UNIVERSITY OF CALIFORNIA,
IRVINE

Machine Learning for SERS Quantitative Detection of Pyocyanin

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Materials Science and Engineering

by

Cuong Q. Nguyen

Thesis Committee:
Associate Professor Regina Ragan, Chair
Assistant Professor Allon Hochbaum
Associate Professor Ali Mohraz

2018

DEDICATION

To my father, the entrepreneur;

To my mother, the provider;

To my brothers, the two troublemakers;

And to Suzy, the best climbing partner I could ever hope for.

TABLE OF CONTENTS

Page

LIST OF FIGURES **v**

LIST OF TABLES **vii**

ACKNOWLEDGMENTS **viii**

ABSTRACT OF THE THESIS **ix**

1 Introduction **1**
1.1 Thesis Organization 3

2 Background **5**
2.1 Localized Surface Plasmon Resonance 5
2.2 Surface-Enhanced Raman Scattering Spectroscopy 7
2.2.1 Raman Spectroscopy 7
2.2.2 SERS Enhancements 9
2.3 Machine Learning for Regression Analysis 10
2.3.1 Partial Least Squares (PLS) Regression 11
2.3.2 Artificial Feedforward Neural Networks 12

3 Self-Assembled Plasmonic SERS Substrates **15**
3.1 Driving Chemical Crosslinking with EHD Flow 16
3.2 Uniform Optical Responses of SERS Substrates 17
3.3 Fabrication and Characterization Procedures 20
3.4 Summary 22

4 Robust SERS Quantitative Detection of Pyocyanin **24**
4.1 Pyocyanin: A Biomarker for PA 25
4.2 Experimental Procedures 25
4.3 Results and Discussions 29
4.3.1 Quantification and Detection of Pyocyanin in Aqueous Media 29
4.3.2 Predicting Pyocyanin Concentration with PLS Regression 31
4.3.3 Pyocyanin Quantitative Detection in Complex Media 33
4.3.4 Monitoring Biofilm Formation *via* Pyocyanin Quantification 35
4.4 Summary 38

5 Deep Learning for Improved SERS Quantitative Detection **40**
5.1 Feedforward Artificial Neural Networks 41
5.1.1 Hyperparameters Tuning 41
5.1.2 Limit of Quantification, R^2, and Prediction Error 41
5.2 Convolutional Neural Networks 44
5.2.1 Hyperparameters Tuning 44
5.2.2 Limit of Quantification, R^2, and Prediction Error 47
5.3 Summary 47

6 Conclusion **49**

Bibliography **50**

A Dose-Dependent Response of Pyocyanin Vibrational Bands **57**

B Desorption Half-Life of Pyocyanin **58**

C Visualizing CNNs Attention **61**

LIST OF FIGURES

Page

2.1 Schematic of physical system used to describe the phenomenon of localized surface plasmon resonance: A nanosphere with permittivity of $\varepsilon(\omega)$ embedded in a uniform dielectric medium with permittivity of $\varepsilon_m(\omega)$. Adapted by permission from Springer: *Plasmonics: fundamentals and applications* by Stefan A. Maier. Copyright 2007. 6
2.2 Energy diagram of Rayleigh and and Raman (Stokes and anti-Stokes) scattering. ν_I and ν represent the frequency of incident light and vibration, respectively. 8
2.3 Components of an artificial neuron. 12
2.4 Schematic representation of a feedforward ANN with 1 hidden layer. 13

3.1 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. 16
3.2 (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 ($\times$), 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%. 18

4.1 (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 calculated from 100 measurements. 31
4.2 (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 Δ*phz*1/2 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 calculated from 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, respectively. Error bars show standard deviation calculated from 400 measurements. (d) Growth curves for planktonic cultures of *P. aeruginosa* PA14 and its phenazine-null mutant strain as determined from optical density. 34
4.3 (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 calculated 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 calculated 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-attached bacteria to carbenicillin treatment at the specified growth times. Error bars depict standard deviation of dead cell fraction calculated from n = 3 biological replicates. 37

5.1 Predicted pyocyanin concentrations and losses during training of feedforward ANNs (a)(b) Model 1, (c)(d) Model 2, and (e)(f) Model 3. 43
5.2 Predicted pyocyanin concentrations and losses during training of 1D CNNs (a)(b) Model 1, (c)(d) Model 2, and (e)(f) Model 3. 46

LIST OF TABLES

Page

5.1 Summary of hyperparameters range and optimized values for feedforward ANNs. 42
5.2 Summary of hyperparameters range and optimized values for 1D CNNs. . . . 45

ACKNOWLEDGMENTS

It gives me great pleasure to acknowledge that this thesis would have not been possible without the contributions of many people.

I want to thank first and foremost my advisor, **Prof. Regina Ragan**, for her open-mindedness. Regina gave me the freedom to explore the vast sea of knowledge while providing continuous trust and support throughout my academic career. Regina had also given me the experience of attending multiple scientific conferences, where I was simultaneously humbled and inspired. Above all, Regina's ability to bring together teams of multidisciplinary researchers is nothing short of miraculous.

I would like to extend my acknowledgments to my committee members, **Prof. Allon Hochbaum** and **Prof. Ali Mohraz** for their patience and insightful comments. It was a privilege to have Allon and Ali as instructors during my undergraduate and graduate studies. Their passions for excellence in teaching is evident in their concise, clear, and effective lectures, which have unquestionably inspired many students year after year.

I would also like to express my deepest appreciation to **Dr. Dima Fishman** at the Laser Spectroscopy Facility for his assistance in troubleshooting my data collection process on the Raman microscope, **Prof. Katrine Whiteson** for her crucial scientific insights on metabolomics, and **Prof. Pierre Baldi** for introducing me to the mesmerizing world of machine learning and deep learning.

I want to also acknowledge my friends and collaborators – **Arunima Bhattacharjee**, **Mahsa Darvishzadeh-Varcheie**, **Tara Gallagher**, **Saba Ranjbar**, and **Qiancheng Zhao** – without whom I would still have many questions unanswered.

I thank **William Thrift** for his mentorship. My Master would have not been the same without the stimulating scientific discussions during our Eureka's *brainstorming* sessions.

I thank my parents, **Viet and Phuong**, for their unconditional support, and finally, **Suzanne Cardenas** for her unwavering belief in me.

Thank you.

Sections of this thesis consist materials reprinted with permissions from the American Chemical Society and Springer.

This work was supported by National Science Foundation EECS-1449397.

ABSTRACT OF THE THESIS

Machine Learning for SERS Quantitative Detection of Pyocyanin

By

Cuong Q. Nguyen

Master of Science in Materials Science and Engineering

University of California, Irvine, 2018

Associate Professor Regina Ragan, Chair

Since its discovery in 1977, surface-enhanced Raman scattering (SERS) spectroscopy has been cemented as a powerful spectroscopic technique. Taking advantage of local electric field enhancement from plasmonic nanostructures, SERS provides vibrational fingerprints down to the single molecule detection limits. Yet fully capitalizing on the technique has proven challenging. The problem is rooted in (1) inherent variances in SERS enhancement factors and (2) dated spectral analysis technique. To address (1), I present a fabrication scheme that produces optically uniform SERS substrates by employing electrohydrodynamic flow to drive chemical crosslinking between colloidal gold nanospheres. The resulting substrates exhibit SERS signals with relative standard deviation of 10.4 % over 100×100 µm^2. With pyocyanin as analyte – a secondary metabolite produced by P. aeruginosa – SERS substrates exhibit limit of quantification of 1 ng·mL^{-1} and robust quantification of concentrations spanning 5-orders of magnitude. To address (2), I implemented three machine learning algorithms to analyze SERS spectra. Partial least squares regression is crucial in monitoring of *P. aeruginosa* biofilm formation in a microfluidic environment, enabling detection as early as 3 h after inoculation in complex media. Feedforward artificial neural networks trained on pyocyanin data produces prediction errors of 6.2 ± 1.1 %. Finally, 1D convolutional neural networks trained with spectra stack further reduces prediction errors to an impressive 4.9 ± 0.9 %. Overall, this thesis demonstrates SERS spectroscopy as a potential diagnostic tool

while laying the foundation to fully exploit its sensing capability by integrating machine learning in the analysis pipeline.

Chapter 1

Introduction

Surface-enhanced Raman scattering (SERS) spectroscopy is an ultrasensitive vibrational spectroscopic technique that can provide molecular fingerprints as a label-free biosensing method. SERS-active surfaces, composed of plasmonic nanoantennas, provide enhancements to the local electric field and thereby optical signals that can reach detection limits down to the single molecule region [1]. Its ability to detect molecules at low concentrations makes SERS an attractive tool for biosensing applications where biologically relevant concentrations often range from nM to mM levels in clinical samples [2, 3].

Throughout the years, SERS *detection* capability has been repeatedly documented in the literature. SERS-based studies have demonstrated detection of small proteins in the nM range [4], DNA/RNA in the fM range using labeled sensing [5], and even differentiation between bacteria strains during their maturation phase [6, 7, 8]. Yet achieving reproducible *quantification* – a crucial aspect of biosensing – in these critical concentration ranges has proven difficult with SERS due to dated techniques used in interpreting SERS data and inherent variances of SERS enhancement factors resulting from non-uniformity in nanostructure architecture.

SERS enhancement factors are highly dependent on the nanogap 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 [9, 10]. At nanogap distances below approximately 0.5 nm, depolarization effects attributed to quantum tunneling reduce enhancements [11]. 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 [12, 13]. 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 [13, 14]. Inherent variances in SERS substrates' enhancement factors only worsen any analyte's location-dependent signal variations and will lead to large SERS intensity fluctuations. Thus, a fabrication scheme that produces SERS surfaces having uniform and dense nanogaps is necessary to increase the 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.

Another crucial component to this issue is the dated techniques used in interpreting SERS data. A SERS spectra is composed of rich vibrational information of the sample – depending on the molecule, it is not uncommon for a spectra to contain of 10 or more peaks. Yet 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 [15, 16, 17, 18, 19, 20, 21, 22, 23]. While efforts in breaking the norms is minimal, multivariate regression analyses have demonstrated some successes for SERS quantification of glucose [24]. Being linear algorithms, they cannot account for the

non-linear nature of SERS enhancement factor variances and non-uniform distribution of analyte molecule in SERS hotspots, and thus still lack performances at low concentrations. It is important then to utilize a nonlinear algorithms to fully capitalize on SERS sensing capability. In this thesis, I present a chemical assembly fabrication method to generate SERS surfaces that overcome such variances in electric field enhancements and a SERS-based microfluidic sensor to monitor biofilm formation of *Pseudomonas aeruginosa*, an opportunistic pathogen commonly associated with medical devices contaminations and immunocompromised patients [25, 26]. This is accomplished *via* SERS quantitative detection of pyocyanin – a secondary metabolite produced by the bacteria that concurrently acts as its biomarker – using statistical learning algorithms for spectral analysis. I also demonstrate the advantages of using deep learning tools in interpreting SERS data for improved quantification of pyocyanin.

1.1 Thesis Organization

Chapter 2 provides a theoretical and fundamental framework of the multidisciplinary topics discussed in this thesis, ranging from Raman scattering to neural networks. In **Chapter 3**, I introduce chemical assembly of SERS surfaces using EHD flow. Characterization and performance of resulting SERS surfaces is also discussed in this chapter to provide further insights on the advantages of using EHD flow to generate SERS surfaces. **Chapter 4** demonstrates instantaneous quantitative detection of pyocyanin using SERS surfaces, enabling early detection of *P. aeruginosa* biofilm formation. Spectra are collected in a microfluidic environment and are analyzed with partial least squares regression, a multivariate linear statistical learning algorithm, to quantify pyocyanin concentrations. Finally in **Chapter 5**, I investigate using deep learning methods to improve quantitative detection of pyocyanin. Performances

of feedforward artificial neural networks and convolutional neural networks are compared with linear algorithms for this task.

Chapter 2

Background

2.1 Localized Surface Plasmon Resonance

Metals, from an electromagnetic point of view, are plasmas of free electron gas oscillating around a fixed core. On a macro scale, these oscillations, termed bulk plasmons, exhibit resonant frequencies dictated in part by the restoring forces exerted on displaced electrons. When light incidents on a metal nanosphere surrounded by a bulk dielectric environment, as depicted in Figure 2.1, a similar phenomenon occurs. The oscillating electric field induces a dipole moment in the nanosphere and displaces the electron gas as permitted by its polarizability. The nanosphere's curved surface exerts a restoring force on the electron gas, as in the case of bulk plasmons, leading to the existence of a resonant frequency which is subsequently termed the localized surface plasmon resonance (LSPR). It is then possible to theoretically approximate this frequency by considering the polarizability of the nanosphere. When the nanosphere is small with respect to the wavelength of the incident light, the polarizability can be described in Equation 2.1 using simple quasi-static approximation where the phase of the time-dependent oscillating incident field is assumed to be constant over the particle

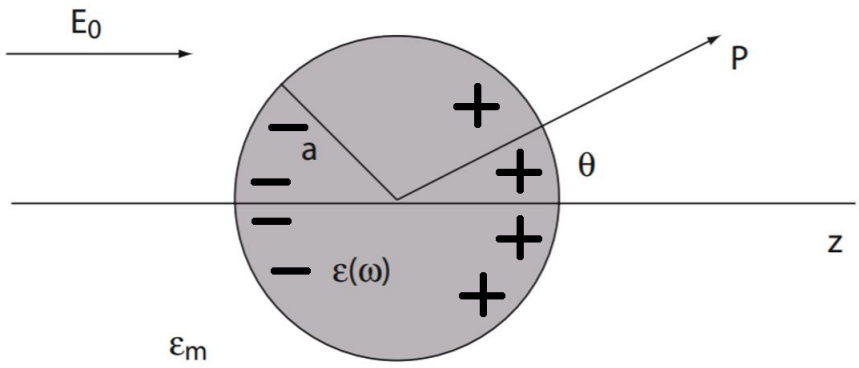

Figure 2.1: Schematic of physical system used to describe the phenomenon of localized surface plasmon resonance: A nanosphere with permittivity of $\varepsilon(\omega)$ embedded in a uniform dielectric medium with permittivity of $\varepsilon_m(\omega)$. Adapted by permission from Springer: *Plasmonics: fundamentals and applications* by Stefan A. Maier. Copyright 2007.

volume, thus simplifying the problem to a particle in an electrostatic field.

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

Here α is the polarizability of the nanosphere, a is the radius of the nanosphere, ε_d is the permittivity of the dielectric environment, ε_m is the permittivity of the metal, and ω is the frequency. The resonance condition is reached at the frequency where α is maximized, which Equation 2.1 shows to take place when $\varepsilon_m(\omega) = -2\varepsilon_d(\omega)$. Interestingly, this condition also implies that the plasmonic response of a metal nanosphere is dependent on the surrounding environment: A gold nanosphere in water will have a resonance that is red-shifted from that in air. It is also worth noting two important aspects of this analysis. First, the quasi-static approximation places an upper bound on the size of a nanosphere that can exhibit this resonance, meaning that when the nanosphere becomes large relative to the incident wavelength, surface plasmon polariton behavior is observed instead of LSPR. Second, when considering a real metal with complex permittivity, it is apparent that α is still confined from being infinite by the imaginary components of the permittivity associated with loss due to absorption.

Further investigation of the induced dipole described above reveals an important effect of the LSPR. Equation 2.2 describing the electric near field of an oscillating dipole will elucidate this.

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

Here $\boldsymbol{n}$ is the unit normal vector, $\boldsymbol{p}$ is the dipole moment (to be differentiated from point P in Figure 2.1), ε_0 is vacuum permittivity, and r is the distance from the dipole. Considering the dipole moment $\boldsymbol{p}$ is a function of the polarizability α, we can observe that, when the resonance condition is met, the local electric field near the nanosphere also experiences an enhancement as a result of the increased polarizability. This electric field enhancement is the foundation to various surface enhanced spectroscopy techniques, including SERS as further elaborated below.

2.2 Surface-Enhanced Raman Scattering Spectroscopy

2.2.1 Raman Spectroscopy

The Raman effect, commonly known as Raman scattering, is an inelastic scattering process between a photon and a phonon, a quantum of a molecular vibration. The scattered photons are reduced by energies characteristic to vibrational transitions within a molecule. Consequently a Raman spectra, constructed by probing the frequency change of the scattered photons with respect to the incident frequency, provides a vibrational fingerprint of a molecule with high specificity. Classical treatment of Raman scattering is sufficient to provide the necessary insight for this work. Consider an oscillating electric field applied to a molecule. The electrons and nuclei will be displaced by the field through interaction with phonon modes, thus resulting in an oscillating dipole moment in the molecule. This induced

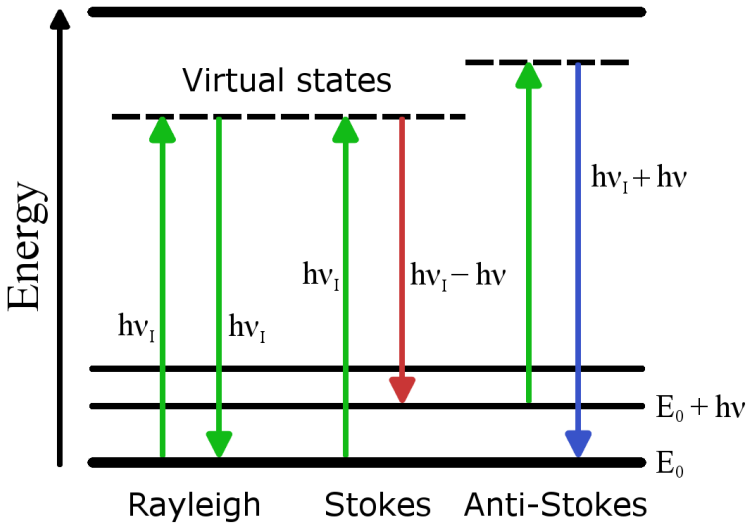

Figure 2.2: Energy diagram of Rayleigh and and Raman (Stokes and anti-Stokes) scattering. ν_I and ν represent the frequency of incident light and vibration, respectively.

dipole moment can be described simply by

$$\mu_{ind} = \alpha E \tag{2.3}$$

where μ_{ind} is the induced dipole moment, E is the incident electric field, and α is the polarizability, a property of the molecule. As a vibrating molecule oscillates in size and shape, its polarizability will also oscillate. Polarizability of the molecule can then be expanded as

$$\alpha = \alpha_0 + \alpha_1 \cos \omega t \tag{2.4}$$

where α_1 is the amplitude of the oscillation around the static polarizability α_0, and ω is the frequency of the molecular vibration. The oscillating incident field can be represented by

$$E = E_0 \cos \omega_I t \tag{2.5}$$

where ω_I is the frequency of the incident field. Substituting Equations 2.4 and 2.5 into 2.3 (and applying a simple trigonometric identity) gives Equation 2.6 [27].

$$\mu_{ind} = \alpha_0 E_0 \cos \omega_I t + \frac{1}{2} \alpha_1 E_0 \cos (\omega_I - \omega)t + \frac{1}{2} \alpha_1 E_0 \cos (\omega_I + \omega)t \tag{2.6}$$

From Equation 2.6, it is apparent that the induced dipoles—and thus vibrations—exist at three frequencies: ω_I, $\omega_I - \omega$, and $\omega_I + \omega$. The first term correlates to the unshifted frequency of Rayleigh scattering, while the second (red-shifted) and third (blue-shifted) terms are Stokes and anti-Stokes scattering, respectively. The energy diagram of the three scattering events is depicted in Figure 2.2. Since most molecules energetically remain at the ground state, Raman scattering is usually measured by the stronger Stokes scattering at $\omega_I - \omega$.

2.2.2 SERS Enhancements

Raman scattering, however, is an inefficient process. Approximately only 1 in 10^6 – 10^8 incident photons scatters inelastically, and thus Raman detection limit is not suitable for biosensing applications [28]. To overcome this, SERS utilizes LSPR of metal nanostructures to achieve detection of single molecules. From Equation 2.6, it is easy to see the benefit of LSPR in SERS: We can replace the incident field E_0 with the local field of the metal nanosphere from Equation 2.2, which is maximized at the resonant frequency. It is also worth noting that Raman spectroscopy instruments measure intensity, which is proportional to $|E|^2$. This intensity at the detector with respect to the incident light also experiences two enhancements: the electric field from external excitation source exciting the molecular vibrations *and* from the light inelastically scattered from the molecule are both increased by Equation 2.3. As a result, SERS exhibits approximately a $|E|^4$ enhancement relative to the incident electric field. This $|E|^4$ enhancement is at the core of SERS impressive

detection capability, consider that a 2-fold increase in electric field increases measured Raman scattering intensity by 16-fold.

In practice, bringing two or more single plasmonic nanospheres closely together can lead to coupling between their localized plasmons [29], which generates greater field enhancement at the nanogaps—up to 1000-fold, equivalent to a 10^{12}-fold increase in scattering intensity—when the incident electric field is along the inter-particle axis [30]. It has been also shown theoretically and experimentally that this enhancement diminishes quickly as the nanogap size increases [31, 32]. Thus it is necessary to have SERS substrates with uniform nanogaps to minimize the variance in field enhancement for reliable analyte quantification with SERS, as demonstrated in **Chapters 3** and **4**.

2.3 Machine Learning for Regression Analysis

Regression analysis is a group of statistical tools used to model relationships between variables. Widely used for prediction and forecasting, it involves understanding mathematically how changes in one or more independent variables affect the dependent variables of interest. In the context of SERS quantification, they provide a means to accurately predict analyte concentrations despite the non-linear nature of SERS enhancement factor variances and the non-uniform distribution of analyte on SERS surfaces. The two tools described below – partial least square (PLS) regression and artificial neural networks (ANNs) – subsequently used in **Chapters 4** and **5** to quantify pyocyanin concentrations from SERS spectra.

2.3.1 Partial Least Squares (PLS) Regression

PLS regression was developed in the 1980s to overcome the shortcomings of popular techniques at the time. PLS regression is now a standard chemometric tool utilized by researchers in a wide range of fields from analytical chemistry to social sciences.

In simple terms, PLS regression performs multiple linear regression on the projections of the inputs and outputs. Suppose a system has m features, p labels, and n training examples, we can set up a $n \times m$ matrix $\mathbf{X}$ consisting of all features and a corresponding $n \times p$ matrix $\mathbf{Y}$ consisting of all labels. In PLS, $\mathbf{X}$ and $\mathbf{Y}$ are decomposed according to

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E} \tag{2.7}$$

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^T + \mathbf{F} \tag{2.8}$$

where $\mathbf{T}$ and $\mathbf{U}$ are the $\mathbf{X}$ and $\mathbf{Y}$ scores, respectively, and $\mathbf{P}$ and $\mathbf{Q}$ are the $\mathbf{X}$ and $\mathbf{Y}$ loading vectors, respectively. The decomposition is carried out so that the covariance of $\mathbf{T}$ and $\mathbf{U}$ are maximized, thus allowing the loading vectors to be rotated relative to each other [33]. Finally the least squares regression estimate of $\mathbf{B}$ is calculated for the relationship $\mathbf{U} = \mathbf{T}\mathbf{B}$. In practice, one needs to choose the appropriate amount of loading vectors to prevent under- and over-fitting; since this value is generally smaller than m, the decomposition in Equation 2.7 is usually understood as a dimensional reduction step [34].

Intuitively speaking, PLS regression finds a best-fit linear model from the most important components of the inputs and outputs. The inputs and outputs are reduced to the components that explain the most changes, or variances, in each other. This step automatically minimizes redundancy within the data, resulting in more accurate models than simple multiple linear regression (MLR) or principal component regression (PCR). As an example, one can look at a Raman spectra as having high redundancy with respect to analyte concentrations: As the concentration increases (decreases), intensities of all peaks associated with

the analyte will also increase (decrease). Performing MLR or PCR in this case would result in large prediction error whereas PLS regression overcomes this by using only the essential information of the spectra.

2.3.2 Artificial Feedforward Neural Networks

Deep learning – having produced disruptively high performance in various artificial intelligence applications – is rooted in the development of artificial neural networks (ANNs). To understand ANNs, we first consider its fundamental building blocks: The artificial neurons (Figure 2.3).

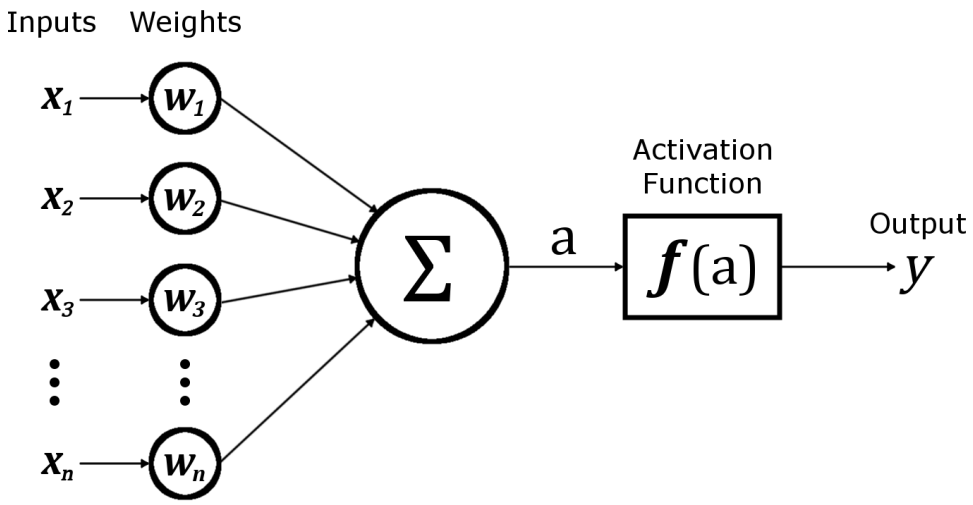

Figure 2.3: Components of an artificial neuron.

It is possible to see from Figure 2.3 that a single neuron takes in multiple inputs and outputs 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.

$$a = \sum x_j w_j \tag{2.9}$$

$$y = f(a) = f\Big(\sum x_j w_j\Big) \tag{2.10}$$

Here a is the activation, x_j is the j^{th} input, w_j is the corresponding j^{th} weight, $f(x)$ is the activation function, and y is the output. Activation functions can be chosen from a range of options, such as sigmoid, tanh, or the popular rectified linear units (ReLU) [35].

ANNs are built by connecting these neurons into larger networks, as visually depicted in Figure 2.4. Here each line represents a weight, and each circle represents a node that operates as a single neuron. The hidden layers, an important aspect in deep learning, are any layers of neurons between the input and output.

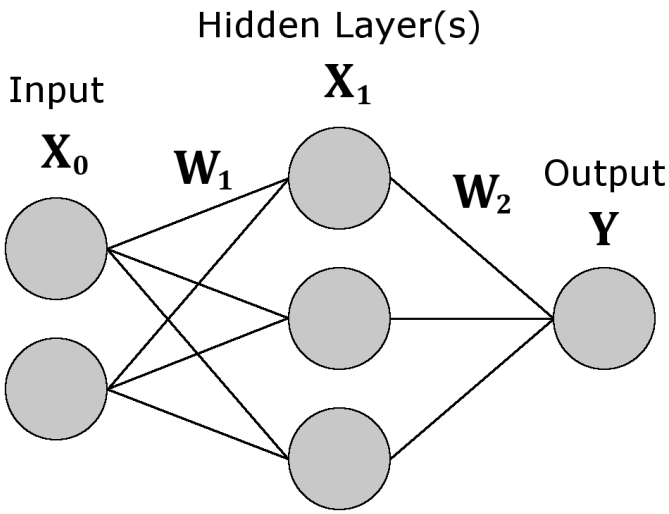

Figure 2.4: Schematic representation of a feedforward ANN with 1 hidden layer.

In matrix notation, one can describe the flow of values as followed

$$\mathbf{A_1} = \mathbf{X_0}\mathbf{W_1} + \mathbf{B_1} \tag{2.11}$$

$$\mathbf{X_1} = f(\mathbf{A_1}) \tag{2.12}$$

$$\mathbf{A_2} = \mathbf{X_1}\mathbf{W_2} + \mathbf{B_2} \tag{2.13}$$

$$\mathbf{X_2} = \mathbf{Y} = f(\mathbf{A_2}) \tag{2.14}$$

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. It is apparent, then, that ANNs are just 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 [36]. In practice, the learning rate, the number of hidden layers (depth) and nodes in each layer (width) are just three parameters from a multitude of application-dependent parameters that require precise tuning for optimal performance.

Chapter 3

Self-Assembled Plasmonic SERS Substrates

Inherent variance in SERS enhancement factors – as described in **Chapter 1** – is a crucial limiting factor when it comes to SERS reproducibility and quantification capability. This knowledge has understandably resulted in many reports demonstrating high uniformity *via* fine control of hotpots size and cluster morphology using top-down approaches [37, 38, 39, 40, 41]. In addition to expense, top down methods have substantial difficulty in achieving interparticle distances less than 5 nm, limiting achievable signal enhancements and detection limits. In this chapter I detail an alternative bottom-up self-assembly technique that minimizes variances in SERS enhancement by using EHD flow to drive chemical crosslinking. The resulting SERS surfaces exhibit optical uniformity over large area and are subsequently used in **Chapters 4** and **5**.

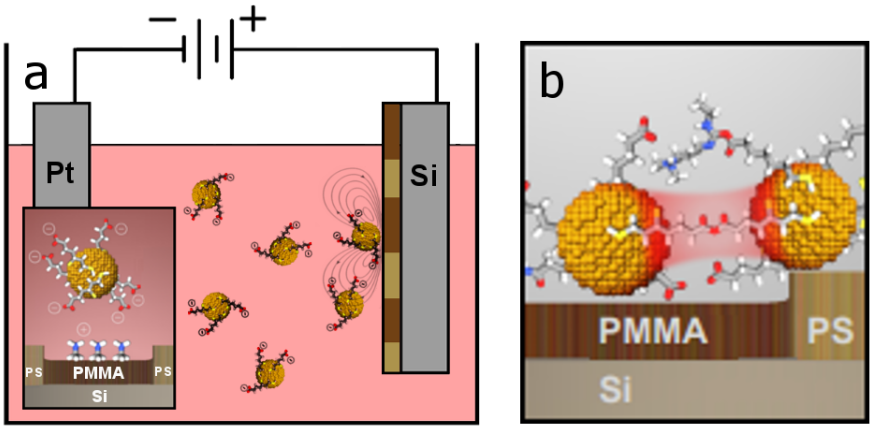

Figure 3.1: Schematic overview of chemical assembly of SERS substrates. (a) Lipoic acid-functionalized Au nanoparticles are electrophoretically driven toward a working electrode. Au seeds chemically crosslink with amine functionalized PMMA regions on self-organized diblock copolymer PS-*b*-PMMA template (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.

3.1 Driving Chemical Crosslinking with EHD Flow

Directing the assembly of Au oligomers with uniformly distributed electromagnetic hotspots over large areas is achieved by utilizing EHD flow to drive chemical crosslinking to form nanogaps with sub-nanometer gap spacing as reported previously [42]. The 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 3.1a, lipoic acid-functionalized Au nanoparticle monomer seeds are chemically assembled on amine-functionalized poly(methyl methacrylate) (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/s-NHS) 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 3.1a, inducing an

osmotic flow toward the seeds' equator, termed the EHD flow [43, 44, 45]. These flow fields entrain the 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 3.1b. 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 and observed in transmission electron microscopy [42].

3.2 Uniform Optical Responses of SERS Substrates

A scanning electron microscopy (SEM) image of a SERS substrate composed of self-assembled oligomers is displayed in Figure 3.2a. The size distribution in percent area (Figure 3.2b inset), determined from image analysis of SEM images using Wolfram Mathematica, 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 the UV-visible measurement as shown in Figure 3.2b. This agreement suggests that the measured attenuation is dominated by frequencies near the trimer resonance with contributions from frequencies associated with dimers and quadrumers that are blue- and red-shifted, respectively. Larger oligomers observed in the SEM image of Figure 3.2a 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.

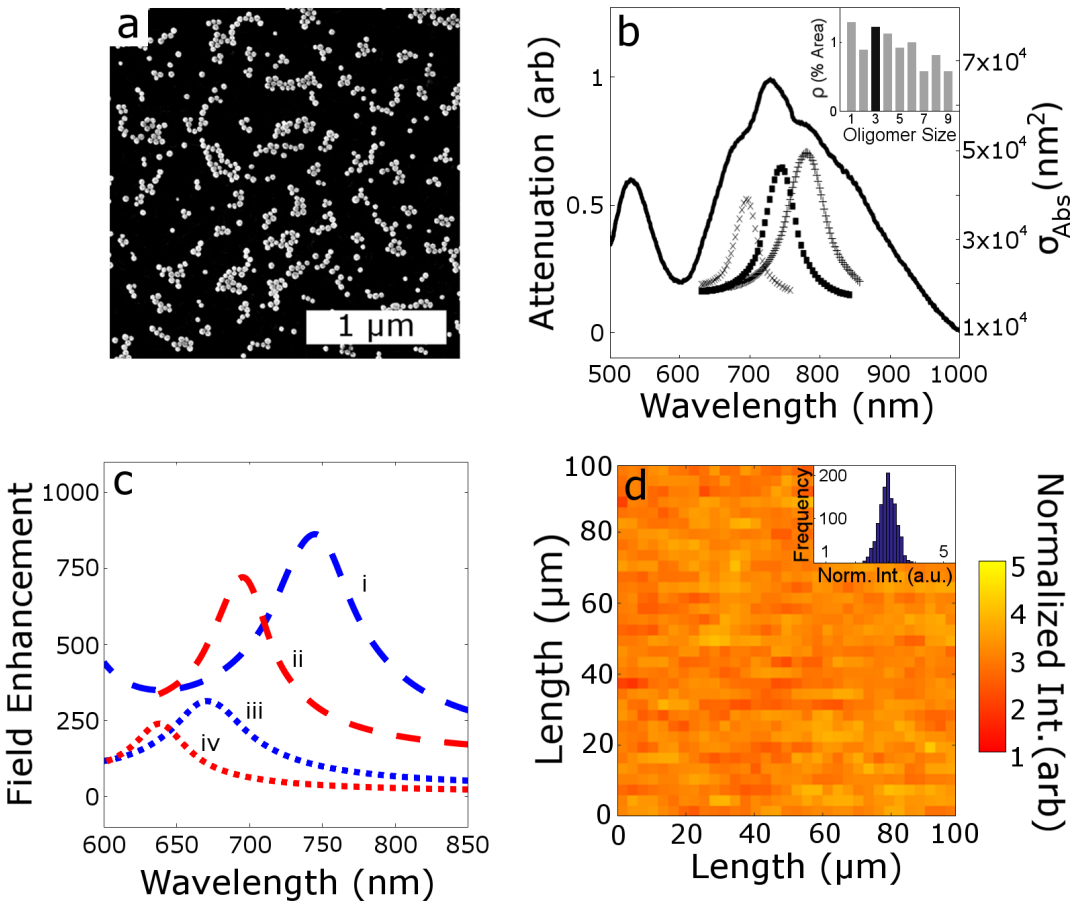

Figure 3.2: (a) SEM micrograph of self-assembled Au oligomers. (b) Attenuation curve of SERS substrates in DI water. Dotted curves represent absorption cross section of a dimer (×), linear trimer (squares), and linear quadrumer (+) from full-wave simulations. Inset shows oligomers distributions in percent 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%.

For example, a hexamer made up of two trimers on top of each other has a plasmon mode that is slightly red-shifted from that of a linear trimer [46]. 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 3.2b inset, it was determined that, by number, 98% (88% by area) of the 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—$|E_{olig}|/|E_0|^4$, where E_{olig} and E_0 are the plane wave field with and without oligomers, respectively—was performed for nanogaps of 0.9 and 2 nm in dimers and linear trimers and are shown in Figure 3.2c. As expected, the resonance red-shifts as the gap distance decreases. Furthermore, increasing the 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 the spacing decreases below 2 nm. Thus, chemically controlled gap spacing enables uniform SERS intensity, as observed in Figure 3.2d.

Figure 3.2d 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 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 625–875 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 data sets needed for statistical analysis enabling quantitative detection.

3.3 Fabrication and Characterization Procedures

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 spin-coated on the flat Si surface and annealed at 198 ^{o}C to form thin films as described in previous work [47]. PMMA regions were 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 was then washed with IPA for 1 minute and dried under nitrogen.

Lipoic acid functionalized 40 nm Au nanospheres suspended in water were 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 suspension was added to a 10 mL beaker, followed by 35 µL of freshly prepared 20 mM s-NHS in 0.1 M MES buffer and 35 µL of freshly prepared 8 mM EDC in 0.1 M MES buffer. The suspension was swirled and moved on a hotplate to be heated to 8o^{o}C. A 1 cm $\times$ 1 cm Pt mesh and 1 cm $\times$ 1 cm Si substrate were stabilized using alligator clips and placed into the suspension 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 was used to covalently bind carboxylic acid functionalized Au nanospheres to the selectively amine-functionalized PMMA regions of the template *via* an amide bond [47]. The substrate, Pt mesh, and beaker weree rinsed with IPA for 1 minute and dried under nitrogen. The process was repeated on the same substrate with a freshly concentrated Au nanospheres suspension, but with 30 µL 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 was 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 were 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.

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) onto a quartz cuvette filled with DI water.

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 is as follows. Au nanospheres of 40 nm in diameter with 0.9 nm interparticle gaps [42] are partially embedded in a 40 nm PMMA layer [48] 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. Briefly, open boundary conditions enable saving computation time on calculations involving large structures with small features by reducing the required computational domain. Here, the software considers the glass, ITO, and PMMA layers to extend beyond the computational domain, thus waves can pass this boundary with minimal

reflection. Au permittivity was calculated from Drude model with parameters extracted from Grady [49]. The relative electric permittivity used for water, PMMA, glass, and ITO are 1.77, 2.47, 2.3207, and adapted from Moerland [50], respectively. The permittivity in the gap region was unknown due to anhydride bonding [42], 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 (3.1).

$$FE = |E_{olig}|/|E_0| \tag{3.1}$$

$|E_{olig}|$ 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.

3.4 Summary

Chemically assembled SERS substrates, using EHD flow to initiate chemical crosslinking yields uniform nanogap spacings of 0.9 nm. Full-wave simulations of near-field enhancements comparing 1 and 2 nm nanogaps show that variation of this parameter is critical to achieve high and reproducible SERS intensity. Consequently, the 1573 cm^{-1} vibrational band of adsorbed benzenethiol produces intensities with RSD of 10.4 % over a 100 µm × 100 µm

area. The substrates exhibit broad band resonance between 625 nm and 875 nm, sufficient to enhance both the excitation and Raman scattered light.

Chapter 4

Robust SERS Quantitative Detection of Pyocyanin

Self-assembly of SERS substrates using EHD flow and EDC crosslinking chemistry minimizes inherent variances in SERS enhancement factor, producing optically uniform SERS surfaces (**Chapter 3**). Yet such uniformity does not guarantee high performance in molecular sensing without appropriate spectral analysis. In this chapter, I show the advantage of using PLS regression – a multivariate statistical learning algorithm – over the dated method of linear regression in analyzing SERS spectra of pyocyanin, a metabolite produced by *P. aeruginosa*. More importantly, I demonstrate an integrated microfluidic SERS device capable of early detection of *P. aeruginosa* biofilm formation *via* instantaneous quantitative detection of pyocyanin.

4.1 Pyocyanin: A Biomarker for PA

During biofilm formation, differential gene expression is regulated through a cell density-dependent mechanism called quorum sensing (QS) [51, 52]. Soon after surface attachment, bacteria begin producing extracellular polymeric substances and QS signaling molecules [53]. Once formed, a combination of physical mechanisms and genetic and metabolic adaptations within the biofilms imparts extreme antibiotic tolerance or resistance to the constituent cells [54], which can withstand up to 1000 times higher doses of antibiotics than their free floating planktonic counterparts [55]. 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 *P. aeruginosa* biofilm formation at a stage where antibiotic treatment has higher efficacy. *P. aeruginosa* is a biofilm-forming, opportunistic pathogen that is associated with the contamination of medical devices and respiratory infections in immunocompromised patients [25, 26] and is one of the most common bacteria isolated in chronic wounds [56]. Among the many virulence factors and QS compounds that *P. aeruginosa* produces is pyocyanin [57], a redox-active secondary metabolite which can act as a terminal signaling factor in the QS process. Pyocyanin has been shown to exist in concentration in the ng·mL^{-1} to µg·mL^{-1} (ppb to ppm) range in clinical samples [58]. Thus quantitative detection of pyocyanin at low concentrations may enable early detection of *P. aeruginosa* infections and earlier, more effective treatment.

4.2 Experimental Procedures

Materials

Random copolymer Poly(styrene-co-methyl methacrylate)--Hydroxyl--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*-144 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 18.2 Mohm·cm was obtained from Milli-Q Millipore System.

P. aeruginosa cell-free supernatant preparation

Wild type *P. aeruginosa* (strain PA14 [59]) 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 ^{o}C. Shaking cultures were inoculated from single colonies on the LB plate into 5 mL of liquid LB and grown at 37 ^{o}C on an orbital shaker overnight. 50 µL of the shaking culture was diluted into 24 mL of 10 g·L^{-1} tryptone media (Bacto tryptone "TB", BD Scientific) and this subculture was grown at 37 ^{o}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 × 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.

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) [60]. *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 surface-enhanced resonance Raman scattering (SERRS) measurements were collected once the in-line device is 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.

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 $\mu g \cdot mL^{-1}$) [61] 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 BacLight kit (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-attached bacteria 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 µg·mL^{-1} carbenicillin in PBS for 3 h. The input line was then changed again to the BacLight solution 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.

Spectroscopic Measurements Instrumentation and Procedure

Surface enhanced Raman spectroscopy (SERS) measurements were conducted using a Renishaw InVia Raman Microscope system. A 785 nm continuous wave laser was chosen to excite near the plasmon resonance of Au nanoparticle assemblies as determined from electromagnetic simulations and ultraviolet-visible absorption measurements [42].

In droplet measurements, as referred to in the main text, a 60X water immersion objective with 1.2 NA was used for illumination and collection. Approximately 150 µL of solution of interest was transferred onto SERS substrates and the measurements were acquired with laser power and acquisition time of 7.3 µW and 0.5 second, respectively, over area specified for each case. SERS substrates were 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 was used. Measurements were taken with laser power and acquisition time of 14.6 µW and 0.1 second, respectively. Illumination and collection were done through the glass slide of SERS microfluidic channel.

Spectra Processing and Analysis

Raman scattering spectra processing and analysis were performed off-line using MATLAB R2016b (The MathWorks Inc, Natick, MA). Each spectrum underwent baseline correction, smoothing with Savitzky-Golay, and normalization to the average intensity of Si second-order vibrational band. While this band was previously reported between 920 – 1045 cm^{-1} [62], the range 920 – 970 cm^{-1} was used due to the appearance of pyocyanin ring stretching vibrational band at 975 cm^{-1}. Preprocessing allows for comparison of different samples where slight intensity variations may arise due to deviations in optical collection in the experimental setup. 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 [33].

4.3 Results and Discussions

4.3.1 Quantification and Detection of Pyocyanin in Aqueous Media

Pyocyanin exhibits a broad absorption band from 550 to 900 nm [16, 63]; thus using a 785 nm laser to excite nanoantennas results in surface-enhanced resonance Raman scattering

(SERRS). Henceforth, the acronym SERS is used to represent the fabricated substrates, while SERRS is used to describe measurements associated with pyocyanin. SERRS spectra of aqueous pyocyanin from 1 µg·mL^{-1} (4.8 µM) to 100 µg·mL^{-1} (480 µM) (Figure 4.1a) displays clear Raman bands similar to pyocyanin spectra reported using surfaces with Ag colloids [15] and Ag and Au nanorods [16, 17] at 552, 1353, 1602, and 1620 cm^{-1}. These bands rise at a concentration as low as of 100 pg·mL^{-1} (480 pM), as exemplified in Figure 4.1b for the band 552 cm^{-1}. The log–log dose–response curve of pyocyanin at 552 cm^{-1} (Figure 4.1c) reveals a linear regime between 1 ng·mL^{-1} and 10 µg·mL^{-1}, consistent with Langmuir adsorption kinetics; similar linear relationships with concentration are also observed for other pyocyanin vibrational bands, and results are shown in Appendix A. Within this linear regime, pyocyanin concentration can be quantified using the formula shown in Equation 4.1 with $R^2 = 0.951$.

$$\log C = a \log I + b \tag{4.1}$$

Here a and b are fitting constants with values 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·mL^{-1}, SERS substrates detection of pyocyanin was determined by comparing the signal and background at 552 cm^{-1} per guideline EP17 of the Clinical and Laboratory Standards Institute [64]. The limit of blank (LOB) is calculated by adding the mean background signal [deionized (DI) water] $\bar{I}_{bg}$ at 552 cm^{-1} to 1.645 its standard deviation σ_{bg}, as shown in Equation 4.2.

$$\text{LOB} = \bar{I}_{bg} + 1.645\,\sigma_{bg} \tag{4.2}$$

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, greater than or equal to the LOB, as shown in Equation

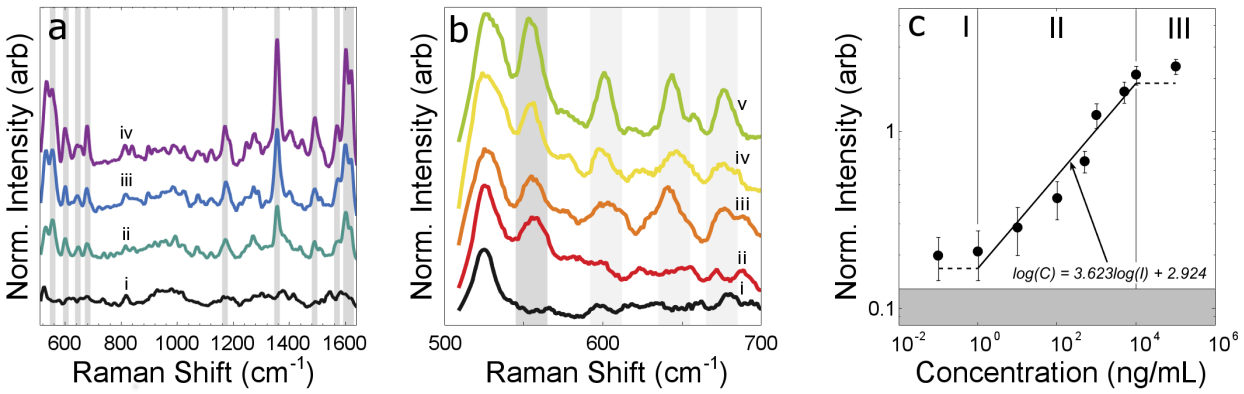

Figure 4.1: (a) SERRS spectra of (i) DI water, (ii) 1 µg·mL^{-1}, (iii) 10 µg·mL^{-1}, and (iv) 100 µg·mL^{-1} pyocyanin in water. Grey bars indicate Raman bands of pyocyanin. (b) SERRS spectra of of (i) DI water, (ii) 100 pg·mL^{-1}, (iii) 1 ng·mL^{-1}, (iv) 10 ng·mL^{-1}, and (v) 100 ng·mL^{-1} pyocyanin 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 calculated from 100 measurements.

4.3.

$$\bar{I}_D - 1.645\,\sigma_D \geq \text{LOB} \tag{4.3}$$

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 4.1c. We did not measure pyocyanin at lower concentration since our interest was in establishing a limit of detection (LOD) in biological media. These results are discussed in the following section.

4.3.2 Predicting Pyocyanin Concentration with PLS Regression

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 [15, 16, 17, 18, 19, 20, 21, 22, 23]. While this method is sufficient for the analysis of a pure analyte in water,

it discards the remaining rich spectral 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 LOQs in complex media [65]. Thus, we capitalize on the uniform SERS response of our surfaces to acquire necessary training data and employ PLS analysis to quantify pyocyanin in a 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 of cell-free supernatants collected from mid-stationary phase cultures of wild-type *P. aeruginosa* PA14 and its phenazine-deficient mutant strain $\Delta phz1/2$ as a control (Figure 4.2a). The $\Delta phz1/2$ 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 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 root mean square error of cross validation (RMSECV) is observed, displayed in inset of Figure 4.2b. With the testing set, the model demonstrates accurate prediction between 1 ng·mL^{-1} and 100 µg·mL^{-1} as shown in Figure 4.2b. 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.

4.3.3 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*phz*1/2 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 4.2c. 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 [66].

The pyocyanin-deficient Δ*phz*1/2 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*phz*1/2 strain. This value agrees well with the reported LOD of 1 µM for UV-vis [67]. The UV-vis data is also plotted in Figure 4.2c for comparison with SERRS data. The LOD of SERRS observed in Figure 4.2b 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

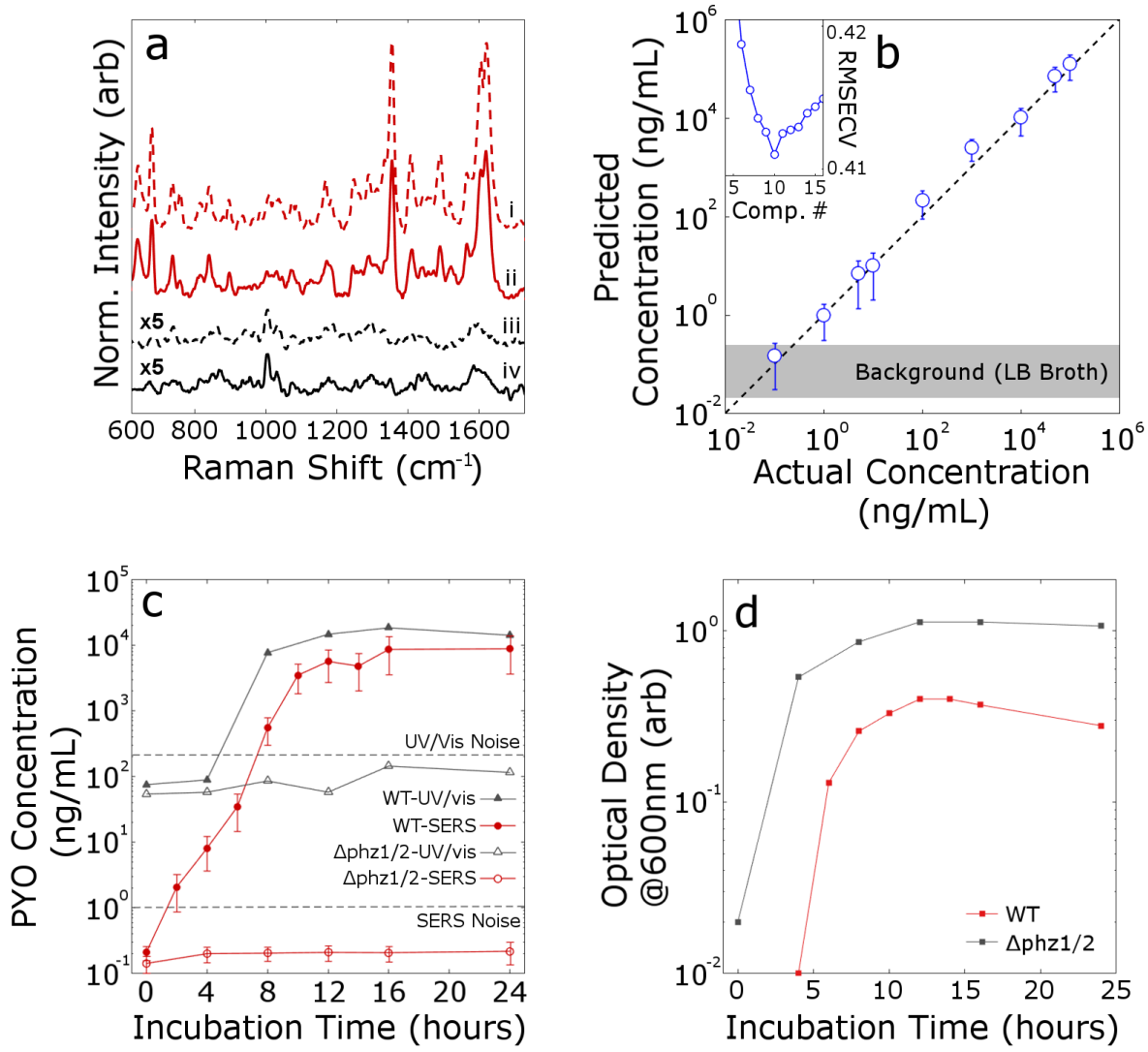

Figure 4.2: (a) SERRS fingerprint of (i) bacteria-free supernatant from mid-stationary phase cultures of wild-type *P. aeruginosa* PA14, (ii) 10 µg·mL^{-1} pyocyanin in LB broth, (iii) phenzine-null mutant strain $\Delta phz1/2$ and (iv) LB broth. Spectra in (i) and (ii) are multipled by a factor of 5. (b) Pyocyanin concentration predicted by PLS model for spectra in testing set. Error bars depict standard deviation calculated from 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, respectively. Error bars show standard deviation calculated from 400 measurements. (d) Growth curves for planktonic cultures of *P. aeruginosa* PA14 and its phenazine-null mutant strain as determined from optical density.

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 4.2d 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.

4.3.4 Monitoring Biofilm Formation *via* Pyocyanin Quantification

Here, we integrate microfluidic channels with SERS substrates to perform in-line sampling of biofilm's 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 [60]. Effluent from the biofilm growth channel is delivered to a microfluidic channel with Au oligomers assembled on the surface as illustrated in Figure 4.3a. Every 3 hours, 200 SERRS spectra are collected in the effluent with a total acquisition time of 2 min, thus eliminating the need for performing chloroform extraction and incubating [16], or evaporating the solvent on SERS surfaces [17]. To quantify pyocyanin, a suitable predictive model was generated by repeating PLS analysis on the training dataset acquired using the appropriate collection parameters for TB media (as opposed 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 and 9 h after inoculating as observed in Figure 4.3b. 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. The half-life of pyocyanin desorption from the surface was determined to be on the order of

11 min from subsequent SERS measurements from samples immersed in water, exposed to 1 µg·mL^{-1}(4.8 µM) pyocyanin, rinsed, and reimmersed in water, as shown in Appendix B. Thus, the calculated concentrations are indicative of the instantaneous production of pyocyanin in the 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 4.3c. This fraction increases to 70.3, 96.7, and 100 % for the subsequent time points. This is consistent with nonuniform distribution of pyocyanin on the substrate surface at low concentrations. Hence, while the averaged concentration of pyocyanin calculated for the set of Raman spectra at 3 and 6 h of the biofilm growth lies below the LOD, a fraction of those spectra shows that pyocyanin is detected at those early time points. Thus, 3 h is set as the time of detection (TOD). The calculated concentrations of pyocyanin correlate with the time-dependent accumulation of biofilm biomass obtained from the analysis of fluorescence images of the microfluidic growth channels (Figure 4.3d). Representative confocal fluorescence images of the growth channels show the initial stages of bacterial cell adhesion (0-6 h), microcolony formation (10 h), and three-dimensional growth of biofilms above the channel surface (16 h). The biomass and imaging data indicates that the 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 that observed in fluorescence confocal images.

Bacterial biofilms impart antibiotic resistance and tolerance on constituent cells *via* several distinct mechanism [54]. These mechanisms are characteristic of bacterial phenotypes and properties of mature biofilms; hence, 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

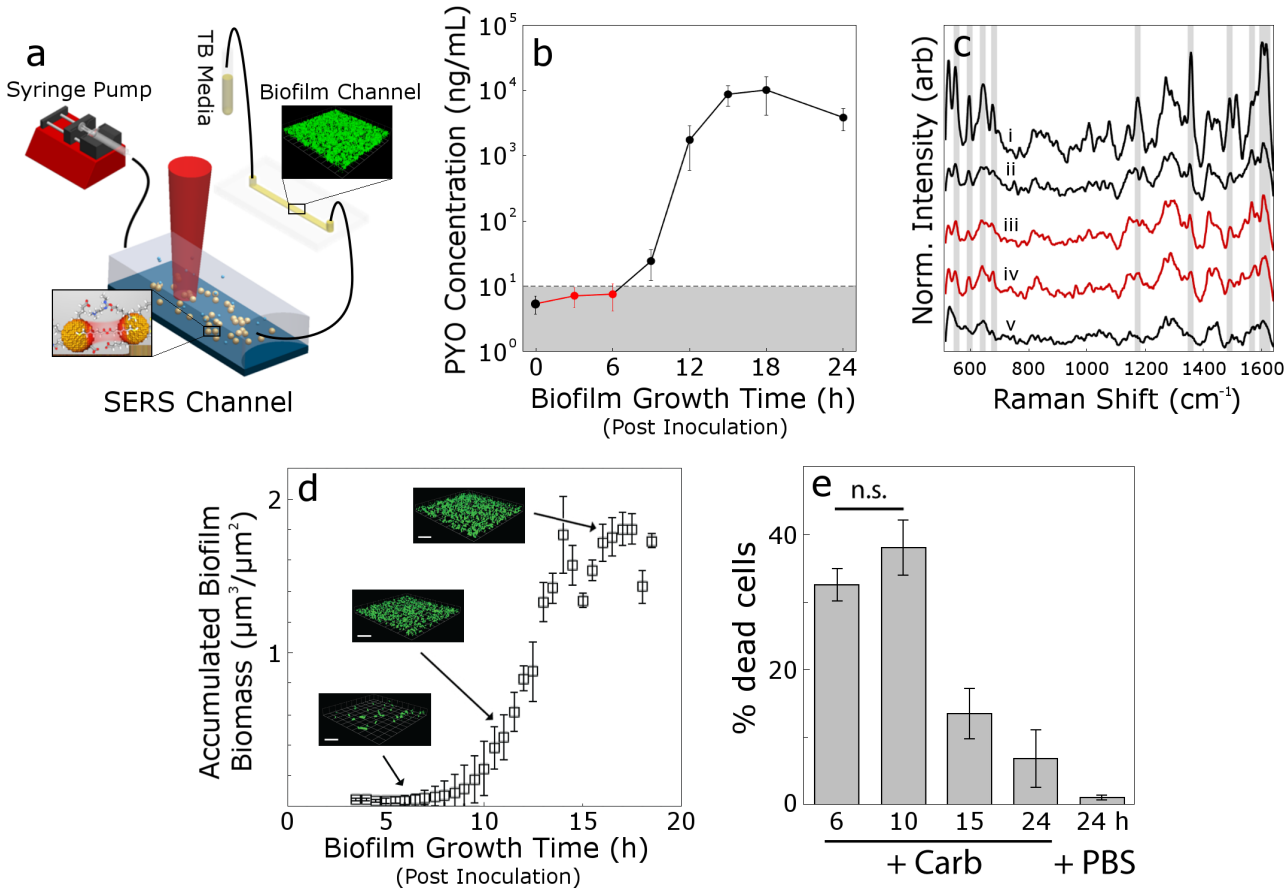

Figure 4.3: (a) Schematic of the biofilm inline measurements setup. (b) Pyocyanin concentration predicted by PLS model from SERS spectra of various time. Each data point represents the averaged of predicted concentrations and error bars show standard deviation calculated from 200 measurements. (c) Individual 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 calculated 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-attached bacteria to carbenicillin treatment at the specified growth times. Error bars depict standard deviation of dead cell fraction calculated from n = 3 biological replicates.

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 antibiotic-tolerant state (Figure 4.3e); this transition is also captured as the sharp increase in pyocyanin concentration measured from SERRS

in Figure 4.3b between 9 and 12 h. A control biofilm was grown for 24 h and exposed to phosphate-buffered saline (PBS) instead of carbenicillin indicates no statistically significant difference between the antibiotic susceptibility of cells grown in flow channels for 6 and 10 h. Treatment of biofilms after 15 and 24 h of growth with the same concentration of carbenicillin, however, resulted in significantly reduced the effectiveness of the antibiotic. Consequently, the detection of a *P. aeruginosa* infection and intervention at early stages of colonization (<10 h growth) show the potential to substantially improve the effectiveness of antibiotic treatment.

4.4 Summary

SERS substrates, with nanostructures assembled using methods of Chapter 3, provide uniform and reproducible SERS intensity over 100 × 100 μm^2. This allows for integration of SERS substrates in a device platform and the rapid acquisition of large data sets for statistical analysis. SERS substrates exhibit the ability to detect pyocyanin in aqueous media at concentration of 100 pg·mL^{-1} when individual pyocyanin Raman bands are investigated. In more complex media, PLS analysis of spectral data enables robust quantification of pyocyanin spanning 5-orders of magnitude in biologically relevant levels between 1 ng·mL^{-1} and 100 μg·mL^{-1}. Integration of microfluidic architectures with SERS substrates facilitates inline sampling of effluent medium for longitudinal monitoring of pyocyanin concentration during *P. aeruginosa* biofilm formation. Measured pyocyanin concentration in effluent medium correlates with accumulated biofilm biomass obtained from confocal fluorescence images. More significantly, pyocyanin can be detected as early as 3 h after inoculation. A sharp increase in pyocyanin concentration is observed between 9 and 12 h after inoculation; this corresponds to decreased susceptibility of surface-attached bacteria to bactericidal antibiotic carbenicillin after 10 h. Thus detecting microbial production of metabolites associated

with quorum sensing with chemically assembled SERS substrates and analysis using machine learning algorithms lays out a promising strategy for detection of metabolites in biological media that can be pursued in future studies for more complex systems.

Chapter 5

Deep Learning for Improved SERS Quantitative Detection

In the last decade, improved computing power (from modern GPUs) and more accessible code (from open-source frameworks) has led to significant developments in the field of artificial intelligence/machine learning (AI/ML). Deep learning, a branch of AI/ML characterized by the use of deep artificial neural networks (ANNs), has been among the largest beneficiaries of this development. Continuous breakthroughs had already resulted in everyday use-cases such as Facebook's face identification feature, improved Google Translate, and even self-driving cars. More significantly, open-source deep learning frameworks developed by institutions ranging from universities (Theano) to Google (TensorFlow) and Microsoft (CNTK) now provide scientists and engineers with access to deep learning tools previously available only to AI/ML researchers. In this chapter, I develop models that map SERS spectra to concentration of pyocyanin using deep learning tools, namely feedforward ANNs and convolutional NNs (CNNs). Using the same training set from **Chapter 4.3.1**, the NNs exhibit superior performances in both quality of fit and accuracy of predictions.

5.1 Feedforward Artificial Neural Networks

5.1.1 Hyperparameters Tuning

In deep learning, choosing the right hyperparameters (i.e. learning rate, activation function, optimization algorithm, etc) can be the difference between a useful model and a useless one. In this work, hyperparameters tuning is carried out using Bayesian optimization with Gaussian process priors as it reduces the amount of trials and time required to search the hyperparameters space [68]. Table 5.1 outlines the parameters to be tuned, their respective value ranges, and the resulting optimal values.

5.1.2 Limit of Quantification, R^2, and Prediction Error

Using the resulting Model 1, 2, and 3 consistently produces LOQs of 100 pg·mL^{-1} (100 ppt in Figure 5.1). Interestingly, it seems that the obtained training data does not include sufficiently low concentrations to reach achievable detection limits. When fitting a line with slope of 1 representing perfect accuracy in prediction R^2 values of 0.991, 0.987, and 0.991 are obtained. Note that they are all higher than the R^2 of 0.951 and 0.956 from linear algorithms used in **Chapter 4**. Moreover, over-fitting is not taking place, as evidenced in the training and validation losses monitored during training. The prediction errors (in %) of each model is calculated in two steps: (1) the mean of prediction errors is taken for each concentration, and (2) the mean of this value over all concentrations is the prediction error for that model. The calculations are summarized in Equation 5.1.

$$\text{Prediction Error (\%)} = \frac{1}{N}\sum_{i}^{N}\frac{\frac{1}{M}\sum_{j}^{M}|C_{i,j}-\hat{C}_i|}{\hat{C}_i} \tag{5.1}$$

$C_{i,j}$ represents the j^{th} predicted concentration for concentration set i^{th} during testing, $\hat{C}_i$

(Layer) Parameter	Range	Model 1	Model 2	Model 3
(1) Hidden nodes	{16, 32, 64, 128, 256}	128	64	64
(1) Activation	{ReLU, Leaky ReLU, ELU}	Leaky ReLU	ReLU	ReLU
(2) Hidden nodes	{16, 32, 64, 128, 256}	16	16	256
(2) Activation	{ReLU, Leaky ReLU, ELU}	ReLU	ELU	ELU
(3) Hidden nodes	{16, 32, 64, 128, 256}	128	64	16
(3) Activation	{ReLU, Leaky ReLU, ELU}	ReLU	Leaky ReLU	ReLU
(4) Hidden nodes	{16, 32, 64, 128, 256}	32	256	256
(4) Activation	{ReLU, Leaky ReLU, ELU}	ReLU	ReLU	ELU
(5) Hidden nodes	{16, 32, 64, 128, 256}	64	128	16
(5) Activation	{ReLU, Leaky ReLU, ELU}	Leaky ReLU	ReLU	ELU
-Log(learning rate)	4—1	2.15	2.93	2.18
Nesterov Momentum	{True, False}	False	False	True
Epochs	{10, 20,. . . ,100}	100	100	60

Table 5.1: Summary of hyperparameters range and optimized values for feedforward ANNs.

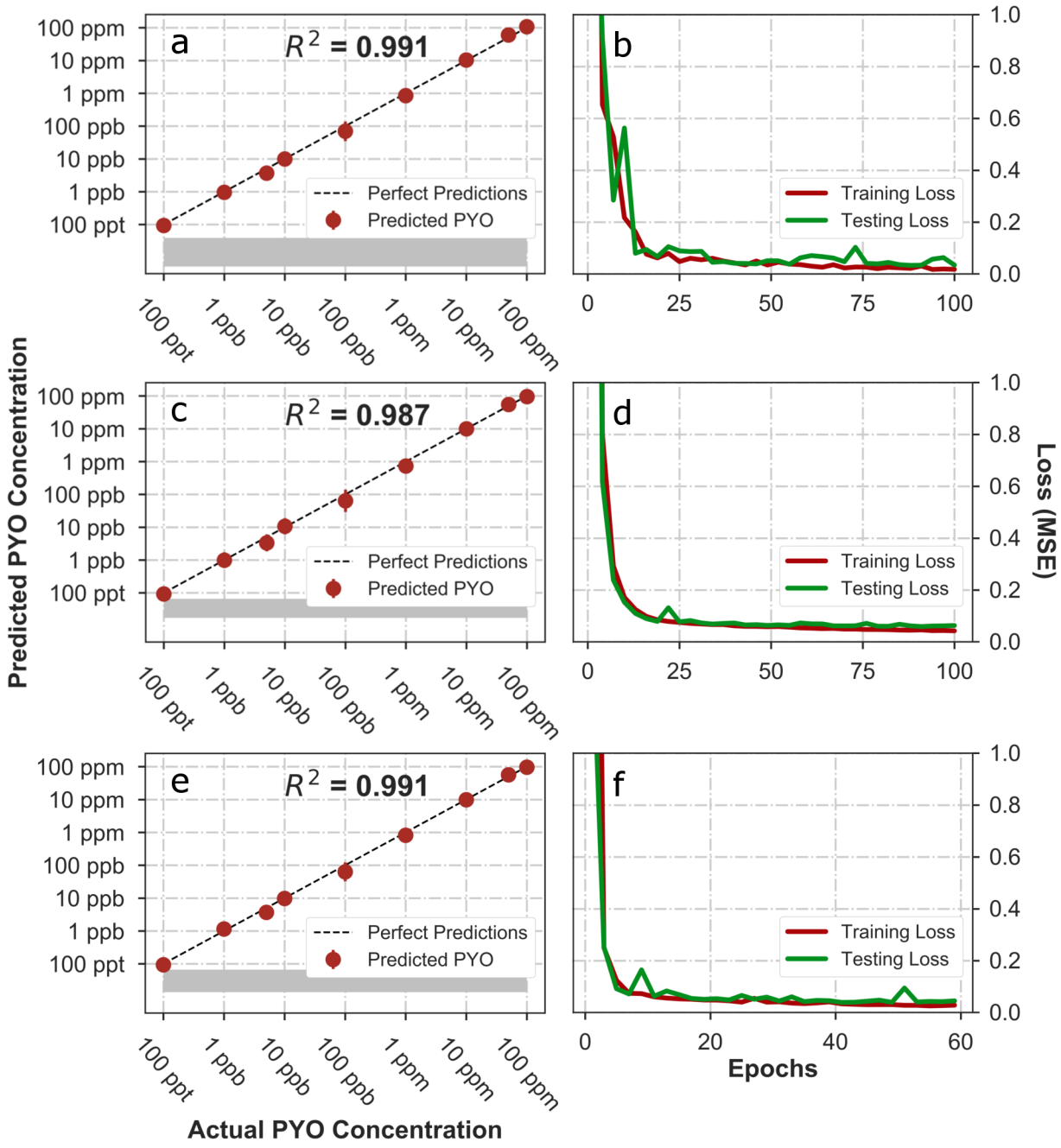

Figure 5.1: Predicted pyocyanin concentrations and losses during training of feedforward ANNs (a)(b) Model 1, (c)(d) Model 2, and (e)(f) Model 3.

represents the true value of concentration set i^{th}, N and M are the number of concentration sets and number of samples in each concentration set, respectively. This calculation produces prediction errors of 5.8 %, 7.7 %, and 5.1 % for an average of 6.2 ± 1.1 %. It is worth noting that spectra only need individual scaling to achieve this performance, unlike the extensive preprocessing required in **Chapter 4**.

5.2 Convolutional Neural Networks

While using feedforward ANNs drastically improve quantification capabilities, the models can further benefit from learning and analyzing multiple spectra at once. This is accomplished by using 1D CNNs, which have been shown to perform well on data having spatial relations in a single dimension, such as stock price [69], electroencephalograms (EEGs) [70], audio signals [71], and spectra. In short, a group of 5 spectra, termed spectra stack, is fed to the network as one training sample with dimensions of (1011 × 5). As shown below, this method increases R^2 to a maximum of 0.998 and reduces the average prediction error by an additional 1.3 % to 4.9 ± 0.9 %.

5.2.1 Hyperparameters Tuning

Hyperparameters are tuned similarly to feedforward ANNs using Bayesian optimization. CNNs traditionally require more epochs to sufficiently train than feedforward ANNs, thus the upper limit of epochs range is increased to 200. Table 5.2 details the parameters, their respective range of values, and the resulting architectures.

(Layer) Parameter	Range	Model 1	Model 2	Model 3
(1) Filter	{8, 16, 32}	16	8	8
(1) Kernel length	{16, 32, 64}	32	16	16
(1) Activation	{ReLU, Leaky ReLU, ELU}	ReLU	ReLU	ReLU
(2) Filter	{16, 32, 64}	64	32	32
(2) Kernel length	{8, 16, 32}	32	8	8
(2) Activation	{ReLU, Leaky ReLU, ELU}	Leaky ReLU	Leaky ReLU	Leaky ReLU
(3) Filter	{32, 64, 128}	128	64	64
(3) Kernel length	{4, 8, 16}	16	8	8
(3) Activation	{ReLU, Leaky ReLU, ELU}	ReLU	ReLU	ReLU
(4) Hidden nodes	{8, 16, 32}	8	16	16
(4) Activation	{ReLU, Leaky ReLU, ELU}	ReLU	ReLU	ReLU
-Log(learning rate)	4—1	3.95	3.35	2.41
Nesterov Momentum	{True, False}	False	True	True
Epochs	{10, 20,. . . ,200}	200	200	200

Table 5.2: Summary of hyperparameters range and optimized values for 1D CNNs.

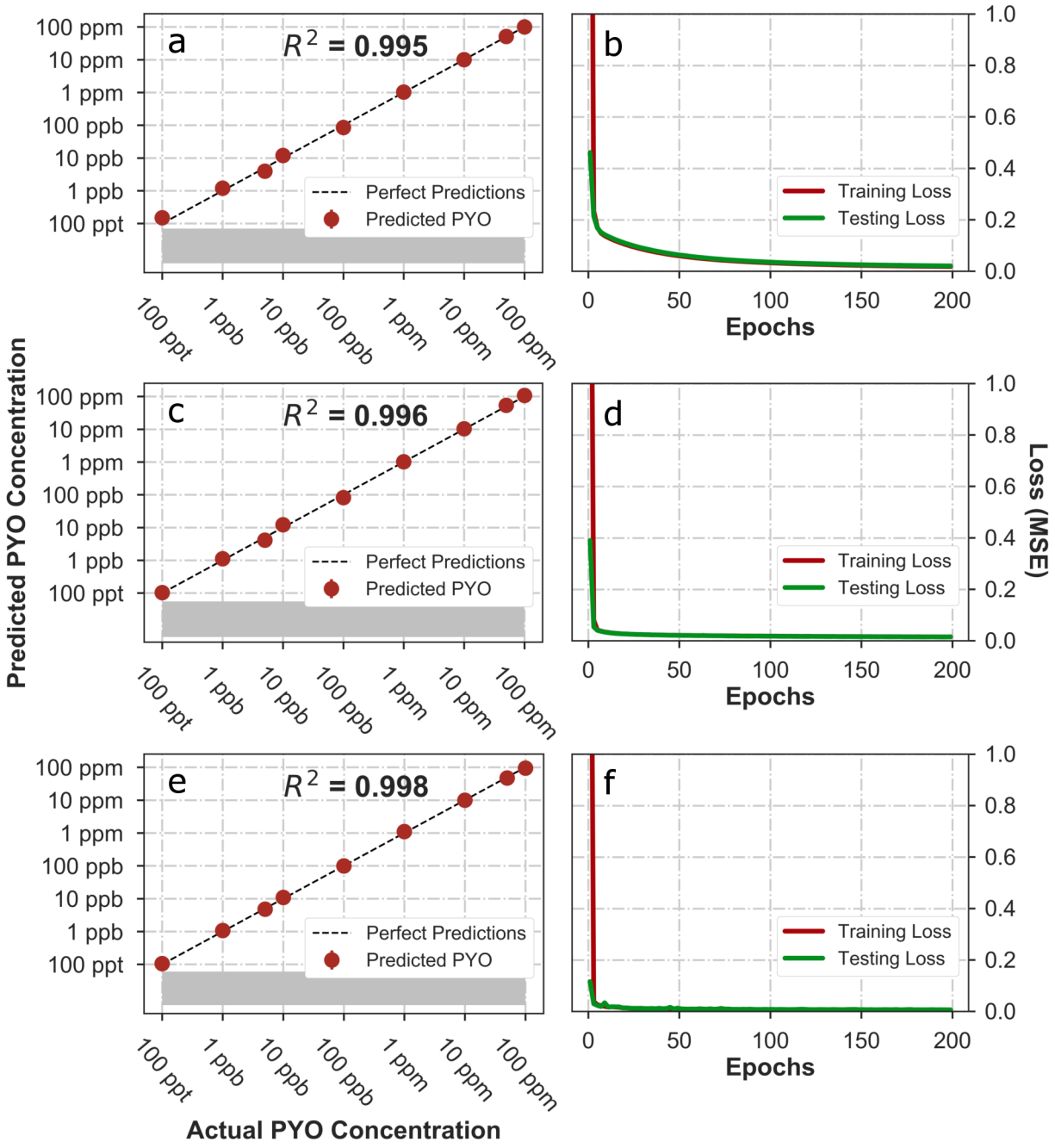

Figure 5.2: Predicted pyocyanin concentrations and losses during training of 1D CNNs (a)(b) Model 1, (c)(d) Model 2, and (e)(f) Model 3.

5.2.2 Limit of Quantification, R^2, and Prediction Error

1D CNNs Model 1, 2, and 3 also reproducibly exhibit LOQs of 100 pg·mL^{-1}, as shown in Figure 5.2. Class activation mapping, a popular technique used in computer vision to visualize how CNNs make predictions [72], shows the networks devote high importance on pyocyanin vibrational bands (Appendix C). Over-fitting is again prevented by monitoring training and testing loss. Here R^2 values are calculated to be 0.995, 0.996, and 0.998 – an improvement over the already accurate results of feedforward ANNs. Finally, using Equation 5.1, prediction error of each models are 5.4 %, 5.7 %, and 3.6 %, thus further reducing the average prediction error to 4.9 ± 0.9 %.

Interestingly 1D CNNs also exhibit better generalizing ability than feedforward ANNs, as evidenced by the lack of spikes in testing losses for all three models during training. This can be attributed to the use of spectra stack. Intuitively, using spectra stacks as training samples is somewhat equivalent to averaging the predicted concentrations from 5 individual spectra; the key difference is that spectra stacks let the models recognize the possible variances before making a prediction.

5.3 Summary

Nonlinearities in deep learning algorithms, when employed to analyze SERS spectra, improve prediction accuracy and precision in comparison to PLS regression. Specifically, two architectures – feedforward ANNs, and 1D CNNs – were investigated. Bayesian optimization was used to search for optimal hyperparameters, resulting in three models per architecture for a total of six models. Conventional feedforward ANNs achieve accurate predictions with $R^2_{max} = 0.991$ and average prediction error of 6.2 ± 1.1 %. Meanwhile 1D CNNs, using spectra stacks as inputs, further improve these values while maintaining low LOQ of 100

pg·mL^{-1} to $R^2_{max} = 0.998$ and average prediction error of 4.9 ± 0.9 %. The results presented in this chapter pave the way for NNs to be a valuable spectral analysis technique. More significantly, they reveal the potential to fully capitalize on SERS sensing capability – a task pursued by many researchers since the discovery of SERS – by employing deep learning algorithms.

Chapter 6

Conclusion

While there still exists a significant amount of hurdles to overcome before a commercial label-free SERS sensing platform is realized, the purpose of this thesis is to present potential solutions to two of those challenges: inherent variances in SERS EF and inadequate spectral analysis techniques. In short, self-assembly using EHD flow and chemical crosslinking minimizes inherent variances of SERS EF by driving reactions in plasmonic nanogaps to produce uniform interparticle distance of 0.9 nm. The resulting SERS substrates exhibit uniform optical response: SERS signals from benzenethiol adsorbed to the nanostructures have RSD of 10.4% over 100 µm × 100 µm area. Regarding spectral analysis, PLS regression enables the detection and quantification of pyocyanin in complex media from 1 ng·mL^{-1} to 100 µg·mL^{-1} ($R^2 = 0.956$) and facilitates early detection of biofilm formation as early as 3 h after inoculation. The most drastic improvement takes place when deep learning is employed. Feedforward ANNs tuned using Bayesian optimization lowers LOQ from 1 ng·mL^{-1} to 100 pg·mL^{-1} with $R^2_{max} = 0.991$ and prediction errors of 6.2 ± 1.1 %. More impressively, 1D CNNs trained with spectra stack has $R^2_{max} = 0.998$ and further reduces prediction errors to 4.9 ± 0.9 %.

Bibliography

[1] K. Kneipp, Y. Wang, H. Kneipp, L. T. Perelman, I. Itzkan, R. R. Dasari, and M. S. Feld, "Single molecule detection using surface-enhanced raman scattering (sers)," *Physical review letters*, vol. 78, no. 9, p. 1667, 1997.

[2] J. A. Bowden, A. Heckert, C. Z. Ulmer, C. M. Jones, J. P. Koelmel, L. Abdullah, L. Ahonen, Y. Alnouti, A. Armando, J. M. Asara, *et al.*, "Harmonizing lipidomics: Nist interlaboratory comparison exercise for lipidomics using standard reference material 1950 metabolites in frozen human plasma," *Journal of Lipid Research*, pp. jlr–M079012, 2017.

[3] A. A. Aksenov, R. da Silva, R. Knight, N. P. Lopes, and P. C. Dorrestein, "Global chemical analysis of biology by mass spectrometry," *Nature Reviews Chemistry*, vol. 1, no. 7, p. 0054, 2017.

[4] I. Pavel, E. McCarney, A. Elkhaled, A. Morrill, K. Plaxco, and M. Moskovits, "Label-free sers detection of small proteins modified to act as bifunctional linkers," *The Journal of Physical Chemistry C*, vol. 112, no. 13, pp. 4880–4883, 2008.

[5] Y. C. Cao, R. Jin, and C. A. Mirkin, "Nanoparticles with raman spectroscopic fingerprints for dna and rna detection," *Science*, vol. 297, no. 5586, pp. 1536–1540, 2002.

[6] B. Lorenz, C. Wichmann, S. Stöckel, P. Rösch, and J. Popp, "Cultivation-free raman spectroscopic investigations of bacteria," *Trends in microbiology*, vol. 25, no. 5, pp. 413–424, 2017.

[7] X. Wu, Y.-W. Huang, B. Park, R. A. Tripp, and Y. Zhao, "Differentiation and classification of bacteria using vancomycin functionalized silver nanorods array based surface-enhanced raman spectroscopy and chemometric analysis," *Talanta*, vol. 139, pp. 96–103, 2015.

[8] M. Ghebremedhin, R. Heitkamp, S. Yesupriya, B. Clay, and N. J. Crane, "Accurate and rapid differentiation of acinetobacter baumannii strains by raman spectroscopy: a comparative study," *Journal of clinical microbiology*, vol. 55, no. 8, pp. 2480–2490, 2017.

[9] J.-M. Nam, J.-W. Oh, H. Lee, and Y. D. Suh, "Plasmonic nanogap-enhanced raman scattering with nanoparticles," *Accounts of chemical research*, vol. 49, no. 12, pp. 2746–2755, 2016.

[10] J.-H. Lee, J.-M. Nam, K.-S. Jeon, D.-K. Lim, H. Kim, S. Kwon, H. Lee, and Y. D. Suh, "Tuning and maximizing the single-molecule surface-enhanced raman scattering from dna-tethered nanodumbbells," *ACS nano*, vol. 6, no. 11, pp. 9574–9584, 2012.

[11] W. Zhu and K. B. Crozier, "Quantum mechanical limit to plasmonic enhancement as observed by surface-enhanced raman scattering," *Nature communications*, vol. 5, p. 5228, 2014.

[12] E. C. Le Ru, J. Grand, I. Sow, W. R. Somerville, P. G. Etchegoin, M. Treguer-Delapierre, G. Charron, N. Félidj, G. Lévi, and J. Aubard, "A scheme for detecting every single target molecule with surface-enhanced raman spectroscopy," *Nano letters*, vol. 11, no. 11, pp. 5013–5019, 2011.

[13] E. C. Le Ru, M. Meyer, and P. G. Etchegoin, "Proof of single-molecule sensitivity in surface enhanced raman scattering (sers) by means of a two-analyte technique," *The journal of physical chemistry B*, vol. 110, no. 4, pp. 1944–1948, 2006.

[14] A. B. Zrimsek, N. L. Wong, and R. P. Van Duyne, "Single molecule surface-enhanced raman spectroscopy: a critical analysis of the bianalyte versus isotopologue proof," *The Journal of Physical Chemistry C*, vol. 120, no. 9, pp. 5133–5142, 2016.

[15] O. Žukovskaja, I. J. Jahn, K. Weber, D. Cialla-May, and J. Popp, "Detection of pseudomonas aeruginosa metabolite pyocyanin in water and saliva by employing the sers technique," *Sensors*, vol. 17, no. 8, p. 1704, 2017.

[16] G. Bodelón, V. Montes-García, V. López-Puente, E. H. Hill, C. Hamon, M. N. Sanz-Ortiz, S. Rodal-Cedeira, C. Costas, S. Celiksoy, I. Pérez-Juste, *et al.*, "Detection and imaging of quorum sensing in pseudomonas aeruginosa biofilm communities by surface-enhanced resonance raman scattering," *Nature materials*, vol. 15, no. 11, p. 1203, 2016.

[17] X. Wu, J. Chen, X. Li, Y. Zhao, and S. M. Zughaier, "Culture-free diagnostics of pseudomonas aeruginosa infection by silver nanorod array based sers from clinical sputum samples," *Nanomedicine: Nanotechnology, Biology and Medicine*, vol. 10, no. 8, pp. 1863–1870, 2014.

[18] G. Qi, K. Jia, C. Fu, S. Xu, and W. Xu, "A highly sensitive sers sensor for quantitative analysis of glucose based on the chemical etching of silver nanoparticles," *Journal of Optics*, vol. 17, no. 11, p. 114020, 2015.

[19] D. S. Grubisha, R. J. Lipert, H.-Y. Park, J. Driskell, and M. D. Porter, "Femtomolar detection of prostate-specific antigen: an immunoassay based on surface-enhanced raman scattering and immunogold labels," *Analytical chemistry*, vol. 75, no. 21, pp. 5936–5943, 2003.

[20] J. Yan, X. Han, J. He, L. Kang, B. Zhang, Y. Du, H. Zhao, C. Dong, H.-L. Wang, and P. Xu, "Highly sensitive surface-enhanced raman spectroscopy (sers) platforms based on silver nanostructures fabricated on polyaniline membrane surfaces," *ACS applied materials & interfaces*, vol. 4, no. 5, pp. 2752–2756, 2012.

[21] D. Lee, S. Lee, G. H. Seong, J. Choo, E. K. Lee, D.-G. Gweon, and S. Lee, "Quantitative analysis of methyl parathion pesticides in a polydimethylsiloxane microfluidic channel using confocal surface-enhanced raman spectroscopy," *Applied spectroscopy*, vol. 60, no. 4, pp. 373–377, 2006.

[22] R. Stevenson, A. Ingram, H. Leung, D. C. McMillan, and D. Graham, "Quantitative serrs immunoassay for the detection of human psa," *Analyst*, vol. 134, no. 5, pp. 842–844, 2009.

[23] K. R. Theis, A. Venkataraman, J. A. Dycus, K. D. Koonter, E. N. Schmitt-Matzen, A. P. Wagner, K. E. Holekamp, and T. M. Schmidt, "Symbiotic bacteria appear to mediate hyena social odors," *Proceedings of the National Academy of Sciences*, vol. 110, no. 49, pp. 19832–19837, 2013.

[24] K. E. Shafer-Peltier, C. L. Haynes, M. R. Glucksberg, and R. P. Van Duyne, "Toward a glucose biosensor based on surface-enhanced raman scattering," *Journal of the American Chemical Society*, vol. 125, no. 2, pp. 588–593, 2003.

[25] R. C. Hunter, V. Klepac-Ceraj, M. M. Lorenzi, H. Grotzinger, T. R. Martin, and D. K. Newman, "Phenazine content in the cystic fibrosis respiratory tract negatively correlates with lung function and microbial complexity," *American journal of respiratory cell and molecular biology*, vol. 47, no. 6, pp. 738–745, 2012.

[26] K. L. Whiteson, S. Meinardi, Y. W. Lim, R. Schmieder, H. Maughan, R. Quinn, D. R. Blake, D. Conrad, and F. Rohwer, "Breath gas metabolites and bacterial metagenomes from cystic fibrosis airways indicate active ph neutral 2, 3-butanedione fermentation," *The ISME journal*, vol. 8, no. 6, p. 1247, 2014.

[27] P. F. Bernath, *Spectra of atoms and molecules*. Oxford University Press, 2015.

[28] E. Smith and G. Dent, *Modern Raman spectroscopy: a practical approach*. John Wiley & Sons, 2013.

[29] S. A. Maier, *Plasmonics: fundamentals and applications*. Springer Science & Business Media, 2007.

[30] N. Liver, A. Nitzan, and J. Gersten, "Local fields in cavity sites of rough dielectric surfaces," *Chemical physics letters*, vol. 111, no. 4-5, pp. 449–454, 1984.

[31] H. Cang, A. Labno, C. Lu, X. Yin, M. Liu, C. Gladden, Y. Liu, and X. Zhang, "Probing the electromagnetic field of a 15-nanometre hotspot by single molecule imaging," *Nature*, vol. 469, no. 7330, p. 385, 2011.

[32] E. Le Ru, P. Etchegoin, and M. Meyer, "Enhancement factor distribution around a single surface-enhanced raman scattering hot spot and its relation to single molecule detection," *The Journal of chemical physics*, vol. 125, no. 20, p. 204701, 2006.

[33] P. Geladi and B. R. Kowalski, "Partial least-squares regression: a tutorial," *Analytica chimica acta*, vol. 185, pp. 1–17, 1986.

[34] M. C. Denham, "Choosing the number of factors in partial least squares regression: estimating and minimizing the mean squared error of prediction," *Journal of Chemometrics: A Journal of the Chemometrics Society*, vol. 14, no. 4, pp. 351–361, 2000.

[35] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proceedings of the 27th international conference on machine learning (ICML-10)*, pp. 807–814, 2010.

[36] P. J. Werbos, "Backpropagation through time: what it does and how to do it," *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550–1560, 1990.

[37] M. Green and F. M. Liu, "Sers substrates fabricated by island lithography: the silver/pyridine system," *The Journal of Physical Chemistry B*, vol. 107, no. 47, pp. 13015–13021, 2003.

[38] L. Petti, R. Capasso, M. Rippa, M. Pannico, P. La Manna, G. Peluso, A. Calarco, E. Bobeico, and P. Musto, "A plasmonic nanostructure fabricated by electron beam lithography as a sensitive and highly homogeneous sers substrate for bio-sensing applications," *Vibrational Spectroscopy*, vol. 82, pp. 22–30, 2016.

[39] N. A. Abu Hatab, J. M. Oran, and M. J. Sepaniak, "Surface-enhanced raman spectroscopy substrates created via electron beam lithography and nanotransfer printing," *ACS nano*, vol. 2, no. 2, pp. 377–385, 2008.

[40] A. Gopinath, S. V. Boriskina, B. M. Reinhard, and L. Dal Negro, "Deterministic aperiodic arrays of metal nanoparticles for surface-enhanced raman scattering (sers)," *Optics Express*, vol. 17, no. 5, pp. 3741–3753, 2009.

[41] C. David, N. Guillot, H. Shen, T. Toury, and M. L. de La Chapelle, "Sers detection of biomolecules using lithographed nanoparticles towards a reproducible sers biosensor," *Nanotechnology*, vol. 21, no. 47, p. 475501, 2010.

[42] W. J. Thrift, C. Q. Nguyen, M. Darvishzadeh-Varcheie, S. Zare, N. Sharac, R. N. Sanderson, T. J. Dupper, A. I. Hochbaum, F. Capolino, M. J. Abdolhosseini Qomi, *et al.*, "Driving chemical reactions in plasmonic nanogaps with electrohydrodynamic flow," *ACS nano*, vol. 11, no. 11, pp. 11317–11329, 2017.

[43] P. J. Sides, "Electrohydrodynamic particle aggregation on an electrode driven by an alternating electric field normal to it," *Langmuir*, vol. 17, no. 19, pp. 5791–5800, 2001.

[44] M. Trau, D. Saville, and I. Aksay, "Assembly of colloidal crystals at electrode interfaces," *Langmuir*, vol. 13, no. 24, pp. 6375–6381, 1997.

[45] T. M. Squires and M. Z. Bazant, "Induced-charge electro-osmosis," *Journal of Fluid Mechanics*, vol. 509, pp. 217–252, 2004.

[46] A. L. Koh, K. Bao, I. Khan, W. E. Smith, G. Kothleitner, P. Nordlander, S. A. Maier, and D. W. McComb, "Electron energy-loss spectroscopy (eels) of surface plasmons in single silver nanoparticles and dimers: influence of beam damage and mapping of dark modes," *ACS nano*, vol. 3, no. 10, pp. 3015–3022, 2009.

[47] S. M. Adams, S. Campione, J. D. Caldwell, F. J. Bezares, J. C. Culbertson, F. Capolino, and R. Ragan, "Non-lithographic sers substrates: Tailoring surface chemistry for au nanoparticle cluster assembly," *Small*, vol. 8, no. 14, pp. 2239–2249, 2012.

[48] S. M. Adams, S. Campione, F. Capolino, and R. Ragan, "Directing cluster formation of au nanoparticles from colloidal solution," *Langmuir*, vol. 29, no. 13, pp. 4242–4251, 2013.

[49] N. K. Grady, N. J. Halas, and P. Nordlander, "Influence of dielectric function properties on the optical response of plasmon resonant metallic nanoparticles," *Chemical Physics Letters*, vol. 399, no. 1-3, pp. 167–171, 2004.

[50] R. J. Moerland and J. P. Hoogenboom, "Subnanometer-accuracy optical distance ruler based on fluorescence quenching by transparent conductors," *Optica*, vol. 3, no. 2, pp. 112–117, 2016.

[51] T. R. De Kievit, R. Gillis, S. Marx, C. Brown, and B. H. Iglewski, "Quorum-sensing genes in pseudomonas aeruginosa biofilms: their role and expression patterns," *Applied and environmental microbiology*, vol. 67, no. 4, pp. 1865–1873, 2001.

[52] M. B. Miller and B. L. Bassler, "Quorum sensing in bacteria," *Annual Reviews in Microbiology*, vol. 55, no. 1, pp. 165–199, 2001.

[53] R. Kolter and E. P. Greenberg, "Microbial sciences: the superficial life of microbes," *Nature*, vol. 441, no. 7091, p. 300, 2006.

[54] G. Anderson and G. O'toole, "Innate and induced resistance mechanisms of bacterial biofilms," in *Bacterial Biofilms*, pp. 85–105, Springer, 2008.

[55] M. Malone, D. M. Goeres, I. Gosbell, K. Vickery, S. Jensen, and P. Stoodley, "Approaches to biofilm-associated infections: the need for standardized and relevant biofilm methods for clinical applications," *Expert review of anti-infective therapy*, vol. 15, no. 2, pp. 147–156, 2017.

[56] R. Serra, R. Grande, L. Butrico, A. Rossi, U. F. Settimio, B. Caroleo, B. Amato, L. Gallelli, and S. de Franciscis, "Chronic wound infections: the role of pseudomonas aeruginosa and staphylococcus aureus," *Expert review of anti-infective therapy*, vol. 13, no. 5, pp. 605–613, 2015.

[57] J. C. Davies, "Pseudomonas aeruginosa in cystic fibrosis: pathogenesis and persistence," *Paediatric respiratory reviews*, vol. 3, no. 2, pp. 128–134, 2002.

[58] Å. Reimer, "Concentrations of the pseudomonas aeruginosa toxin pyocyanin in human ear secretions," *Acta Oto-laryngologica*, vol. 120, no. 543, pp. 86–88, 2000.

[59] L. G. Rahme, E. J. Stevens, S. F. Wolfort, J. Shao, R. G. Tompkins, and F. M. Ausubel, "Common virulence factors for bacterial pathogenicity in plants and animals," *Science*, vol. 268, no. 5219, pp. 1899–1902, 1995.

[60] A. Bhattacharjee, M. Khan, M. Kleiman, and A. I. Hochbaum, "Effects of growth surface topography on bacterial signaling in coculture biofilms," *ACS applied materials & interfaces*, vol. 9, no. 22, pp. 18531–18539, 2017.

[61] L. G. Monahan, L. Turnbull, S. R. Osvath, D. Birch, I. G. Charles, and C. B. Whitchurch, "Rapid conversion of pseudomonas aeruginosa to a spherical cell morphotype facilitates tolerance to carbapenems and penicillins but increases susceptibility to antimicrobial peptides," *Antimicrobial agents and chemotherapy*, vol. 58, no. 4, pp. 1956–1962, 2014.

[62] J. Parker Jr, D. Feldman, and M. Ashkin, "Raman scattering by silicon and germanium," *Physical Review*, vol. 155, no. 3, p. 712, 1967.

[63] D. WATSON, J. MacDERMOT, R. WILSON, P. J. COLE, and G. W. TAYLOR, "Purification and structural analysis of pyocyanin and 1-hydroxyphenazine," *The FEBS Journal*, vol. 159, no. 2, pp. 309–313, 1986.

[64] D. A. Armbruster and T. Pry, "Limit of blank, limit of detection and limit of quantitation," *The Clinical Biochemist Reviews*, vol. 29, no. Suppl 1, p. S49, 2008.

[65] C. Nguyen, W. Thrift, A. Bhattacharjee, K. Whiteson, A. Hochbaum, and R. Ragan, "Robust sers spectral analysis for quantitative detection of pycocyanin in biological fluids," in *Biosensing and Nanomedicine X*, vol. 10352, p. 1035205, International Society for Optics and Photonics, 2017.

[66] K. J. Reszka, Y. O'Malley, M. L. McCormick, G. M. Denning, and B. E. Britigan, "Oxidation of pyocyanin, a cytotoxic product from pseudomonas aeruginosa, by microperoxidase 11 and hydrogen peroxide," *Free Radical Biology and Medicine*, vol. 36, no. 11, pp. 1448–1459, 2004.

[67] L. E. Dietrich, A. Price-Whelan, A. Petersen, M. Whiteley, and D. K. Newman, "The phenazine pyocyanin is a terminal signalling factor in the quorum sensing network of pseudomonas aeruginosa," *Molecular microbiology*, vol. 61, no. 5, pp. 1308–1321, 2006.

[68] J. Snoek, H. Larochelle, and R. P. Adams, "Practical bayesian optimization of machine learning algorithms," in *Advances in neural information processing systems*, pp. 2951–2959, 2012.

[69] A. Tsantekidis, N. Passalis, A. Tefas, J. Kanniainen, M. Gabbouj, and A. Iosifidis, "Forecasting stock prices from the limit order book using convolutional neural networks," in *Business Informatics (CBI), 2017 IEEE 19th Conference on*, vol. 1, pp. 7–12, IEEE, 2017.

[70] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of eeg motor imagery signals," *Journal of neural engineering*, vol. 14, no. 1, p. 016003, 2016.

[71] J. Lee, J. Park, K. L. Kim, and J. Nam, "Sample-level deep convolutional neural networks for music auto-tagging using raw waveforms," *arXiv preprint arXiv:1703.01789*, 2017.

[72] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," *See https://arxiv. org/abs/1610.02391 v3*, vol. 7, no. 8, 2016.

Appendix A

Dose-Dependent Response of Pyocyanin Vibrational Bands

The log-log dose-response curves of commonly observed pyocyanin vibrational modes (1620 cm^{-1}, 1602 cm^{-1}, and 1353 cm^{-1}) were determined from the corresponding normalized SERRS intensity. The intensity versus concentration was also fit to Equation 4.3 in the main text. The linear regression results are plotted in Figure A.1.

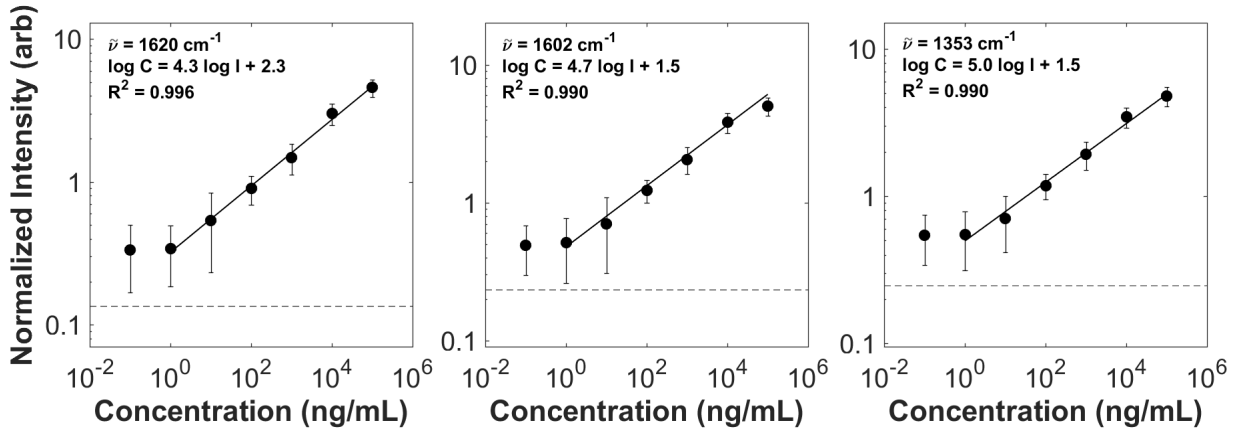

Figure A.1: Dose-response response of SERS normalized intensity at 1620 cm^{-1}, 1602 cm^{-1}, and 1353 cm^{-1}.

Appendix B

Desorption Half-Life of Pyocyanin

In order to measure the time for desorption of pyocyanin from sample surface, a SERS substrate was cleaned with IPA and DI water for 1 minute each. Droplet measurements were carried out to collect SERS signals from samples exposed to DI water and 1 $\mu g \cdot mL^{-1}$ (4.8 µM) pyocyanin. The substrate and immersion objective were subsequently rinsed with IPA and DI water, and methanol and DI water, respectively, after exposure to pyocyanin. The sample was then immersed in DI water and SERS spectra were acquired at 0, 10, and 20 min after rinsing to monitor the decay of pyocyanin signal. 100 spectra were collected at each step using the collection parameters described in the methods section. The resulting dose-response curves for major pyocyanin bands at 1620 cm^{-1}, 1602 cm^{-1}, 1353 cm^{-1}, and 552 cm^{-1} are illustrated below in Figure B.1.

After 20 minutes, two-sample t-tests with significance level $\alpha = 0.01$ show no statistically significant differences between the clean and used substrates at listed vibrational bands (p-value $> \alpha$). Thus this indicates that pyocyanin has appeared to desorp from the surface. To roughly estimate the half-life of this desorption, we fit an exponential decay function to the averaged intensities of each vibrational bands after rinsing (shown in last 3 data points

of each curve in Figure B.1).

$$y = ae^{rt} + b \tag{B.1}$$

We assume 2 fixed values: (1) the decay begins at the average intensity at $t = 0$ and (2) the decaying intensity will ultimately reach that of the substrate before exposure to pyocyanin, I_{clean}. Hence

$$a + b = I_{(t=0)} \tag{B.2}$$

$$b = I_{clean} \tag{B.3}$$

where $I_{(t=0)}$, and I_{clean} represent SERS signal initially after rinsing (third data point in Figure B.1), and the average intensity at the specified wavenumber on the clean surface (first data point in Figure B.1), respectively.

The resulting decay rates and their corresponding half-life are summarized in Figure B.1. The half-life of desorption is approximated as the mean of all calculated half-life to be 11 ± 2 min.

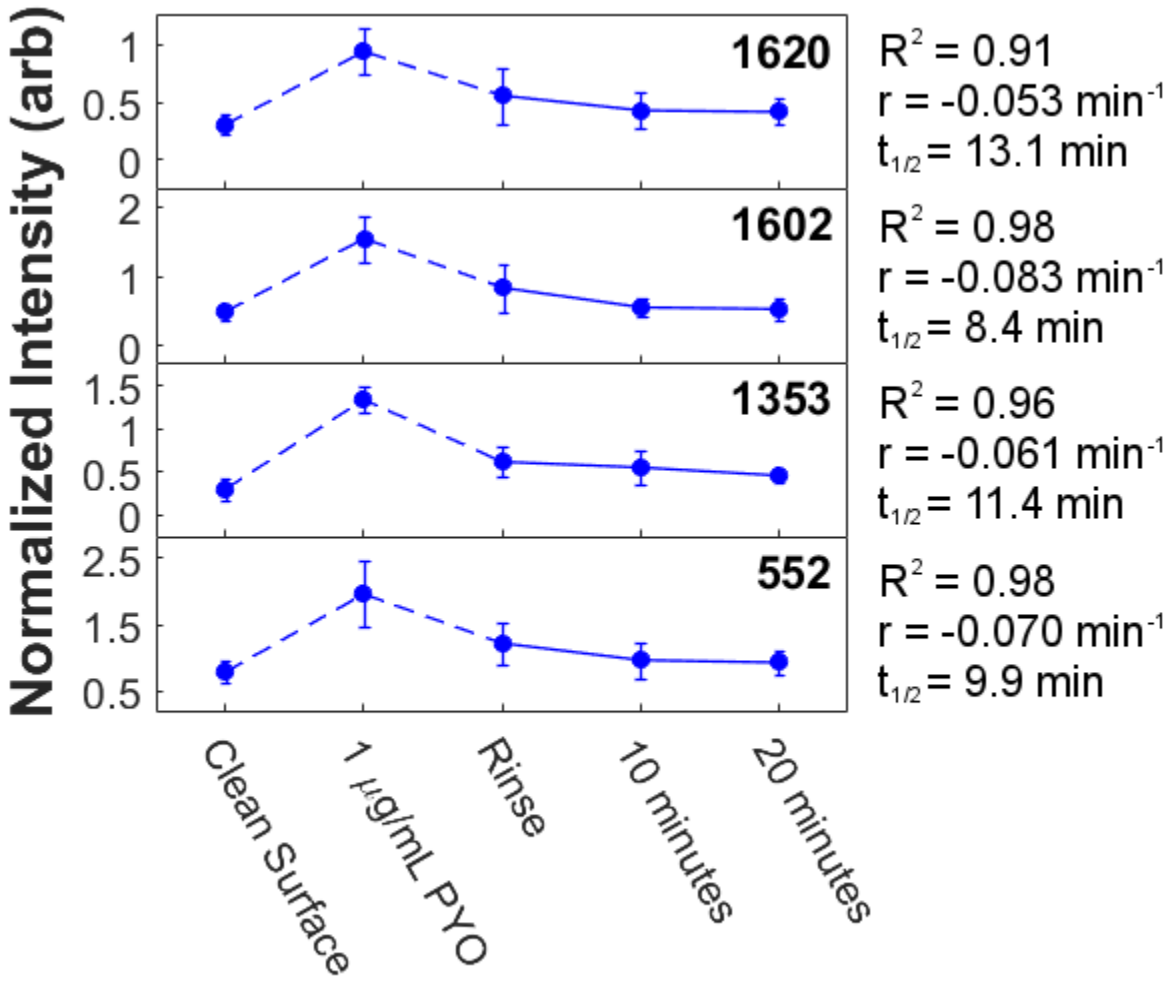

Figure B.1: Desorption kinetics of pyocyanin.

Appendix C

Visualizing CNNs Attention

With CNNs it is possible to provide a visual explanation of how the network makes its predictions *via* class activation mapping, or CAM. Figure C.1 shows SERS spectra in a spectra stack overlaid with their corresponding CAM heatmap. Note that all hot regions—indicating higher importance to the final prediction—coincide with pyocyanin peaks, most notably around the 1353 and 1620 cm^{-1} vibrational bands.

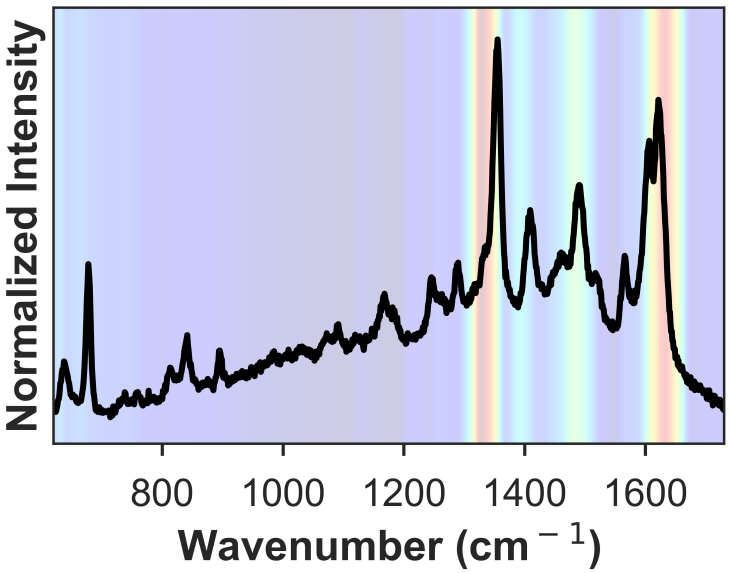

Figure C.1: CAM of 1D-CNN Model 3 ($R^2 = 0.998$) for SERS spectra of pyocyanin.