), 15434–15442. https://doi.org/10.1021/jacs.6b08787.

(165) Shalabaeva, V.; Lovato, L.; Rocca, R. L.; Messina, G. C.; Dipalo, M.; Miele, E.; Perrone, M.; Gentile, F.; Angelis, F. D. Time Resolved and Label Free Monitoring of Extracellular Metabolites by Surface Enhanced Raman Spectroscopy. *PLOS ONE* **2017**, *12* (4), e0175581. https://doi.org/10.1371/journal.pone.0175581.

(166) Etchegoin, P. G.; Meyer, M.; Blackie, E.; Le Ru, E. C. Statistics of Single-Molecule Surface Enhanced Raman Scattering Signals: Fluctuation Analysis with Multiple Analyte Techniques. *Anal. Chem.* **2007**, *79* (21), 8411–8415. https://doi.org/10.1021/ac071231s.

(167) Feng, S.; Chen, R.; Lin, J.; Pan, J.; Chen, G.; Li, Y.; Cheng, M.; Huang, Z.; Chen, J.; Zeng, H. Nasopharyngeal Cancer Detection Based on Blood Plasma Surface-Enhanced Raman Spectroscopy and Multivariate Analysis. *Biosens. Bioelectron.* **2010**, *25* (11), 2414–2419. https://doi.org/10.1016/j.bios.2010.03.033.

(168) Lin, D.; Feng, S.; Pan, J.; Chen, Y.; Lin, J.; Chen, G.; Xie, S.; Zeng, H.; Chen, R. Colorectal Cancer Detection by Gold Nanoparticle Based Surface-Enhanced Raman Spectroscopy of Blood Serum and Statistical Analysis. *Opt. Express* **2011**, *19* (14), 13565–13577. https://doi.org/10.1364/OE.19.013565.

(169) Feng, S.; Chen, R.; Lin, J.; Pan, J.; Wu, Y.; Li, Y.; Chen, J.; Zeng, H. Gastric Cancer Detection Based on Blood Plasma Surface-Enhanced Raman Spectroscopy Excited by Polarized Laser Light. *Biosens. Bioelectron.* **2011**, *26* (7), 3167–3174. https://doi.org/10.1016/j.bios.2010.12.020.

(170) Feng, S.; Lin, D.; Lin, J.; Li, B.; Huang, Z.; Chen, G.; Zhang, W.; Wang, L.; Pan, J.; Chen, R.; et al. Blood Plasma Surface-Enhanced Raman Spectroscopy for Non-Invasive Optical Detection of Cervical Cancer. *Analyst* **2013**, *138* (14), 3967–3974. https://doi.org/10.1039/C3AN36890D.

(171) Lee, D.; Lee, S.; Seong, G. H.; Choo, J.; Lee, E. K.; Gweon, D.-G.; Lee, S. Quantitative Analysis of Methyl Parathion Pesticides in a Polydimethylsiloxane Microfluidic Channel Using Confocal Surface-Enhanced Raman Spectroscopy. *Appl. Spectrosc.* **2006**, *60* (4), 373–377. https://doi.org/10.1366/000370206776593762.

(172) Stevenson, R.; Ingram, A.; Leung, H.; C. McMillan, D.; Graham, D. Quantitative SERRS Immunoassay for the Detection of Human PSA. *Analyst* **2009**, *134* (5), 842–844. https://doi.org/10.1039/B902174D.

(173) Bodelón, G.; Montes-García, V.; López-Puente, V.; Hill, E. H.; Hamon, C.; Sanz-Ortiz, M. N.; Rodal-Cedeira, S.; Costas, C.; Celiksoy, S.; Pérez-Juste, I.; et al. Detection and Imaging of Quorum Sensing in *Pseudomonas Aeruginosa* Biofilm Communities by Surface-Enhanced Resonance Raman Scattering. *Nat. Mater.* **2016**, *15* (11), 1203–1211. https://doi.org/10.1038/nmat4720.

(174) Shafer-Peltier, K. E.; Haynes, C. L.; Glucksberg, M. R.; Van Duyne, R. P. Toward a Glucose Biosensor Based on Surface-Enhanced Raman Scattering. *J. Am. Chem. Soc.* **2003**, *125* (2), 588–593. https://doi.org/10.1021/ja028255v.

(175) Langmuir, I. THE ADSORPTION OF GASES ON PLANE SURFACES OF GLASS, MICA AND PLATINUM. *J. Am. Chem. Soc.* **1918**, *40* (9), 1361–1403. https://doi.org/10.1021/ja02242a004.

(176) Li, S.; Zhang, Y.; Xu, J.; Li, L.; Zeng, Q.; Lin, L.; Guo, Z.; Liu, Z.; Xiong, H.; Liu, S. Noninvasive Prostate Cancer Screening Based on Serum Surface-Enhanced Raman Spectroscopy and Support Vector Machine. *Appl. Phys. Lett.* **2014**, *105* (9), 091104. https://doi.org/10.1063/1.4892667.

(177) Dong, R.; Weng, S.; Yang, L.; Liu, J. Detection and Direct Readout of Drugs in Human Urine Using

Dynamic Surface-Enhanced Raman Spectroscopy and Support Vector Machines. *Anal. Chem.* **2015**, *87* (5), 2937–2944. https://doi.org/10.1021/acs.analchem.5b00137.

(178) Albuquerque, C. D. L.; Nogueira, R. B.; Poppi, R. J. Determination of 17β-Estradiol and Noradrenaline in Dog Serum Using Surface-Enhanced Raman Spectroscopy and Random Forest. *Microchem. J.* **2016**, *128*, 95–101. https://doi.org/10.1016/j.microc.2016.04.012.

(179) Huang, C.; Tsai, T.; Wen, B.; Chung, C.; Li, Y.; Chuang, Y.; Lin, W.; Li, L.; Wang, J.; Wang, Y.; et al. Hybrid SVM/CART Classification of Pathogenic Species of Bacterial Meningitis with Surface-Enhanced Raman Scattering. In *2010 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*; 2010; pp 406–409. https://doi.org/10.1109/BIBM.2010.5706600.

(180) E. Stephen, K.; Homrighausen, D.; DePalma, G.; H. Nakatsu, C.; Irudayaraj, J. Surface Enhanced Raman Spectroscopy (SERS) for the Discrimination of Arthrobacter Strains Based on Variations in Cell Surface Composition. *Analyst* **2012**, *137* (18), 4280–4286. https://doi.org/10.1039/C2AN35578G.

(181) Li, X.; Yang, T.; Li, S.; Wang, D.; Song, Y.; Yu, K. Different Classification Algorithms and Serum Surface Enhanced Raman Spectroscopy for Noninvasive Discrimination of Gastric Diseases. *J. Raman Spectrosc.* **2016**, *47* (8), 917–925. https://doi.org/10.1002/jrs.4924.

(182) Weng, S.; Li, M.; Chen, C.; Gao, X.; Zheng, S.; Zeng, X. Fast and Accurate Determination of Organophosphate Pesticides Using Surface-Enhanced Raman Scattering and Chemometrics. *Anal. Methods* **2015**, *7* (6), 2563–2567. https://doi.org/10.1039/C4AY03067B.

(183) Matschulat, A.; Drescher, D.; Kneipp, J. Surface-Enhanced Raman Scattering Hybrid Nanoprobe Multiplexing and Imaging in Biological Systems. *ACS Nano* **2010**, *4* (6), 3259–3269. https://doi.org/10.1021/nn100280z.

(184) Shanmukh, S.; Jones, L.; Zhao, Y.-P.; Driskell, J. D.; Tripp, R. A.; Dluhy, R. A. Identification and Classification of Respiratory Syncytial Virus (RSV) Strains by Surface-Enhanced Raman Spectroscopy and Multivariate Statistical Techniques. *Anal. Bioanal. Chem.* **2008**, *390* (6), 1551–1555. https://doi.org/10.1007/s00216-008-1851-0.

(185) Pearman, W. F.; Fountain, A. W. Classification of Chemical and Biological Warfare Agent Simulants by Surface-Enhanced Raman Spectroscopy and Multivariate Statistical Techniques. *Appl. Spectrosc.* **2006**, *60* (4), 356–365.

(186) Hu, Y.; Feng, S.; Gao, F.; Li-Chan, E. C. Y.; Grant, E.; Lu, X. Detection of Melamine in Milk Using Molecularly Imprinted Polymers–Surface Enhanced Raman Spectroscopy. *Food Chem.* **2015**, *176*, 123–129. https://doi.org/10.1016/j.foodchem.2014.12.051.

(187) Alharbi, O.; Xu, Y.; Goodacre, R. Simultaneous Multiplexed Quantification of Nicotine and Its Metabolites Using Surface Enhanced Raman Scattering. *Analyst* **2014**, *139* (19), 4820–4827. https://doi.org/10.1039/C4AN00879K.

(188) Lussier, F.; Missirlis, D.; Spatz, J. P.; Masson, J.-F. Machine-Learning-Driven Surface-Enhanced Raman Scattering Optophysiology Reveals Multiplexed Metabolite Gradients Near Cells. *ACS Nano* **2019**, *13* (2), 1403–1411. https://doi.org/10.1021/acsnano.8b07024.

(189) Jackson, J. D. *Classical Electrodynamics*; John Wiley & Sons, Inc., 2007.

(190) Raman Scattering. *Wikipedia*; 2017.

(191) Squires, T. M.; Bazant, M. Z. Induced-Charge Electro-Osmosis. *J. Fluid Mech.* **2004**, *509*, 217–252. https://doi.org/10.1017/S0022112004009309.

(192) Sides, P. J. Calculation of Electrohydrodynamic Flow around a Single Particle on an Electrode. *Langmuir* **2003**, *19* (7), 2745–2751. https://doi.org/10.1021/la026509k.

(193) Artificial Neuron. *Wikipedia*; 2019.

(194) Wasserman, L. *All of Statistics: A Concise Course in Statistical Inference*; Springer Science & Business Media, 2013.

(195) Werbos, P. J. Applications of Advances in Nonlinear Sensitivity Analysis. In *System Modeling and Optimization*; Drenick, R. F., Kozin, F., Eds.; Lecture Notes in Control and Information Sciences; Springer Berlin Heidelberg, 1982; pp 762–770.

(196) Murray, C. B.; Kagan, C. R.; Bawendi, M. G. Self-Organization of CdSe Nanocrystallites into Three-Dimensional Quantum Dot Superlattices. *Science* **1995**, *270* (5240), 1335–1338. https://doi.org/10.1126/science.270.5240.1335.

(197) Gaulding, E. A.; Diroll, B. T.; Goodwin, E. D.; Vrtis, Z. J.; Kagan, C. R.; Murray, C. B. Deposition of Wafer-Scale Single-Component and Binary Nanocrystal Superlattice Thin Films Via Dip-Coating. *Adv. Mater.* **2015**, *27* (18), 2846–2851. https://doi.org/10.1002/adma.201405575.

(198) Gong, J.; Newman, R. S.; Engel, M.; Zhao, M.; Bian, F.; Glotzer, S. C.; Tang, Z. Shape-Dependent Ordering

of Gold Nanocrystals into Large-Scale Superlattices. *Nat. Commun.* **2017**, *8*, 14038. https://doi.org/10.1038/ncomms14038.

(199) Vogel, N.; Retsch, M.; Fustin, C.-A.; del Campo, A.; Jonas, U. Advances in Colloidal Assembly: The Design of Structure and Hierarchy in Two and Three Dimensions. *Chem. Rev.* **2015**, *115* (13), 6265–6311. https://doi.org/10.1021/cr400081d.

(200) Ye, J.; Wen, F.; Sobhani, H.; Lassiter, J. B.; Van Dorpe, P.; Nordlander, P.; Halas, N. J. Plasmonic Nanoclusters: Near Field Properties of the Fano Resonance Interrogated with SERS. *Nano Lett.* **2012**, *12* (3), 1660–1667. https://doi.org/10.1021/nl3000453.

(201) Zhang, Y.; Wen, F.; Zhen, Y.-R.; Nordlander, P.; Halas, N. J. Coherent Fano Resonances in a Plasmonic Nanocluster Enhance Optical Four-Wave Mixing. *Proc. Natl. Acad. Sci.* **2013**, *110* (23), 9215–9219. https://doi.org/10.1073/pnas.1220304110.

(202) Campione, S.; Guclu, C.; Ragan, R.; Capolino, F. Fano Resonances in Metasurfaces Made of Linear Trimers of Plasmonic Nanoparticles. *Opt. Lett.* **2013**, *38* (24), 5216–5219. https://doi.org/10.1364/OL.38.005216.

(203) Campione, S.; Guclu, C.; Ragan, R.; Capolino, F. Enhanced Magnetic and Electric Fields via Fano Resonances in Metasurfaces of Circular Clusters of Plasmonic Nanoparticles. *ACS Photonics* **2014**, *1* (3), 254–260. https://doi.org/10.1021/ph4001313.

(204) Chikkaraddy, R.; de Nijs, B.; Benz, F.; Barrow, S. J.; Scherman, O. A.; Rosta, E.; Demetriadou, A.; Fox, P.; Hess, O.; Baumberg, J. J. Single-Molecule Strong Coupling at Room Temperature in Plasmonic Nanocavities. *Nature* **2016**, *535* (7610), 127–130. https://doi.org/10.1038/nature17974.

(205) Biswas, S.; Liu, X.; Jarrett, J. W.; Brown, D.; Pustovit, V.; Urbas, A.; Knappenberger, K. L.; Nealey, P. F.; Vaia, R. A. Nonlinear Chiro-Optical Amplification by Plasmonic Nanolens Arrays Formed via Directed Assembly of Gold Nanoparticles. *Nano Lett.* **2015**, *15* (3), 1836–1842. https://doi.org/10.1021/nl504613q.

(206) Hamon, C.; Liz-Marzán, L. M. Hierarchical Assembly of Plasmonic Nanoparticles. *Chem. – Eur. J.* **2015**, *21* (28), 9956–9963. https://doi.org/10.1002/chem.201500149.

(207) Nam, J.-M.; Oh, J.-W.; Lee, H.; Suh, Y. D. Plasmonic Nanogap-Enhanced Raman Scattering with Nanoparticles. *Acc. Chem. Res.* **2016**, *49* (12), 2746–2755. https://doi.org/10.1021/acs.accounts.6b00409.

(208) Greybush, N. J.; Liberal, I.; Malassis, L.; Kikkawa, J. M.; Engheta, N.; Murray, C. B.; Kagan, C. R. Plasmon Resonances in Self-Assembled Two-Dimensional Au Nanocrystal Metamolecules. *ACS Nano* **2017**, *11* (3), 2917–2927. https://doi.org/10.1021/acsnano.6b08189.

(209) Flauraud, V.; Mastrangeli, M.; Bernasconi, G. D.; Butet, J.; Alexander, D. T. L.; Shahrabi, E.; Martin, O. J. F.; Brugger, J. Nanoscale Topographical Control of Capillary Assembly of Nanoparticles. *Nat. Nanotechnol.* **2017**, *12* (1), 73–80. https://doi.org/10.1038/nnano.2016.179.

(210) Ni, S.; Klein, M. J. K.; Spencer, N. D.; Wolf, H. Cascaded Assembly of Complex Multiparticle Patterns. *Langmuir* **2014**, *30* (1), 90–95. https://doi.org/10.1021/la403956e.

(211) Fleck, N. A.; McMeeking, R. M.; Kraus, T. Convective Assembly of a Particle Monolayer. *Langmuir* **2015**, *31* (51), 13655–13663. https://doi.org/10.1021/acs.langmuir.5b03635.

(212) Ye, R.; Ye, Y.-H.; Zhou, Z.; Xu, H. Gravity-Assisted Convective Assembly of Centimeter-Sized Uniform Two-Dimensional Colloidal Crystals. *Langmuir* **2013**, *29* (6), 1796–1801. https://doi.org/10.1021/la3040227.

(213) Jaquay, E.; Martínez, L. J.; Huang, N.; Mejia, C. A.; Sarkar, D.; Povinelli, M. L. Light-Assisted, Templated Self-Assembly of Gold Nanoparticle Chains. *Nano Lett.* **2014**, *14* (9), 5184–5188. https://doi.org/10.1021/nl502083m.

(214) Oberdick, S. D.; Majetich, S. A. Electrophoretic Deposition of Iron Oxide Nanoparticles on Templates. *J. Phys. Chem. C* **2013**, *117* (36), 18709–18718. https://doi.org/10.1021/jp405395y.

(215) Yilmaz, C.; Cetin, A. E.; Goutzamanidis, G.; Huang, J.; Somu, S.; Altug, H.; Wei, D.; Busnaina, A. Three-Dimensional Crystalline and Homogeneous Metallic Nanostructures Using Directed Assembly of Nanoparticles. *ACS Nano* **2014**, *8* (5), 4547–4558. https://doi.org/10.1021/nn500084g.

(216) Work, A. H.; Williams, S. J. Characterization of 2D Colloids Assembled by Optically-Induced Electrohydrodynamics. *Soft Matter* **2015**, *11* (21), 4266–4272. https://doi.org/10.1039/C5SM00184F.

(217) Wang, K.-C.; Kumar, A.; J. Williams, S.; G. Green, N.; Chun Kim, K.; Chuang, H.-S. An Optoelectrokinetic Technique for Programmable Particle Manipulation and Bead-Based Biosignal Enhancement. *Lab. Chip* **2014**, *14* (20), 3958–3967. https://doi.org/10.1039/C4LC00661E.

(218) Ristenpart, W. D.; Jiang, P.; Slowik, M. A.; Punckt, C.; Saville, D. A.; Aksay, I. A. Electrohydrodynamic Flow and Colloidal Patterning near Inhomogeneities on Electrodes. *Langmuir* **2008**, *24* (21), 12172–12180. https://doi.org/10.1021/la801419k.

(219) Mirkin, C. A.; Letsinger, R. L.; Mucic, R. C.; Storhoff, J. J. A DNA-Based Method for Rationally

Assembling Nanoparticles into Macroscopic Materials. *Nature* **1996**, *382* (6592), 607–609. https://doi.org/10.1038/382607a0.

(220) Thacker, V. V.; Herrmann, L. O.; Sigle, D. O.; Zhang, T.; Liedl, T.; Baumberg, J. J.; Keyser, U. F. DNA Origami Based Assembly of Gold Nanoparticle Dimers for Surface-Enhanced Raman Scattering. *Nat. Commun.* **2014**, *5*. https://doi.org/10.1038/ncomms4448.

(221) Gür, F. N.; Schwarz, F. W.; Ye, J.; Diez, S.; Schmidt, T. L. Toward Self-Assembled Plasmonic Devices: High-Yield Arrangement of Gold Nanoparticles on DNA Origami Templates. *ACS Nano* **2016**, *10* (5), 5374–5382. https://doi.org/10.1021/acsnano.6b01537.

(222) Taylor, R. W.; Lee, T.-C.; Scherman, O. A.; Esteban, R.; Aizpurua, J.; Huang, F. M.; Baumberg, J. J.; Mahajan, S. Precise Subnanometer Plasmonic Junctions for SERS within Gold Nanoparticle Assemblies Using Cucurbit[n]Uril "Glue." *ACS Nano* **2011**, *5* (5), 3878–3887. https://doi.org/10.1021/nn200250v.

(223) Sigle, D. O.; Kasera, S.; Herrmann, L. O.; Palma, A.; de Nijs, B.; Benz, F.; Mahajan, S.; Baumberg, J. J.; Scherman, O. A. Observing Single Molecules Complexing with Cucurbit[7]Uril through Nanogap Surface-Enhanced Raman Spectroscopy. *J. Phys. Chem. Lett.* **2016**, *7* (4), 704–710. https://doi.org/10.1021/acs.jpclett.5b02535.

(224) Van Haute, D.; Longmate, J. M.; Berlin, J. M. Controlled Assembly of Biocompatible Metallic Nanoaggregates Using a Small Molecule Crosslinker. *Adv. Mater.* **2015**, *27* (35), 5158–5164. https://doi.org/10.1002/adma.201501602.

(225) Novak, J. P.; Feldheim, D. L. Assembly of Phenylacetylene-Bridged Silver and Gold Nanoparticle Arrays. *J. Am. Chem. Soc.* **2000**, *122* (16), 3979–3980. https://doi.org/10.1021/ja000477a.

(226) Sardar, R.; Heap, T. B.; Shumaker-Parry, J. S. Versatile Solid Phase Synthesis of Gold Nanoparticle Dimers Using an Asymmetric Functionalization Approach. *J. Am. Chem. Soc.* **2007**, *129* (17), 5356–5357. https://doi.org/10.1021/ja070933w.

(227) Zon, V. B.; Sachsenhauser, M.; Rant, U. Preparation of Gold Nanoparticle Dimers via Streptavidin-Induced Interlinking. *J. Nanoparticle Res.* **2013**, *15* (10), 1–10. https://doi.org/10.1007/s11051-013-1974-x.

(228) Kim, N. H.; Lee, S. J.; Moskovits, M. Reversible Tuning of SERS Hot Spots with Aptamers. *Adv. Mater.* **2011**, *23* (36), 4152–4156. https://doi.org/10.1002/adma.201101847.

(229) Zhang, L.; Dai, L.; Rong, Y.; Liu, Z.; Tong, D.; Huang, Y.; Chen, T. Light-Triggered Reversible Self-Assembly of Gold Nanoparticle Oligomers for Tunable SERS. *Langmuir* **2015**, *31* (3), 1164–1171. https://doi.org/10.1021/la504365b.

(230) Qian, X.; Li, J.; Nie, S. Stimuli-Responsive SERS Nanoparticles: Conformational Control of Plasmonic Coupling and Surface Raman Enhancement. *J. Am. Chem. Soc.* **2009**, *131* (22), 7540–7541. https://doi.org/10.1021/ja902226z.

(231) Zhu, M.-Q.; Wang, L.-Q.; Exarhos, G. J.; Li, A. D. Q. Thermosensitive Gold Nanoparticles. *J. Am. Chem. Soc.* **2004**, *126* (9), 2656–2657. https://doi.org/10.1021/ja038544z.

(232) Trau, M.; Saville, D. A.; Aksay, I. A. Assembly of Colloidal Crystals at Electrode Interfaces. *Langmuir* **1997**, *13* (24), 6375–6381. https://doi.org/10.1021/la970568u.

(233) Saini, S.; Bukosky, S. C.; Ristenpart, W. D. Influence of Electrolyte Concentration on the Aggregation of Colloidal Particles near Electrodes in Oscillatory Fields. *Langmuir* **2016**, *32* (17), 4210–4216. https://doi.org/10.1021/acs.langmuir.5b04636.

(234) Prieve, D. C.; Sides, P. J.; Wirth, C. L. 2-D Assembly of Colloidal Particles on a Planar Electrode. *Curr. Opin. Colloid Interface Sci.* **2010**, *15* (3), 160–174. https://doi.org/10.1016/j.cocis.2010.01.005.

(235) Kanipe, K. N.; Chidester, P. P. F.; Stucky, G. D.; Moskovits, M. Large Format Surface-Enhanced Raman Spectroscopy Substrate Optimized for Enhancement and Uniformity. *ACS Nano* **2016**, *10*, 7566–7571. https://doi.org/10.1021/acsnano.6b02564.

(236) Habermehl, A.; Liu, X.; Eschenbaum, C.; Lemmer, U. Fabrication of SERS Substrates by Roll-to-Roll Hot Embossing. In *Nano-Optics: Principles Enabling Basic Research and Applications*; NATO Science for Peace and Security Series B: Physics and Biophysics; Springer, Dordrecht, 2017; pp 513–515. https://doi.org/10.1007/978-94-024-0850-8_55.

(237) Nakajima, N.; Ikada, Y. Mechanism of Amide Formation by Carbodiimide for Bioconjugation in Aqueous Media. *Bioconjug. Chem.* **1995**, *6* (1), 123–130. https://doi.org/10.1021/bc00031a015.

(238) Solomentsev, Y.; Bevan, M.; Anderson, J. L. Aggregation Dynamics for Two Particles during Electrophoretic Deposition under Steady Fields. *Langmuir* **2000**, *16* (24), 9208–9216. https://doi.org/10.1021/la0005199.

(239) Yan, Q.; Zheng, H.-N.; Jiang, C.; Li, K.; Xiao, S.-J. EDC/NHS Activation Mechanism of Polymethacrylic Acid: Anhydride versus NHS-Ester. *RSC Adv.* **2015**, *5* (86), 69939–69947.

https://doi.org/10.1039/C5RA13844B.

(240) Frenkel, D.; Smit, B. *Understanding Molecular Simulation, Second Edition: From Algorithms to Applications*, 2 edition.; Academic Press: San Diego, 2001.

(241) Monti, S.; Carravetta, V.; Ågren, H. Simulation of Gold Functionalization with Cysteine by Reactive Molecular Dynamics. *J. Phys. Chem. Lett.* **2016**, *7* (2), 272–276. https://doi.org/10.1021/acs.jpclett.5b02769.

(242) van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A. ReaxFF: A Reactive Force Field for Hydrocarbons. *J. Phys. Chem. A* **2001**, *105* (41), 9396–9409. https://doi.org/10.1021/jp004368u.

(243) Abad, J. M.; Mertens, S. F. L.; Pita, M.; Fernández, V. M.; Schiffrin, D. J. Functionalization of Thioctic Acid-Capped Gold Nanoparticles for Specific Immobilization of Histidine-Tagged Proteins. *J. Am. Chem. Soc.* **2005**, *127* (15), 5689–5694. https://doi.org/10.1021/ja042717i.

(244) Henkelman, G.; Jónsson, H. Improved Tangent Estimate in the Nudged Elastic Band Method for Finding Minimum Energy Paths and Saddle Points. *J. Chem. Phys.* **2000**, *113* (22), 9978–9985. https://doi.org/10.1063/1.1323224.

(245) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. A Climbing Image Nudged Elastic Band Method for Finding Saddle Points and Minimum Energy Paths. *J. Chem. Phys.* **2000**, *113* (22), 9901–9904. https://doi.org/10.1063/1.1329672.

(246) Kennemur, J. G.; Novak, B. M. Advances in Polycarbodiimide Chemistry. *Polymer* **2011**, *52* (8), 1693–1710. https://doi.org/10.1016/j.polymer.2011.02.040.

(247) Sides, P. J. Electrohydrodynamic Particle Aggregation on an Electrode Driven by an Alternating Electric Field Normal to It. *Langmuir* **2001**, *17* (19), 5791–5800. https://doi.org/10.1021/la0105376.

(248) Russel, W. B.; Saville, D. A.; Schowalter, W. R. *Colloidal Dispersions*; Cambridge University Press, 1989.

(249) Darvishzadeh-Varcheie, M.; Guclu, C.; Ragan, R.; Boyraz, O.; Capolino, F. Electric Field Enhancement with Plasmonic Colloidal Nanoantennas Excited by a Silicon Nitride Waveguide. *Opt. Express* **2016**, *24* (25), 28337–28352. https://doi.org/10.1364/OE.24.028337.

(250) Sam, S.; Touahir, L.; Salvador Andresa, J.; Allongue, P.; Chazalviel, J.-N.; Gouget-Laemmel, A. C.; Henry de Villeneuve, C.; Moraillon, A.; Ozanam, F.; Gabouze, N.; et al. Semiquantitative Study of the EDC/NHS Activation of Acid Terminal Groups at Modified Porous Silicon Surfaces. *Langmuir* **2010**, *26* (2), 809–814. https://doi.org/10.1021/la902220a.

(251) Cai, W. B.; Ren, B.; Li, X. Q.; She, C. X.; Liu, F. M.; Cai, X. W.; Tian, Z. Q. Investigation of Surface-Enhanced Raman Scattering from Platinum Electrodes Using a Confocal Raman Microscope: Dependence of Surface Roughening Pretreatment. *Surf. Sci.* **1998**, *406* (1–3), 9–22. https://doi.org/10.1016/S0039-6028(97)01030-3.

(252) Aggarwal, R. L.; Farrar, L. W.; Diebold, E. D.; Polla, D. L. Measurement of the Absolute Raman Scattering Cross Section of the 1584-Cm−1 Band of Benzenethiol and the Surface-Enhanced Raman Scattering Cross Section Enhancement Factor for Femtosecond Laser-Nanostructured Substrates. *J. Raman Spectrosc.* **2009**, *40* (9), 1331–1333. https://doi.org/10.1002/jrs.2396.

(253) Grady, N. K.; Halas, N. J.; Nordlander, P. Influence of Dielectric Function Properties on the Optical Response of Plasmon Resonant Metallic Nanoparticles. *Chem. Phys. Lett.* **2004**, *399* (1–3), 167–171. https://doi.org/10.1016/j.cplett.2004.09.154.

(254) Moerland, R. J.; Hoogenboom, J. P. Subnanometer-Accuracy Optical Distance Ruler Based on Fluorescence Quenching by Transparent Conductors. *Optica* **2016**, *3* (2), 112–117. https://doi.org/10.1364/OPTICA.3.000112.

(255) Campione, S.; Adams, S. M.; Ragan, R.; Capolino, F. Comparison of Electric Field Enhancements: Linear and Triangular Oligomers versus Hexagonal Arrays of Plasmonic Nanospheres. *Opt. Express* **2013**, *21* (7), 7957–7973. https://doi.org/10.1364/OE.21.007957.

(256) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminformatics* **2012**, *4* (1), 17. https://doi.org/10.1186/1758-2946-4-17.

(257) Avogadro: an open-source molecular builder and visualization tool. Version 1.1.1 https://avogadro.cc/cite/.

(258) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114* (25), 10024–10035.

(259) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117* (1), 1–19. https://doi.org/10.1006/jcph.1995.1039.

(260) Nakano, A. A Space–Time-Ensemble Parallel Nudged Elastic Band Algorithm for Molecular Kinetics Simulation. *Comput. Phys. Commun.* **2008**, *178* (4), 280–289. https://doi.org/10.1016/j.cpc.2007.09.011.

(261) Burresi, M.; Oosten, D. van; Kampfrath, T.; Schoenmaker, H.; Heideman, R.; Leinse, A.; Kuipers, L. Probing the Magnetic Field of Light at Optical Frequencies. *Science* **2009**, *326* (5952), 550–553. https://doi.org/10.1126/science.1177096.

(262) Kasperczyk, M.; Person, S.; Ananias, D.; Carlos, L. D.; Novotny, L. Excitation of Magnetic Dipole Transitions at Optical Frequencies. *Phys. Rev. Lett.* **2015**, *114* (16), 163903. https://doi.org/10.1103/PhysRevLett.114.163903.

(263) Kamandi, M.; Guclu, C.; Luk, T. S.; Wang, G. T.; Capolino, F. Giant Field Enhancement in Longitudinal Epsilon-near-Zero Films. *Phys. Rev. B* **2017**, *95* (16), 161105. https://doi.org/10.1103/PhysRevB.95.161105.

(264) Klein, M. W.; Enkrich, C.; Wegener, M.; Soukoulis, C. M.; Linden, S. Single-Slit Split-Ring Resonators at Optical Frequencies: Limits of Size Scaling. *Opt. Lett.* **2006**, *31* (9), 1259. https://doi.org/10.1364/OL.31.001259.

(265) Liu, N.; Mukherjee, S.; Bao, K.; Li, Y.; Brown, L. V.; Nordlander, P.; Halas, N. J. Manipulating Magnetic Plasmon Propagation in Metallic Nanocluster Networks. *ACS Nano* **2012**, *6* (6), 5482–5488. https://doi.org/10.1021/nn301393x.

(266) Darvishzadeh-Varcheie, M.; Guclu, C.; Capolino, F. Magnetic Nanoantennas Made of Plasmonic Nanoclusters for Photoinduced Magnetic Field Enhancement. *Phys. Rev. Appl.* **2017**, *8* (2), 024033. https://doi.org/10.1103/PhysRevApplied.8.024033.

(267) Guclu, C.; Veysi, M.; Darvishzadeh-Varcheie, M.; Capolino, F. Artificial Magnetism via Nanoantennas under Azimuthally Polarized Vector Beam Illumination. In *Conference on Lasers and Electro-Optics (2016), paper JW2A.21*; Optical Society of America, 2016; p JW2A.21. https://doi.org/10.1364/CLEO_AT.2016.JW2A.21.

(268) Guclu, C.; Veysi, M.; Darvishzadeh-Vercheie, M.; Capolino, F. Optical Nanoantennas as Magnetic Nanoprobes for Enhancing Light-Matter Interaction. In *2016 10th International Congress on Advanced Electromagnetic Materials in Microwaves and Optics (METAMATERIALS)*; 2016; pp 391–393. https://doi.org/10.1109/MetaMaterials.2016.7746411.

(269) Nadal, F.; Argoul, F.; Hanusse, P.; Pouligny, B.; Ajdari, A. Electrically Induced Interactions between Colloidal Particles in the Vicinity of a Conducting Plane. *Phys. Rev. E* **2002**, *65* (6), 061409. https://doi.org/10.1103/PhysRevE.65.061409.

(270) Böhmer. In Situ Observation of 2-Dimensional Clustering during Electrophoretic Deposition. *Langmuir* **1996**, *12* (24), 5747–5750. https://doi.org/10.1021/la960183w.

(271) Wang, C.; Yan, Q.; Liu, H.-B.; Zhou, X.-H.; Xiao, S.-J. Different EDC/NHS Activation Mechanisms between PAA and PMAA Brushes and the Following Amidation Reactions. *Langmuir* **2011**, *27* (19), 12058–12068. https://doi.org/10.1021/la202267p.

(272) Thrift, W. J.; Nguyen, C. Q.; Darvishzadeh-Varcheie, M.; Zare, S.; Sharac, N.; Sanderson, R. N.; Dupper, T. J.; Hochbaum, A. I.; Capolino, F.; Abdolhosseini Qomi, M. J.; et al. Driving Chemical Reactions in Plasmonic Nanogaps with Electrohydrodynamic Flow. *ACS Nano* **2017**, *11* (11), 11317–11329. https://doi.org/10.1021/acsnano.7b05815.

(273) Nguyen, C. Q.; Thrift, W. J.; Bhattacharjee, A.; Ranjbar, S.; Gallagher, T.; Darvishzadeh-Varcheie, M.; Sanderson, R. N.; Capolino, F.; Whiteson, K.; Baldi, P.; et al. Longitudinal Monitoring of Biofilm Formation via Robust Surface-Enhanced Raman Scattering Quantification of Pseudomonas Aeruginosa-Produced Metabolites. *ACS Appl. Mater. Interfaces* **2018**, *10* (15), 12364–12373. https://doi.org/10.1021/acsami.7b18592.

(274) Thrift, W.; Bhattacharjee, A.; Darvishzadeh-Varcheie, M.; Lu, Y.; Hochbaum, A.; Capolino, F.; Whiteson, K.; Ragan, R. Surface Enhanced Raman Scattering for Detection of Pseudomonas Aeruginosa Quorum Sensing Compounds. In *Biosensing and Nanomedicine VIII*; 2015; Vol. 9550, pp 95500B-95500B – 13. https://doi.org/10.1117/12.2188806.

(275) Bastús, N. G.; Comenge, J.; Puntes, V. Kinetically Controlled Seeded Growth Synthesis of Citrate-Stabilized Gold Nanoparticles of up to 200 Nm: Size Focusing versus Ostwald Ripening. *Langmuir* **2011**, *27* (17), 11098–11105. https://doi.org/10.1021/la201938u.

(276) Kievit, T. R. D.; Gillis, R.; Marx, S.; Brown, C.; Iglewski, B. H. Quorum-Sensing Genes in Pseudomonas Aeruginosa Biofilms: Their Role and Expression Patterns. *Appl. Environ. Microbiol.* **2001**, *67* (4), 1865–1873. https://doi.org/10.1128/AEM.67.4.1865-1873.2001.

(277) Kolter, R.; Greenberg, E. P. Microbial Sciences: The Superficial Life of Microbes. *Nature* **2006**, *441* (7091), 300–302. https://doi.org/10.1038/441300a.

(278) Anderson, G. G.; O'Toole, G. A. Innate and Induced Resistance Mechanisms of Bacterial Biofilms. In *Bacterial Biofilms*; Current Topics in Microbiology and Immunology; Springer, Berlin, Heidelberg, 2008;

pp 85–105. https://doi.org/10.1007/978-3-540-75418-3_5.

(279) Malone, M.; Goeres, D. M.; Gosbell, I.; Vickery, K.; Jensen, S.; Stoodley, P. Approaches to Biofilm-Associated Infections: The Need for Standardized and Relevant Biofilm Methods for Clinical Applications. *Expert Rev. Anti Infect. Ther.* **2017**, *15* (2), 147–156. https://doi.org/10.1080/14787210.2017.1262257.

(280) Hunter, R. C.; Klepac-Ceraj, V.; Lorenzi, M. M.; Grotzinger, H.; Martin, T. R.; Newman, D. K. Phenazine Content in the Cystic Fibrosis Respiratory Tract Negatively Correlates with Lung Function and Microbial Complexity. *Am. J. Respir. Cell Mol. Biol.* **2012**, *47* (6), 738–745. https://doi.org/10.1165/rcmb.2012-0088OC.

(281) Whiteson, K. L.; Meinardi, S.; Lim, Y. W.; Schmieder, R.; Maughan, H.; Quinn, R.; Blake, D. R.; Conrad, D.; Rohwer, F. Breath Gas Metabolites and Bacterial Metagenomes from Cystic Fibrosis Airways Indicate Active PH Neutral 2,3-Butanedione Fermentation. *ISME J.* **2014**, *8* (6), 1247–1258. https://doi.org/10.1038/ismej.2013.229.

(282) Serra, R.; Grande, R.; Butrico, L.; Rossi, A.; Settimio, U. F.; Caroleo, B.; Amato, B.; Gallelli, L.; Franciscis, S. de. Chronic Wound Infections: The Role of Pseudomonas Aeruginosa and Staphylococcus Aureus. *Expert Rev. Anti Infect. Ther.* **2015**, *13* (5), 605–613. https://doi.org/10.1586/14787210.2015.1023291.

(283) Davies, J. C. Pseudomonas Aeruginosa in Cystic Fibrosis: Pathogenesis and Persistence. *Paediatr. Respir. Rev.* **2002**, *3* (2), 128–134. https://doi.org/10.1016/S1526-0550(02)00003-3.

(284) Lorenz, B.; Wichmann, C.; Stöckel, S.; Rösch, P.; Popp, J. Cultivation-Free Raman Spectroscopic Investigations of Bacteria. *Trends Microbiol.* **2017**, *25* (5), 413–424. https://doi.org/10.1016/j.tim.2017.01.002.

(285) Wu, X.; Huang, Y.-W.; Park, B.; Tripp, R. A.; Zhao, Y. Differentiation and Classification of Bacteria Using Vancomycin Functionalized Silver Nanorods Array Based Surface-Enhanced Raman Spectroscopy and Chemometric Analysis. *Talanta* **2015**, *139* (Supplement C), 96–103. https://doi.org/10.1016/j.talanta.2015.02.045.

(286) Berry, D.; Mader, E.; Lee, T. K.; Woebken, D.; Wang, Y.; Zhu, D.; Palatinszky, M.; Schintlmeister, A.; Schmid, M. C.; Hanson, B. T.; et al. Tracking Heavy Water (D2O) Incorporation for Identifying and Sorting Active Microbial Cells. *Proc. Natl. Acad. Sci.* **2015**, *112* (2), E194–E203. https://doi.org/10.1073/pnas.1420406112.

(287) Bowden, J. A.; Heckert, A.; Ulmer, C. Z.; Jones, C. M.; Koelmel, J. P.; Abdullah, L.; Ahonen, L.; Alnouti, Y.; Armando, A.; Asara, J. M.; et al. Harmonizing Lipidomics: NIST Interlaboratory Comparison Exercise for Lipidomics Using Standard Reference Material 1950 Metabolites in Frozen Human Plasma. *J. Lipid Res.* **2017**. https://doi.org/10.1194/jlr.M079012.

(288) Aksenov, A. A.; Silva, R. da; Knight, R.; Lopes, N. P.; Dorrestein, P. C. Global Chemical Analysis of Biology by Mass Spectrometry. *Nat. Rev. Chem.* **2017**, *1* (7), s41570-017-0054–017. https://doi.org/10.1038/s41570-017-0054.

(289) Fessenden, M. Metabolomics: Small Molecules, Single Cells. *Nature* **2016**, *540* (7631), 540153a. https://doi.org/10.1038/540153a.

(290) Yu, Z.; Kastenmüller, G.; He, Y.; Belcredi, P.; Möller, G.; Prehn, C.; Mendes, J.; Wahl, S.; Roemisch-Margl, W.; Ceglarek, U.; et al. Differences between Human Plasma and Serum Metabolite Profiles. *PLOS ONE* **2011**, *6* (7), e21230. https://doi.org/10.1371/journal.pone.0021230.

(291) Keleştemur, S.; Çulha, M. Understanding and Discrimination of Biofilms of Clinically Relevant Microorganisms Using Surface-Enhanced Raman Scattering. *Appl. Spectrosc.* **2017**, *71* (6), 1180–1188. https://doi.org/10.1177/0003702816670916.

(292) Žukovskaja, O.; Jahn, I. J.; Weber, K.; Cialla-May, D.; Popp, J. Detection of Pseudomonas Aeruginosa Metabolite Pyocyanin in Water and Saliva by Employing the SERS Technique. *Sensors* **2017**, *17* (8), 1704. https://doi.org/10.3390/s17081704.

(293) Bodelón, G.; Montes-García, V.; López-Puente, V.; Hill, E. H.; Hamon, C.; Sanz-Ortiz, M. N.; Rodal-Cedeira, S.; Costas, C.; Celiksoy, S.; Pérez-Juste, I.; et al. Detection and Imaging of Quorum Sensing in Pseudomonas Aeruginosa Biofilm Communities by Surface-Enhanced Resonance Raman Scattering. *Nat. Mater.* **2016**, *15* (11), 1203–1211. https://doi.org/10.1038/nmat4720.

(294) Lee, J.-H.; Nam, J.-M.; Jeon, K.-S.; Lim, D.-K.; Kim, H.; Kwon, S.; Lee, H.; Suh, Y. D. Tuning and Maximizing the Single-Molecule Surface-Enhanced Raman Scattering from DNA-Tethered Nanodumbbells. *ACS Nano* **2012**, *6* (11), 9574–9584. https://doi.org/10.1021/nn3028216.

(295) Zhu, W.; Crozier, K. B. Quantum Mechanical Limit to Plasmonic Enhancement as Observed by Surface-Enhanced Raman Scattering. *Nat. Commun.* **2014**, *5*, 5228. https://doi.org/10.1038/ncomms6228.

(296) Le Ru, E. C.; Grand, J.; Sow, I.; Somerville, W. R. C.; Etchegoin, P. G.; Treguer-Delapierre, M.; Charron,

G.; Félidj, N.; Lévi, G.; Aubard, J. A Scheme for Detecting Every Single Target Molecule with Surface-Enhanced Raman Spectroscopy. *Nano Lett.* **2011**, *11* (11), 5013–5019. https://doi.org/10.1021/nl2030344.

(297) Le Ru, E. C.; Meyer, M.; Etchegoin, P. G. Proof of Single-Molecule Sensitivity in Surface Enhanced Raman Scattering (SERS) by Means of a Two-Analyte Technique. *J. Phys. Chem. B* **2006**, *110* (4), 1944–1948. https://doi.org/10.1021/jp054732v.

(298) Zrimsek, A. B.; Wong, N. L.; Van Duyne, R. P. Single Molecule Surface-Enhanced Raman Spectroscopy: A Critical Analysis of the Bianalyte versus Isotopologue Proof. *J. Phys. Chem. C* **2016**, *120* (9), 5133–5142. https://doi.org/10.1021/acs.jpcc.6b00606.

(299) Woehl, T. J.; Heatley, K. L.; Dutcher, C. S.; Talken, N. H.; Ristenpart, W. D. Electrolyte-Dependent Aggregation of Colloidal Particles near Electrodes in Oscillatory Electric Fields. *Langmuir* **2014**, *30* (17), 4887–4894. https://doi.org/10.1021/la4048243.

(300) Lyandres, O.; Shah, N. C.; Yonzon, C. R.; Walsh, J. T.; Glucksberg, M. R.; Van Duyne, R. P. Real-Time Glucose Sensing by Surface-Enhanced Raman Spectroscopy in Bovine Plasma Facilitated by a Mixed Decanethiol/Mercaptohexanol Partition Layer. *Anal. Chem.* **2005**, *77* (19), 6134–6139. https://doi.org/10.1021/ac051357u.

(301) Cheung, W.; Shadi, I. T.; Xu, Y.; Goodacre, R. Quantitative Analysis of the Banned Food Dye Sudan-1 Using Surface Enhanced Raman Scattering with Multivariate Chemometrics. *J. Phys. Chem. C* **2010**, *114* (16), 7285–7290. https://doi.org/10.1021/jp908892n.

(302) Watson, D.; MacDERMOT, J.; Wilson, R.; Cole, P. J.; Taylor, G. W. Purification and Structural Analysis of Pyocyanin and 1-Hydroxyphenazine. *Eur. J. Biochem.* **1986**, *159* (2), 309–313. https://doi.org/10.1111/j.1432-1033.1986.tb09869.x.

(303) Wu, X.; Chen, J.; Li, X.; Zhao, Y.; Zughaier, S. M. Culture-Free Diagnostics of Pseudomonas Aeruginosa Infection by Silver Nanorod Array Based SERS from Clinical Sputum Samples. *Nanomedicine Nanotechnol. Biol. Med.* **2014**, *10* (8), 1863–1870. https://doi.org/10.1016/j.nano.2014.04.010.

(304) Armbruster, D. A.; Pry, T. Limit of Blank, Limit of Detection and Limit of Quantitation. *Clin. Biochem. Rev.* **2008**, *29* (Suppl 1), S49–S52.

(305) Qi, G.; Jia, K.; Fu, C.; Xu, S.; Xu, W. A Highly Sensitive SERS Sensor for Quantitative Analysis of Glucose Based on the Chemical Etching of Silver Nanoparticles. *J. Opt.* **2015**, *17* (11), 114020. https://doi.org/10.1088/2040-8978/17/11/114020.

(306) Grubisha, D. S.; Lipert, R. J.; Park, H.-Y.; Driskell, J.; Porter, M. D. Femtomolar Detection of Prostate-Specific Antigen: An Immunoassay Based on Surface-Enhanced Raman Scattering and Immunogold Labels. *Anal. Chem.* **2003**, *75* (21), 5936–5943. https://doi.org/10.1021/ac034356f.

(307) Srichan, C.; Ekpanyapong, M.; Horprathum, M.; Eiamchai, P.; Nuntawong, N.; Phokharatkul, D.; Danvirutai, P.; Bohez, E.; Wisitsoraat, A.; Tuantranont, A. Highly-Sensitive Surface-Enhanced Raman Spectroscopy (SERS)-Based Chemical Sensor Using 3D Graphene Foam Decorated with Silver Nanoparticles as SERS Substrate. *Sci. Rep.* **2016**, *6*, 23733. https://doi.org/10.1038/srep23733.

(308) Nguyen, C.; Thrift, W.; Bhattacharjee, A.; Whiteson, K.; Hochbaum, A.; Ragan, R. Robust SERS Spectral Analysis for Quantitative Detection of Pycocyanin in Biological Fluids. In *Biosensing and Nanomedicine X*; International Society for Optics and Photonics, 2017; Vol. 10352, p 1035205. https://doi.org/10.1117/12.2267958.

(309) Reszka, K. J.; O'Malley, Y.; McCormick, M. L.; Denning, G. M.; Britigan, B. E. Oxidation of Pyocyanin, a Cytotoxic Product from Pseudomonas Aeruginosa, by Microperoxidase 11 and Hydrogen Peroxide. *Free Radic. Biol. Med.* **2004**, *36* (11), 1448–1459. https://doi.org/10.1016/j.freeradbiomed.2004.03.011.

(310) Dietrich, L. E. P.; Price-Whelan, A.; Petersen, A.; Whiteley, M.; Newman, D. K. The Phenazine Pyocyanin Is a Terminal Signalling Factor in the Quorum Sensing Network of Pseudomonas Aeruginosa. *Mol. Microbiol.* **2006**, *61* (5), 1308–1321. https://doi.org/10.1111/j.1365-2958.2006.05306.x.

(311) Bhattacharjee, A.; Khan, M.; Kleiman, M.; Hochbaum, A. I. Effects of Growth Surface Topography on Bacterial Signaling in Coculture Biofilms. *ACS Appl. Mater. Interfaces* **2017**, *9* (22), 18531–18539. https://doi.org/10.1021/acsami.7b04223.

(312) Rahme, L. G.; Stevens, E. J.; Wolfort, S. F.; Shao, J.; Tompkins, R. G.; Ausubel, F. M. Common Virulence Factors for Bacterial Pathogenicity in Plants and Animals. *Science* **1995**, *268* (5219), 1899–1902. https://doi.org/10.1126/science.7604262.

(313) Monahan, L. G.; Turnbull, L.; Osvath, S. R.; Birch, D.; Charles, I. G.; Whitchurch, C. B. Rapid Conversion of Pseudomonas Aeruginosa to a Spherical Cell Morphotype Facilitates Tolerance to Carbapenems and Penicillins but Increases Susceptibility to Antimicrobial Peptides. *Antimicrob. Agents Chemother.* **2014**, *58* (4), 1956–1962. https://doi.org/10.1128/AAC.01901-13.

(314) Parker, J. H.; Feldman, D. W.; Ashkin, M. Raman Scattering by Silicon and Germanium. *Phys. Rev.* **1967**, *155* (3), 712–714. https://doi.org/10.1103/PhysRev.155.712.

(315) Geladi, P.; Kowalski, B. R. Partial Least-Squares Regression: A Tutorial. *Anal. Chim. Acta* **1986**, *185*, 1–17. https://doi.org/10.1016/0003-2670(86)80028-9.

(316) Grady, N. K.; Halas, N. J.; Nordlander, P. Influence of Dielectric Function Properties on the Optical Response of Plasmon Resonant Metallic Nanoparticles. *Chem. Phys. Lett.* **2004**, *399* (1), 167–171. https://doi.org/10.1016/j.cplett.2004.09.154.

(317) Conley, S.; Franco, G.; Faloona, I.; Blake, D. R.; Peischl, J.; Ryerson, T. B. Methane Emissions from the 2015 Aliso Canyon Blowout in Los Angeles, CA. *Science* **2016**, aaf2348. https://doi.org/10.1126/science.aaf2348.

(318) Rozas, R.; Morales, J.; Vega, D. Artificial Smell Detection for Robotic Navigation. In *Fifth International Conference on Advanced Robotics 'Robots in Unstructured Environments*; 1991; pp 1730–1733 vol.2. https://doi.org/10.1109/ICAR.1991.240354.

(319) Nakamoto, T.; Ishida, H.; Moriizumi, T. An Odor Compass for Localizing an Odor Source. *Sens. Actuators B Chem.* **1996**, *35* (1), 32–36. https://doi.org/10.1016/S0925-4005(96)02009-6.

(320) Ishida, H.; Kobayashi, A.; Nakamoto, T.; Moriizumi, T. Three-Dimensional Odor Compass. *IEEE Trans. Robot. Autom.* **1999**, *15* (2), 251–257. https://doi.org/10.1109/70.760346.

(321) Ishida, H.; Wada, Y.; Matsukura, H. Chemical Sensing in Robotic Applications: A Review. *IEEE Sens. J.* **2012**, *12* (11), 3163–3173. https://doi.org/10.1109/JSEN.2012.2208740.

(322) Kowadlo, G.; Russell, R. A. Robot Odor Localization: A Taxonomy and Survey. *Int. J. Robot. Res.* **2008**, *27* (8), 869–894. https://doi.org/10.1177/0278364908095118.

(323) Braitenberg, V. *Vehicles: Experiments in Synthetic Psychology*; MIT Press, 1986.

(324) de Croon, G. C. H. E.; O'Connor, L. M.; Nicol, C.; Izzo, D. Evolutionary Robotics Approach to Odor Source Localization. *Neurocomputing* **2013**, *121*, 481–497. https://doi.org/10.1016/j.neucom.2013.05.028.

(325) Villarreal, B. L.; Olague, G.; Gordillo, J. L. Synthesis of Odor Tracking Algorithms with Genetic Programming. *Neurocomputing* **2016**, *175*, 1019–1032. https://doi.org/10.1016/j.neucom.2015.09.108.

(326) Jiang, P.; Wang, Y.; Ge, A. Multivariable Fuzzy Control Based Mobile Robot Odor Source Localization via Semitensor Product https://www.hindawi.com/journals/mpe/2015/736720/abs/ (accessed Oct 24, 2018). https://doi.org/10.1155/2015/736720.

(327) Husni, N. L.; Handayani, A. S.; Nurmaini, S.; Yani, I. Odor Localization Using Gas Sensor for Mobile Robot. In *2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*; 2017; pp 1–6. https://doi.org/10.1109/EECSI.2017.8239176.

(328) Vergassola, M.; Villermaux, E.; Shraiman, B. I. 'Infotaxis' as a Strategy for Searching without Gradients. *Nature* **2007**, *445* (7126), 406–409. https://doi.org/10.1038/nature05464.

(329) Luo, B.; Meng, Q.; Wang, J.; Zeng, M. A Flying Odor Compass to Autonomously Locate the Gas Source. *IEEE Trans. Instrum. Meas.* **2018**, *67* (1), 137–149. https://doi.org/10.1109/TIM.2017.2759378.

(330) Neumann, P. P.; Bennetts, V. H.; Lilienthal, A. J.; Bartholmai, M.; Schiller, J. H. Gas Source Localization with a Micro-Drone Using Bio-Inspired and Particle Filter-Based Algorithms. *Adv. Robot.* **2013**, *27* (9), 725–738. https://doi.org/10.1080/01691864.2013.779052.

(331) Grasso, F. W.; Consi, T. R.; Mountain, D. C.; Atema, J. Biomimetic Robot Lobster Performs Chemo-Orientation in Turbulence Using a Pair of Spatially Separated Sensors: Progress and Challenges. *Robot. Auton. Syst.* **2000**, *30* (1), 115–131. https://doi.org/10.1016/S0921-8890(99)00068-8.

(332) Consi, T. R.; Atema, J.; Goudey, C. A.; Cho, J.; Chryssostomidis, C. AUV Guidance with Chemical Signals. In *Proceedings of IEEE Symposium on Autonomous Underwater Vehicle Technology (AUV'94)*; 1994; pp 450–455. https://doi.org/10.1109/AUV.1994.518659.

(333) Miyatani, I.; Ishida, H. Active Stereo Nose: Using Air Curtain to Enhance the Directivity. In *2010 IEEE SENSORS*; 2010; pp 1522–1525. https://doi.org/10.1109/ICSENS.2010.5690333.

(334) Duckett, T.; Axelsson, M.; Saffiotti, A. Learning to Locate an Odour Source with a Mobile Robot. In *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No.01CH37164)*; 2001; Vol. 4, pp 4017–4022 vol.4. https://doi.org/10.1109/ROBOT.2001.933245.

(335) Kazadi, S.; Goodman, R.; Tsikata, D.; Green, D.; Lin, H. An Autonomous Water Vapor Plume Tracking Robot Using Passive Resistive Polymer Sensors. *Auton. Robots* **2000**, *9* (2), 175–188. https://doi.org/10.1023/A:1008970418316.

(336) Hayes, A. T.; Martinoli, A.; Goodman, R. M. Distributed Odor Source Localization. *IEEE Sens. J.* **2002**, *2* (3), 260–271. https://doi.org/10.1109/JSEN.2002.800682.

(337) Ishida, H.; Tokuhiro, T.; Nakamoto, T.; Moriizumi, T. Improvement of Olfactory Video Camera: Gas/Odor

Flow Visualization System. *Sens. Actuators B Chem.* **2002**, *83* (1), 256–261. https://doi.org/10.1016/S0925-4005(01)01058-9.

(338) Deveza, R.; Thiel, D.; Russell, A.; Mackay-Sim, A. Odor Sensing for Robot Guidance. *Int. J. Robot. Res.* **1994**, *13* (3), 232–239. https://doi.org/10.1177/027836499401300305.

(339) Kuwana, Y.; Nagasawa, S.; Shimoyama, I.; Kanzaki, R. Synthesis of the Pheromone-Oriented Behaviour of Silkworm Moths by a Mobile Robot with Moth Antennae as Pheromone Sensors1This Paper Was Presented at the Fifth World Congress on Biosensors, Berlin, Germany, 3–5 June 1998.1. *Biosens. Bioelectron.* **1999**, *14* (2), 195–202. https://doi.org/10.1016/S0956-5663(98)00106-7.

(340) Trincavelli, M.; Coradeschi, S.; Loutfi, A. Classification of Odours with Mobile Robots Based on Transient Response. In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*; 2008; pp 4110–4115. https://doi.org/10.1109/IROS.2008.4650713.

(341) Trincavelli, M.; Coradeschi, S.; Loutfi, A. Odour Classification System for Continuous Monitoring Applications. *Sens. Actuators B Chem.* **2009**, *139* (2), 265–273. https://doi.org/10.1016/j.snb.2009.03.018.

(342) Baietto, M.; Wilson, A.; Baietto, M.; Wilson, A. D. Electronic-Nose Applications for Fruit Identification, Ripeness and Quality Grading. *Sensors* **2015**, *15* (1), 899–931. https://doi.org/10.3390/s150100899.

(343) Meij, T. G. de; Larbi, I. B.; Schee, M. P. van der; Lentferink, Y. E.; Paff, T.; Droste, J. S. T. sive; Mulder, C. J.; Bodegraven, A. A. van; Boer, N. K. de. Electronic Nose Can Discriminate Colorectal Carcinoma and Advanced Adenomas by Fecal Volatile Biomarker Analysis: Proof of Principle Study. *Int. J. Cancer* **2014**, *134* (5), 1132–1138. https://doi.org/10.1002/ijc.28446.

(344) Burgués, J.; Jiménez-Soto, J. M.; Marco, S. Estimation of the Limit of Detection in Semiconductor Gas Sensors through Linearized Calibration Models. *Anal. Chim. Acta* **2018**, *1013*, 13–25. https://doi.org/10.1016/j.aca.2018.01.062.

(345) Szulczyński, B.; Rybarczyk, P.; Gębicki, J. Monitoring of N-Butanol Vapors Biofiltration Process Using an Electronic Nose Combined with Calibration Models. *Monatshefte Chem.* **2018**, *149* (9), 1693–1699. https://doi.org/10.1007/s00706-018-2243-6.

(346) Walker, D. B.; Walker, J. C.; Cavnar, P. J.; Taylor, J. L.; Pickel, D. H.; Hall, S. B.; Suarez, J. C. Naturalistic Quantification of Canine Olfactory Sensitivity. *Appl. Anim. Behav. Sci.* **2006**, *97* (2), 241–254. https://doi.org/10.1016/j.applanim.2005.07.009.

(347) Zheng, J. B.; Yang, L.; Chen, J. B.; Wang, Y. M. Study on Odor Source Localization Method Based on Bionic Olfaction. *Appl. Mech. Mater.* **2013**, *448–453*, 391–395. https://doi.org/10.4028/www.scientific.net/AMM.448-453.391.

(348) Sánchez-Purrà, M.; Roig-Solvas, B.; Rodriguez-Quijada, C.; Leonardo, B. M.; Hamad-Schifferli, K. Reporter Selection for Nanotags in Multiplexed Surface Enhanced Raman Spectroscopy Assays. *ACS Omega* **2018**, *3* (9), 10733–10742. https://doi.org/10.1021/acsomega.8b01499.

(349) de Albuquerque, C. D. L.; Sobral-Filho, R. G.; Poppi, R. J.; Brolo, A. G. Digital Protocol for Chemical Analysis at Ultralow Concentrations by Surface-Enhanced Raman Scattering. *Anal. Chem.* **2018**, *90* (2), 1248–1254. https://doi.org/10.1021/acs.analchem.7b03968.

(350) Rahmani, M.; Lei, D. Y.; Giannini, V.; Lukiyanchuk, B.; Ranjbar, M.; Liew, T. Y. F.; Hong, M.; Maier, S. A. Subgroup Decomposition of Plasmonic Resonances in Hybrid Oligomers: Modeling the Resonance Lineshape. *Nano Lett.* **2012**, *12* (4), 2101–2106. https://doi.org/10.1021/nl3003683.

(351) Jones, S.; Sinha, S. S.; Pramanik, A.; Ray, P. C. Three-Dimensional (3D) Plasmonic Hot Spots for Label-Free Sensing and Effective Photothermal Killing of Multiple Drug Resistant Superbugs. *Nanoscale* **2016**, *8* (43), 18301–18308. https://doi.org/10.1039/c6nr05888d.

(352) Felten, J.; Hall, H.; Jaumot, J.; Tauler, R.; de Juan, A.; Gorzsás, A. Vibrational Spectroscopic Image Analysis of Biological Material Using Multivariate Curve Resolution–Alternating Least Squares (MCR-ALS). *Nat. Protoc.* **2015**, *10* (2), 217–240. https://doi.org/10.1038/nprot.2015.008.

(353) Hong, K. Y.; de Albuquerque, C. D. L.; Poppi, R. J.; Brolo, A. G. Determination of Aqueous Antibiotic Solutions Using SERS Nanogratings. *Anal. Chim. Acta* **2017**, *982*, 148–155. https://doi.org/10.1016/j.aca.2017.05.025.

(354) Ioffe, S.; Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *ArXiv150203167 Cs* **2015**.

(355) Lee, D. D.; Seung, H. S. Learning the Parts of Objects by Non-Negative Matrix Factorization. *Nature* **1999**, *401* (6755), 788–791. https://doi.org/10.1038/44565.

(356) Laing, D. G. Natural Sniffing Gives Optimum Odour Perception for Humans. *Perception* **1983**, *12* (2), 99–117. https://doi.org/10.1068/p120099.

(357) Delahunt, C. B.; Kutz, J. N. Putting a Bug in ML: The Moth Olfactory Network Learns to Read MNIST.

ArXiv180205405 Cs Q-Bio **2018**.

(358) Galvan, D. D.; Yu, Q. Surface-Enhanced Raman Scattering for Rapid Detection and Characterization of Antibiotic-Resistant Bacteria. *Adv. Healthc. Mater.* **2018**, *7* (13), 1701335. https://doi.org/10.1002/adhm.201701335.

(359) Liu, Y.; Zhou, H.; Hu, Z.; Yu, G.; Yang, D.; Zhao, J. Label and Label-Free Based Surface-Enhanced Raman Scattering for Pathogen Bacteria Detection: A Review. *Biosens. Bioelectron.* **2017**, *94*, 131–140. https://doi.org/10.1016/j.bios.2017.02.032.

(360) Bodelón, G.; Montes-García, V.; Costas, C.; Pérez-Juste, I.; Pérez-Juste, J.; Pastoriza-Santos, I.; Liz-Marzán, L. M. Imaging Bacterial Interspecies Chemical Interactions by Surface-Enhanced Raman Scattering. *ACS Nano* **2017**, *11* (5), 4631–4640. https://doi.org/10.1021/acsnano.7b00258.

(361) Huang, W. Y.; Zhao, Q. C.; Sharac, N.; Capolino, F.; Ragan, R.; Boyraz, O. Highly Nonlinear Sub-Micron Silicon Nitride Trench Waveguide Coated with Gold Nanoparticles. *Nonlinear Opt. Appl. Vol 9503 Pp 95030H SPIE - Int. Soc. Opt. Eng. USA SPIE - Int. Soc. Opt. Eng.* **2015**. https://doi.org/10.1117/12.2182290.

(362) Mosier-Boss, P.; Mosier-Boss, P. A. Review of SERS Substrates for Chemical Sensing. *Nanomaterials* **2017**, *7* (6), 142. https://doi.org/10.3390/nano7060142.

(363) Hakonen, A.; Andersson, P. O.; Stenbæk Schmidt, M.; Rindzevicius, T.; Käll, M. Explosive and Chemical Threat Detection by Surface-Enhanced Raman Scattering: A Review. *Anal. Chim. Acta* **2015**, *893*, 1–13. https://doi.org/10.1016/j.aca.2015.04.010.

(364) Si, K. J.; Guo, P.; Shi, Q.; Cheng, W. Self-Assembled Nanocube-Based Plasmene Nanosheets as Soft Surface-Enhanced Raman Scattering Substrates toward Direct Quantitative Drug Identification on Surfaces. *Anal. Chem.* **2015**, *87* (10), 5263–5269. https://doi.org/10.1021/acs.analchem.5b00328.

(365) Pang, S.; Yang, T.; He, L. Review of Surface Enhanced Raman Spectroscopic (SERS) Detection of Synthetic Chemical Pesticides. *TrAC Trends Anal. Chem.* **2016**, *85*, 73–82. https://doi.org/10.1016/j.trac.2016.06.017.

(366) Kasera, S.; Herrmann, L. O.; Barrio, J. del; Baumberg, J. J.; Scherman, O. A. Quantitative Multiplexing with Nano-Self-Assemblies in SERS. *Sci. Rep.* **2014**, *4*, 6785. https://doi.org/10.1038/srep06785.

(367) Song, B.; Yao, Y.; Groenewald, R. E.; Wang, Y.; Liu, H.; Wang, Y.; Li, Y.; Liu, F.; Cronin, S. B.; Schwartzberg, A. M.; et al. Probing Gap Plasmons Down to Subnanometer Scales Using Collapsible Nanofingers. *ACS Nano* **2017**, *11* (6), 5836–5843. https://doi.org/10.1021/acsnano.7b01468.

(368) Cha, H.; Yoon, J. H.; Yoon, S. Probing Quantum Plasmon Coupling Using Gold Nanoparticle Dimers with Tunable Interparticle Distances Down to the Subnanometer Range. *ACS Nano* **2014**, *8* (8), 8554–8563. https://doi.org/10.1021/nn5032438.

(369) Nguyen, C. Q. Machine Learning for SERS Quantitative Detection of Pyocyanin, UC Irvine, 2018.

(370) Alharbi, O.; Xu, Y.; Goodacre, R. Simultaneous Multiplexed Quantification of Nicotine and Its Metabolites Using Surface Enhanced Raman Scattering. *Analyst* **2014**, *139* (19), 4820–4827. https://doi.org/10.1039/C4AN00879K.

(371) Basheer, I. A.; Hajmeer, M. Artificial Neural Networks: Fundamentals, Computing, Design, and Application. *J. Microbiol. Methods* **2000**, *43* (1), 3–31. https://doi.org/10.1016/S0167-7012(00)00201-3.

(372) Hornik, K.; Stinchcombe, M.; White, H. Multilayer Feedforward Networks Are Universal Approximators. *Neural Netw.* **1989**, *2* (5), 359–366. https://doi.org/10.1016/0893-6080(89)90020-8.

(373) Hassoun, M. H. *Fundamentals of Artificial Neural Networks*; MIT Press, 1995.

(374) Le Ru, E. C.; Grand, J.; Sow, I.; Somerville, W. R. C.; Etchegoin, P. G.; Treguer-Delapierre, M.; Charron, G.; Félidj, N.; Lévi, G.; Aubard, J. A Scheme for Detecting Every Single Target Molecule with Surface-Enhanced Raman Spectroscopy. *Nano Lett.* **2011**, *11* (11), 5013–5019. https://doi.org/10.1021/nl2030344.

(375) Tsantekidis, A.; Passalis, N.; Tefas, A.; Kanniainen, J.; Gabbouj, M.; Iosifidis, A. Forecasting Stock Prices from the Limit Order Book Using Convolutional Neural Networks. In *2017 IEEE 19th Conference on Business Informatics (CBI)*; 2017; Vol. 01, pp 7–12. https://doi.org/10.1109/CBI.2017.23.

(376) Tabar, Y. R.; Halici, U. A Novel Deep Learning Approach for Classification of EEG Motor Imagery Signals. *J. Neural Eng.* **2017**, *14* (1), 016003. https://doi.org/10.1088/1741-2560/14/1/016003.

(377) Lee, J.; Park, J.; Kim, K. L.; Nam, J. Sample-Level Deep Convolutional Neural Networks for Music Auto-Tagging Using Raw Waveforms. *ArXiv170301789 Cs* **2017**.

(378) Selvaraju, R. R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *ArXiv161002391 Cs* **2016**.

(379) Moerner, W. E.; Kador, L. Optical Detection and Spectroscopy of Single Molecules in a Solid. *Phys. Rev. Lett.* **1989**, *62* (21), 2535–2538. https://doi.org/10.1103/PhysRevLett.62.2535.

(380) Betzig, E.; Chichester, R. J. Single Molecules Observed by Near-Field Scanning Optical Microscopy.

Science **1993**, *262* (5138), 1422–1425. https://doi.org/10.1126/science.262.5138.1422.

(381) Macklin, J. J.; Trautman, J. K.; Harris, T. D.; Brus, L. E. Imaging and Time-Resolved Spectroscopy of Single Molecules at an Interface. *Science* **1996**, *272* (5259), 255–258. https://doi.org/10.1126/science.272.5259.255.

(382) Schmidt, T.; Schütz, G. J.; Baumgartner, W.; Gruber, H. J.; Schindler, H. Imaging of Single Molecule Diffusion. *Proc. Natl. Acad. Sci.* **1996**, *93* (7), 2926–2929. https://doi.org/10.1073/pnas.93.7.2926.

(383) Dieringer, J. A.; Wustholz, K. L.; Masiello, D. J.; Camden, J. P.; Kleinman, S. L.; Schatz, G. C.; Van Duyne, R. P. Surface-Enhanced Raman Excitation Spectroscopy of a Single Rhodamine 6G Molecule. *J. Am. Chem. Soc.* **2009**, *131* (2), 849–854. https://doi.org/10.1021/ja8080154.

(384) Stranahan, S. M.; Willets, K. A. Super-Resolution Optical Imaging of Single-Molecule SERS Hot Spots. *Nano Lett.* **2010**, *10* (9), 3777–3784. https://doi.org/10.1021/nl102559d.

(385) Nguyen, C. Q.; Thrift, W. J.; Bhattacharjee, A.; Ranjbar, S.; Gallagher, T.; Darvishzadeh-Varcheie, M.; Sanderson, R. N.; Capolino, F.; Whiteson, K.; Baldi, P.; et al. Longitudinal Monitoring of Biofilm Formation via Robust Surface-Enhanced Raman Scattering Quantification of Pseudomonas Aeruginosa-Produced Metabolites. *ACS Appl. Mater. Interfaces* **2018**, *10* (15), 12364–12373. https://doi.org/10.1021/acsami.7b18592.

(386) Bouatra, S.; Aziat, F.; Mandal, R.; Guo, A. C.; Wilson, M. R.; Knox, C.; Bjorndahl, T. C.; Krishnamurthy, R.; Saleem, F.; Liu, P.; et al. The Human Urine Metabolome. *PLoS ONE* **2013**, *8* (9). https://doi.org/10.1371/journal.pone.0073076.

(387) Yampolsky, S.; Fishman, D. A.; Dey, S.; Hulkko, E.; Banik, M.; Potma, E. O.; Apkarian, V. A. Seeing a Single Molecule Vibrate through Time-Resolved Coherent Anti-Stokes Raman Scattering. *Nat. Photonics* **2014**, *8* (8), 650–656. https://doi.org/10.1038/nphoton.2014.143.

(388) Lombardi, J. R.; Birke, R. L.; Haran, G. Single Molecule SERS Spectral Blinking and Vibronic Coupling. *J. Phys. Chem. C* **2011**, *115* (11), 4540–4545. https://doi.org/10.1021/jp111345u.

(389) Etchegoin, P. G.; Le Ru, E. C. Resolving Single Molecules in Surface-Enhanced Raman Scattering within the Inhomogeneous Broadening of Raman Peaks. *Anal. Chem.* **2010**, *82* (7), 2888–2892. https://doi.org/10.1021/ac9028888.

(390) G. Etchegoin, P.; Ru, E. C. L.; Fainstein, A. Bi-Analyte Single Molecule SERS Technique with Simultaneous Spatial Resolution. *Phys. Chem. Chem. Phys.* **2011**, *13* (10), 4500–4506. https://doi.org/10.1039/C0CP02335C.

(391) Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. *ArXiv14094842 Cs* **2014**.

(392) S. Indrasekara, A. S. D.; Meyers, S.; Shubeita, S.; C. Feldman, L.; Gustafsson, T.; Fabris, L. Gold Nanostar Substrates for SERS-Based Chemical Sensing in the Femtomolar Regime. *Nanoscale* **2014**, *6* (15), 8891–8899. https://doi.org/10.1039/C4NR02513J.

(393) van Dyk, D. A.; Meng, X.-L. The Art of Data Augmentation. *J. Comput. Graph. Stat.* **2001**, *10* (1), 1–50. https://doi.org/10.1198/10618600152418584.

(394) Rochford, C.; Sridhar, D.; Woods, N.; Saleh, Z.; Hartenstein, L.; Ahlawat, H.; Whiting, E.; Dybul, M.; Cars, O.; Goosby, E.; et al. Global Governance of Antimicrobial Resistance. *The Lancet* **2018**, *391* (10134), 1976–1978. https://doi.org/10.1016/S0140-6736(18)31117-6.

(395) Fleming-Dutra, K. E.; Hersh, A. L.; Shapiro, D. J.; Bartoces, M.; Enns, E. A.; File, T. M.; Finkelstein, J. A.; Gerber, J. S.; Hyun, D. Y.; Linder, J. A.; et al. Prevalence of Inappropriate Antibiotic Prescriptions Among US Ambulatory Care Visits, 2010-2011. *JAMA* **2016**, *315* (17), 1864–1873. https://doi.org/10.1001/jama.2016.4151.

(396) Leonard, H.; Colodner, R.; Halachmi, S.; Segal, E. Recent Advances in the Race to Design a Rapid Diagnostic Test for Antimicrobial Resistance. *ACS Sens.* **2018**, *3* (11), 2202–2217. https://doi.org/10.1021/acssensors.8b00900.

(397) Sugden, R.; Kelly, R.; Davies, S. Combatting Antimicrobial Resistance Globally. *Nat. Microbiol.* **2016**, *1*, 16187. https://doi.org/10.1038/nmicrobiol.2016.187.

(398) Jim O'Neill. *Tackling Drug-Resistant Infections Globally: Final Report and Recommendations*; AMR Review.

(399) Reller, L. B.; Weinstein, M.; Jorgensen, J. H.; Ferraro, M. J. Antimicrobial Susceptibility Testing: A Review of General Principles and Contemporary Practices. *Clin. Infect. Dis.* **2009**, *49* (11), 1749–1755. https://doi.org/10.1086/647952.

(400) Nault, V.; Pepin, J.; Beaudoin, M.; Perron, J.; Moutquin, J.-M.; Valiquette, L. Sustained Impact of a Computer-Assisted Antimicrobial Stewardship Intervention on Antimicrobial Use and Length of Stay. *J.*

Antimicrob. Chemother. **2017**, *72* (3), 933–940. https://doi.org/10.1093/jac/dkw468.

(401) de With, K.; Allerberger, F.; Amann, S.; Apfalter, P.; Brodt, H.-R.; Eckmanns, T.; Fellhauer, M.; Geiss, H. K.; Janata, O.; Krause, R.; et al. Strategies to Enhance Rational Use of Antibiotics in Hospital: A Guideline by the German Society for Infectious Diseases. *Infection* **2016**, *44* (3), 395–439. https://doi.org/10.1007/s15010-016-0885-z.

(402) Spellberg, B. The Future of Antibiotics. *Crit. Care* **2014**, *18* (3), 228. https://doi.org/10.1186/cc13948.

(403) O'Dwyer, K.; Spivak, A. T.; Ingraham, K.; Min, S.; Holmes, D. J.; Jakielaszek, C.; Rittenhouse, S.; Kwan, A. L.; Livi, G. P.; Sathe, G.; et al. Bacterial Resistance to Leucyl-TRNA Synthetase Inhibitor GSK2251052 Develops during Treatment of Complicated Urinary Tract Infections. *Antimicrob. Agents Chemother.* **2015**, *59* (1), 289–298. https://doi.org/10.1128/AAC.03774-14.

(404) Nijhuis, R. H. T.; Guerendiain, D.; Claas, E. C. J.; Templeton, K. E. Comparison of EPlex Respiratory Pathogen Panel with Laboratory-Developed Real-Time PCR Assays for Detection of Respiratory Pathogens. *J. Clin. Microbiol.* **2017**, *55* (6), 1938–1945. https://doi.org/10.1128/JCM.00221-17.

(405) Caliendo, A. M.; Hanson, K. E. Point-Counterpoint: The FDA Has a Role in Regulation of Laboratory-Developed Tests. *J. Clin. Microbiol.* **2016**, *54* (4), 829–833. https://doi.org/10.1128/JCM.00063-16.

(406) Steingart, K. R.; Sohn, H.; Schiller, I.; Kloda, L. A.; Boehme, C. C.; Pai, M.; Dendukuri, N. Xpert® MTB/RIF Assay for Pulmonary Tuberculosis and Rifampicin Resistance in Adults. *Cochrane Database Syst. Rev.* **2013**, No. 1. https://doi.org/10.1002/14651858.CD009593.pub2.

(407) Liu, C.-Y.; Han, Y.-Y.; Shih, P.-H.; Lian, W.-N.; Wang, H.-H.; Lin, C.-H.; Hsueh, P.-R.; Wang, J.-K.; Wang, Y.-L. Rapid Bacterial Antibiotic Susceptibility Test Based on Simple Surface-Enhanced Raman Spectroscopic Biomarkers. *Sci. Rep.* **2016**, *6*, 23375. https://doi.org/10.1038/srep23375.

(408) Pulido, M. R.; García-Quintanilla, M.; Martín-Peña, R.; Cisneros, J. M.; McConnell, M. J. Progress on the Development of Rapid Methods for Antimicrobial Susceptibility Testing. *J. Antimicrob. Chemother.* **2013**, *68* (12), 2710–2717. https://doi.org/10.1093/jac/dkt253.

(409) Holcomb, Z. E.; Tsalik, E. L.; Woods, C. W.; McClain, M. T. Host-Based Peripheral Blood Gene Expression Analysis for Diagnosis of Infectious Diseases. *J. Clin. Microbiol.* **2017**, *55* (2), 360–368. https://doi.org/10.1128/JCM.01057-16.

(410) de Albuquerque, C. D. L.; Sobral-Filho, R. G.; Poppi, R. J.; Brolo, A. G. Digital Protocol for Chemical Analysis at Ultralow Concentrations by Surface-Enhanced Raman Scattering. *Anal. Chem.* **2018**, *90* (2), 1248–1254. https://doi.org/10.1021/acs.analchem.7b03968.

(411) Wang, Y.; Lee, K.; Irudayaraj, J. Silver Nanosphere SERS Probes for Sensitive Identification of Pathogens. *J. Phys. Chem. C* **2010**, *114* (39), 16122–16128. https://doi.org/10.1021/jp1015406.

(412) Yang, L.; Yan, B.; Premasiri, W. R.; Ziegler, L. D.; Negro, L. D.; Reinhard, B. M. Engineering Nanoparticle Cluster Arrays for Bacterial Biosensing: The Role of the Building Block in Multiscale SERS Substrates. *Adv. Funct. Mater.* **2010**, *20* (16), 2619–2628. https://doi.org/10.1002/adfm.201000630.

(413) Wen-Yu Chiu, S.; Cheng, H.-W.; Chen, Z.-X.; Wang, H.-H.; Lai, M.-Y.; Wang, J.-K.; Wang, Y.-L. Quantification of Biomolecules Responsible for Biomarkers in the Surface-Enhanced Raman Spectra of Bacteria Using Liquid Chromatography-Mass Spectrometry. *Phys. Chem. Chem. Phys.* **2018**, *20* (12), 8032–8041. https://doi.org/10.1039/C7CP07103E.

(414) Bodelón, G.; Montes-García, V.; López-Puente, V.; Hill, E. H.; Hamon, C.; Sanz-Ortiz, M. N.; Rodal-Cedeira, S.; Costas, C.; Celiksoy, S.; Pérez-Juste, I.; et al. Detection and Imaging of Quorum Sensing in Pseudomonas Aeruginosa Biofilm Communities by Surface-Enhanced Resonance Raman Scattering. *Nat. Mater.* **2016**, *15* (11), 1203–1211. https://doi.org/10.1038/nmat4720.

(415) Esteva, A.; Kuprel, B.; Novoa, R. A.; Ko, J.; Swetter, S. M.; Blau, H. M.; Thrun, S. Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks. *Nature* **2017**, *542* (7639), 115–118. https://doi.org/10.1038/nature21056.

(416) Cheng, J.-Z.; Ni, D.; Chou, Y.-H.; Qin, J.; Tiu, C.-M.; Chang, Y.-C.; Huang, C.-S.; Shen, D.; Chen, C.-M. Computer-Aided Diagnosis with Deep Learning Architecture: Applications to Breast Lesions in US Images and Pulmonary Nodules in CT Scans. *Sci. Rep.* **2016**, *6*, 24454. https://doi.org/10.1038/srep24454.

(417) Kassahun, Y.; Yu, B.; Tibebu, A. T.; Stoyanov, D.; Giannarou, S.; Metzen, J. H.; Vander Poorten, E. Surgical Robotics beyond Enhanced Dexterity Instrumentation: A Survey of Machine Learning Techniques and Their Role in Intelligent and Autonomous Surgical Actions. *Int. J. Comput. Assist. Radiol. Surg.* **2016**, *11* (4), 553–568. https://doi.org/10.1007/s11548-015-1305-z.

(418) Mayer, H.; Gomez, F.; Wierstra, D.; Nagy, I.; Knoll, A.; Schmidhuber, J. A System for Robotic Heart Surgery That Learns to Tie Knots Using Recurrent Neural Networks. *Adv. Robot.* **2008**, *22* (13–14), 1521–1537. https://doi.org/10.1163/156855308X360604.

(419) Miotto, R.; Li, L.; Kidd, B. A.; Dudley, J. T. Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records. *Sci. Rep.* **2016**, *6*, 26094. https://doi.org/10.1038/srep26094.

(420) Razavian, N.; Marcus, J.; Sontag, D. Multi-Task Prediction of Disease Onsets from Longitudinal Laboratory Tests. In *Machine Learning for Healthcare Conference*; 2016; pp 73–100.

(421) Thodoroff, P.; Pineau, J.; Lim, A. Learning Robust Features Using Deep Learning for Automatic Seizure Detection. In *Machine Learning for Healthcare Conference*; 2016; pp 178–190.

(422) Oh, S. L.; Hagiwara, Y.; Raghavendra, U.; Yuvaraj, R.; Arunkumar, N.; Murugappan, M.; Acharya, U. R. A Deep Learning Approach for Parkinson’s Disease Diagnosis from EEG Signals. *Neural Comput. Appl.* **2018**. https://doi.org/10.1007/s00521-018-3689-5.

(423) Le Ru, E. C.; Meyer, M.; Etchegoin, P. G. Proof of Single-Molecule Sensitivity in Surface Enhanced Raman Scattering (SERS) by Means of a Two-Analyte Technique. *J. Phys. Chem. B* **2006**, *110* (4), 1944–1948. https://doi.org/10.1021/jp054732v.

(424) Shi, H.; Wang, H.; Meng, X.; Chen, R.; Zhang, Y.; Su, Y.; He, Y. Setting Up a Surface-Enhanced Raman Scattering Database for Artificial-Intelligence-Based Label-Free Discrimination of Tumor Suppressor Genes. *Anal. Chem.* **2018**, *90* (24), 14216–14221. https://doi.org/10.1021/acs.analchem.8b03080.

(425) Li, J.; Skeete, Z.; Shan, S.; Yan, S.; Kurzatkowska, K.; Zhao, W.; Ngo, Q. M.; Holubovska, P.; Luo, J.; Hepel, M.; et al. Surface Enhanced Raman Scattering Detection of Cancer Biomarkers with Bifunctional Nanocomposite Probes. *Anal. Chem.* **2015**, *87* (21), 10698–10702. https://doi.org/10.1021/acs.analchem.5b03456.

(426) Li, X.; Yang, T.; Li, S.; Jin, L.; Wang, D.; Guan, D.; Ding, J. Noninvasive Liver Diseases Detection Based on Serum Surface Enhanced Raman Spectroscopy and Statistical Analysis. *Opt. Express* **2015**, *23* (14), 18361–18372. https://doi.org/10.1364/OE.23.018361.

(427) Belkum, A. van; Bachmann, T. T.; Lüdke, G.; Lisby, J. G.; Kahlmeter, G.; Mohess, A.; Becker, K.; Hays, J. P.; Woodford, N.; Mitsakakis, K.; et al. Developmental Roadmap for Antimicrobial Susceptibility Testing Systems. *Nat. Rev. Microbiol.* **2019**, *17* (1), 51. https://doi.org/10.1038/s41579-018-0098-9.

(428) Belkum, A. van; Dunne, W. M. Next-Generation Antimicrobial Susceptibility Testing. *J. Clin. Microbiol.* **2013**, *51* (7), 2018–2024. https://doi.org/10.1128/JCM.00313-13.

(429) Ardila, D.; Kiraly, A. P.; Bharadwaj, S.; Choi, B.; Reicher, J. J.; Peng, L.; Tse, D.; Etemadi, M.; Ye, W.; Corrado, G.; et al. End-to-End Lung Cancer Screening with Three-Dimensional Deep Learning on Low-Dose Chest Computed Tomography. *Nat. Med.* **2019**, *25* (6), 954. https://doi.org/10.1038/s41591-019-0447-x.

(430) Burrello, A.; Schindler, K. A.; Benini, L.; Rahimi, A. Hyperdimensional Computing with Local Binary Patterns: One-Shot Learning for Seizure Onset Detection and Identification of Ictogenic Brain Regions from Short-Time IEEG Recordings. *IEEE Trans. Biomed. Eng.* **2019**, 1–1. https://doi.org/10.1109/TBME.2019.2919137.

(431) Kim, M.; Zuallaert, J.; De Neve, W. Few-Shot Learning Using a Small-Sized Dataset of High-Resolution FUNDUS Images for Glaucoma Diagnosis. In *Proceedings of the 2Nd International Workshop on Multimedia for Personal Health and Health Care*; MMHealth ’17; ACM: New York, NY, USA, 2017; pp 89–92. https://doi.org/10.1145/3132635.3132650.

(432) Rajkomar, A.; Oren, E.; Chen, K.; Dai, A. M.; Hajaj, N.; Hardt, M.; Liu, P. J.; Liu, X.; Marcus, J.; Sun, M.; et al. Scalable and Accurate Deep Learning with Electronic Health Records. *Npj Digit. Med.* **2018**, *1* (1), 18. https://doi.org/10.1038/s41746-018-0029-1.

(433) Ballinger, B.; Hsieh, J.; Singh, A.; Sohoni, N.; Wang, J.; Tison, G. H.; Marcus, G. M.; Sanchez, J. M.; Maguire, C.; Olgin, J. E.; et al. DeepHeart: Semi-Supervised Sequence Learning for Cardiovascular Risk Prediction. In *Thirty-Second AAAI Conference on Artificial Intelligence*; 2018.

(434) Kingma, D. P.; Welling, M. Auto-Encoding Variational Bayes. *ArXiv13126114 Cs Stat* **2013**.

(435) Gómez-Bombarelli, R.; Wei, J. N.; Duvenaud, D.; Hernández-Lobato, J. M.; Sánchez-Lengeling, B.; Sheberla, D.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Adams, R. P.; Aspuru-Guzik, A. Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules. *ACS Cent. Sci.* **2018**, *4* (2), 268–276. https://doi.org/10.1021/acscentsci.7b00572.

(436) Li, J.; Wang, C.; Kang, H.; Shao, L.; Hu, L.; Xiao, R.; Wang, S.; Gu, B. Label-Free Identification Carbapenem-Resistant Escherichia Coli Based on Surface-Enhanced Resonance Raman Scattering. *RSC Adv.* **2018**, *8* (9), 4761–4765. https://doi.org/10.1039/C7RA13063E.

(437) Tien, N.; Chen, H.-C.; Gau, S.-L.; Lin, T.-H.; Lin, H.-S.; You, B.-J.; Tsai, P.-C.; Chen, I.-R.; Tsai, M.-F.; Wang, I.-K.; et al. Diagnosis of Bacterial Pathogens in the Dialysate of Peritoneal Dialysis Patients with

Peritonitis Using Surface-Enhanced Raman Spectroscopy. *Clin. Chim. Acta* **2016**, *461*, 69–75. https://doi.org/10.1016/j.cca.2016.07.026.
(438) Maaten, L. van der; Hinton, G. Visualizing Data Using T-SNE. *J. Mach. Learn. Res.* **2008**, *9* (Nov), 2579–2605.
(439) Pan, S. J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22* (10), 1345–1359. https://doi.org/10.1109/TKDE.2009.191.
(440) Bos, L. D. J.; Sterk, P. J.; Schultz, M. J. Volatile Metabolites of Pathogens: A Systematic Review. *PLOS Pathog.* **2013**, *9* (5), e1003311. https://doi.org/10.1371/journal.ppat.1003311.
(441) Savitzky, Abraham.; Golay, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **1964**, *36* (8), 1627–1639. https://doi.org/10.1021/ac60214a047.
(442) Eilers, P.; Boelens, H. Baseline Correction with Asymmetric Least Squares Smoothing. *Leiden Univ. Med. Cent. Rep.* **2005**.
(443) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. *ArXiv14126980 Cs* **2014**.
(444) Liu, F. T.; Ting, K. M.; Zhou, Z. Isolation Forest. In *2008 Eighth IEEE International Conference on Data Mining*; 2008; pp 413–422. https://doi.org/10.1109/ICDM.2008.17.
(445) Roberts, S. J.; Husmeier, D.; Rezek, I.; Penny, W. Bayesian Approaches to Gaussian Mixture Modeling. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20* (11), 1133–1142. https://doi.org/10.1109/34.730550.