UNIVERSITY OF CALIFORNIA
RIVERSIDE

Advancing Surface Plasmon Resonance Biomarker Detection in Complex Matrices With Machine Learning and Novel Biomimetic Interfaces

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Chemistry

by

Alexander Scott Malinick

June 2023

Dissertation Committee:
Dr. Quan Cheng, Chairperson
Dr. Min Xue
Dr. Joseph Genereux

The Dissertation of Alexander Scott Malinick is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

The University of California, Riverside has in many ways been a second home to me for the last five years. I have so many fond memories that I owe to this amazing institution. At the nexus of all these positive relationships and experiences in my graduate career has been Professor Quan "Jason" Cheng.

Jason, your advice and support has been integral in shaping me into the scientist I am today. You have provided me the perfect amounts of independence, advice, and opportunities to achieve the skillsets necessary to obtain my goals. Learning how to conduct science by you has been one of the most rewarding experiences in my life. Looking back on this journey it is hard to believe that five years ago I had no idea what surface plasmon resonance was. What I did know then was that the research conducted in your lab was and is incredibly exciting, important, and ignited a passion for science within me. I cannot thank you enough for the wonderful journey you have facilitated.

I would also like to thank Jason's wonderful family, who have all been incredibly inviting to all of us in Cheng lab over the years, as well as to our friends and families. Some of the best home cooked meals, karaoke competitions, and beach parties, have been hosted by Jason's family. I want to emphasize two things to those reading this. First, I am now a fairly decent volleyball player, which was not the case when I began this journey for a Ph.D., nor a skillset I expected to gain along the way; and second, please do not ask me to sing for you, as I definitely did not improve in karaoke at all during my graduate studies.

To the faculty and staff at UCR, thank you for being incredibly supportive and helpful. Professor Joseph Genereux and Professor Min Xue, you both have provided me

immense amounts of advice, encouragement, and support throughout my graduate career. The success of my studies and of the projects discussed in the forthcoming Chapters would not be in the format they are without both of you. To UCR's Graduate Quantitative Methods Center thank you for helping me with the development, modification, and implementation of the code used in my studies, to UCR's Graduate Division thank you for awarding me the Dissertation Year Fellowship Award and for including me in the Graduate Jump Start Program, to The Department of Chemistry thank you for awarding me the 2022 Advancing Faculty Diversity Graduate Award and accepting me into UCR's graduate chemistry program, to PROTEOMASS Scientific Society thank you for awarding me the Excellent Communication Prize, and to the Graduate Student Association thank you for awarding me Travel Grant Awards in 2020, 2021, 2022, and 2023 to help me attend conferences all around the world.

To the current and past graduate student members of Cheng lab that I have had the honor to know and work with: Dr. Bryce Davis, Dr. Pete Shanta, Dr. Nor Akmaliza Rais, Dr. Kelvin Tran, Dr. Andrew Burris, Dr. Zhengdong Yang, Dr. Fatimah Abouhajar, Dr. Alexander Scott Lambert, Dr. Bochao Li, Santino Valiulis, Daniel Stuart, Cole Ebel, and Xin Wen, thank you. You have all been incredibly helpful in the projects I have worked on. I truly came into lab everyday excited to see all of you, you are all lifelong friends. Alex, thank you for training me how to use the various instruments and strategies of our lab and for bouncing numerous ideas off of no matter how outlandish. You have been a huge influence in my development as a person and as a scientist. Daniel, it is not a typo that you will appear on every one of my first author manuscripts written during my pursuit

for a Ph.D. You have been incredibly helpful in the design and execution of all of my research projects. Santino, your advice and guidance has been extremely influential in helping me grow into a more effective presenter and teacher. Cole, your input has always been much appreciated and you are a great teaching partner. I have shared so many laughs and happy memories with all of you.

To all the students I have mentored, Ellie Puente who is currently a Ph.D. candidate at the University of California, Los Angeles, Kanitin Khamnong who is a graduate student at the University of Massachusetts Amherst, Jasmine Gutierrez a recent UCR graduate preparing to apply to medical school, David Rizvi and Michlle Wu who are both still undergraduates at UCR, and Suzette Aguayo who is an undergraduate at Riverside Community College. I thank all of you for your time and trusting me as a mentor.

A huge thank you is in order for two other graduate students, Desiree Aispuro and Alexa Canchola, who while officially are not in Cheng lab, are honorary members. Little BA, Desiree, who this dissertation is also dedicated to, your love, patience, and advice have been integral to the success of my academic career. I really could not have completed this journey without all of your love and support. I love you with all that I am and am so very excited to begin our next adventure together! Alexa, I cannot express how thankful I am to have you in my life, from all your assistance and advice in editing and reviewing all of my assignments, manuscripts, and this Dissertation to being my best friend. From undergrad to graduate school you both have been my closest friends and most trusted colleagues.

Daniel Naimey, your advice, support, and willingness to share your experiences means so much to me. The trust you have in me as a resource regarding various scientific

topics is what made me realize I was becoming an expert in my field of study. I admire you, strive to be like you, and most importantly see you as a father. I love you dearly.

To my parents, who this Dissertation is also dedicated to, I cannot express how much I cherish both of you. Without either of you I would not be here today. In addition to providing me with the vital genetic material and nutrients needed for my existence, I can confidently say you both did an amazing job raising me. If it wasn't for your love, support, and patience I would not be the person I am and would not have achieved everything I have to date. Dad, you are one of the most entertaining and smartest people I know, you taught me confidence, determination, and integrity. Mom, your unending love and ability to see good in any situation and anyone is truly inspiring, you taught me patience, empathy, and charisma. I love you both greatly and know how much each of you have sacrificed for me, you both mean everything to me and are amazing parents. Thank you for everything.

To the rest of my family, I thank each and every one of you for your unending love and support, including the newest additions Rocky and Grace, I never thought I would have cats and dogs in the same home. In hindsight, I guess chasing all those butterflies and reading about science did pay off. I thank all of you for believing in me and providing me the support, resources, love, guidance, and joy that has been integral in this journey. I wish I could keep writing my acknowledgements to say thank you to each and every one of you, but if there was a way for a Dissertation to play music saying that you have gone over the allotted pages, it would probably be starting to play around now. I love all of you so very much.

And with that, as Jason likes to say, "Now that that's out of the way let's talk science"

COPYRIGHT ACKNOWLEDGEMENT

I dedicate this Dissertation equally to six of the most important beings in my life. Desiree Aispuro (soon to be Dr. Desiree Aispuro-Malinick, Dr. DAM) who is the love of my life, Littlest BA, and soulmate. To my Mom and Dad, Cathie and Dr. Charles Malinick, who without their unending love and support I would not be where I am today. To my Grandpa, Robert Foss, who helped raise me during my early childhood, he helped me catch all those butterflies. To Zorro and Nimbus Malinick who for the past 13 years have brought so much love and joy into my life; they also ate those butterflies.

I thank all of you for everything and I love each of you with all that I am.

ABSTRACT OF THE DISSERTATION

Advancing Surface Plasmon Resonance Biomarker Detection in Complex Matrices With Machine Learning and Novel Biomimetic Interfaces

by

Alexander Scott Malinick

Doctor of Philosophy, Graduate Program in Chemistry
University of California, Riverside, June 2023
Dr. Quan Cheng, Chairperson

Recent advancements in life science research have greatly improved our understanding of various intricate biological systems, some of which are affiliated with complex diseases. While a great deal of progress has been made towards the comprehension of relevant biophysical interactions, many remain poorly understood. This is especially true for protein interactions involving the cellular membrane. Of the analytical strategies developed for investigating molecular interactions, surface plasmon resonance (SPR) provides marked technical advantages and has become a cornerstone for these studies. However, the SPR method is still facing challenges, from both technical aspects and investigation of diseases, which are the focus of this Dissertation. SPR is susceptible to misidentification in biological matrices, due to cross-reactive and nonspecific interactions. The lack of reliable curved biomimetic membrane platforms has limited the investigation on various disease related protein interactions. The presented Dissertation aims to provide solutions to these challenges through developing novel biomimetic

membrane platforms, robust post data acquisitions analysis strategies, and antifouling protocols.

Chapter 2 showcases the development of a self-assembled pseudo-myelin sheath microarray. The platform and developed antifouling protocol were shown to be capable of detecting three multiple sclerosis (MS) specific anti-ganglioside antibodies in 10 % serum. Chapter 3 expands and improves upon the platform presented in Chapter 2. The MS specific antibodies were detected at disease relevant concentrations, 3 ng/mL to 25 ng/mL, in undiluted human serum. Machine learning algorithms were applied to facilitate the differentiation and identification of the highly cross reactive analyte-antigen interactions in a clinical setting. In Chapter 4, the characterization and development of a tunable curved membrane mimicking platform is presented. Through the use of statistical analysis, simulations, and SPR, we have demonstrated that the curved membrane platform could be uniquely applied to quantify protein-membrane interactions that require curvature. Chapter 5 reports the use of the curved membrane platform for the investigation of bridging-integrator-1's biophysical interactions, and their detection in biological fluids such as urine, which has been affiliated with muscular dystrophy. This study is the first quantitative evaluation of these interactions, and could aid in the diagnosis of muscular dystrophy. The works presented in this Dissertation have laid a solid foundation for advanced SPR biosensing that focuses on biomarker detection and disease diagnosis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1: Introduction and Background

1.1 Introduction:

In this Chapter, a discussion of the relevant topics of this Dissertation will be presented. Principal areas of focus include: current disease detection methods and their limitations, technical aspects of biosensors, fundamentals and applications of surface plasmon resonance (SPR) techniques, the challenges SPR-based methods face regarding biosensing and bioanalysis, as well as an overview of the statistical and machine learning algorithms that are essential to resolving these challenges. The combination of these approaches enabled reliable detection of disease biomarkers in complex biological matrices and allowed for characterization of poorly understood membrane-protein biophysical interactions that were not previously possible due to lack of proper techniques.

1.2 Current Disease Detection Processes and Their Limitations:

With the vast amount of diseases and disorders humanity has faced throughout history ranging from the recent COVID-19 pandemic to cancers and neurological diseases, such as multiple sclerosis, the development of fast, reliable, and direct detection methods that enable diagnosis has become a major area of interest.[1-4] Recent advancements in disease detection have drastically improved the quality of life for patients, as these methods facilitate the early implementation of life saving drugs and therapies.[5-7] Generally the first stage in disease detection begins with a patient sharing their symptoms with a health care professional. After an evaluation of the patients medical history, they will be prescribed to undergo a series of screening tests.[2] The most common screening tests include cholesterol measurements,[8] fecal occult blood tests,[9] pap smears,[10] detection of prostate specific

antigens in blood,[11] mammography,[12] colonoscopy,[13] and checking blood sugar levels.[14] The majority of these screening tests rely on monitoring changes to and/or the detection of specific biomarkers in blood.[15] The reliance on blood for evaluating a patient's health can largely be attributed to it being easy to collect, especially compared to other biological samples, and relatively large quantities can be collected.[16, 17] The ease and quantity of blood sample collection facilitates the use of many highly informative analytical investigations.[18] However, when a patient has symptoms that are not easily affiliated with a disease or detectable with the listed screening test, more robust analytical techniques need to be implemented.[19-21]

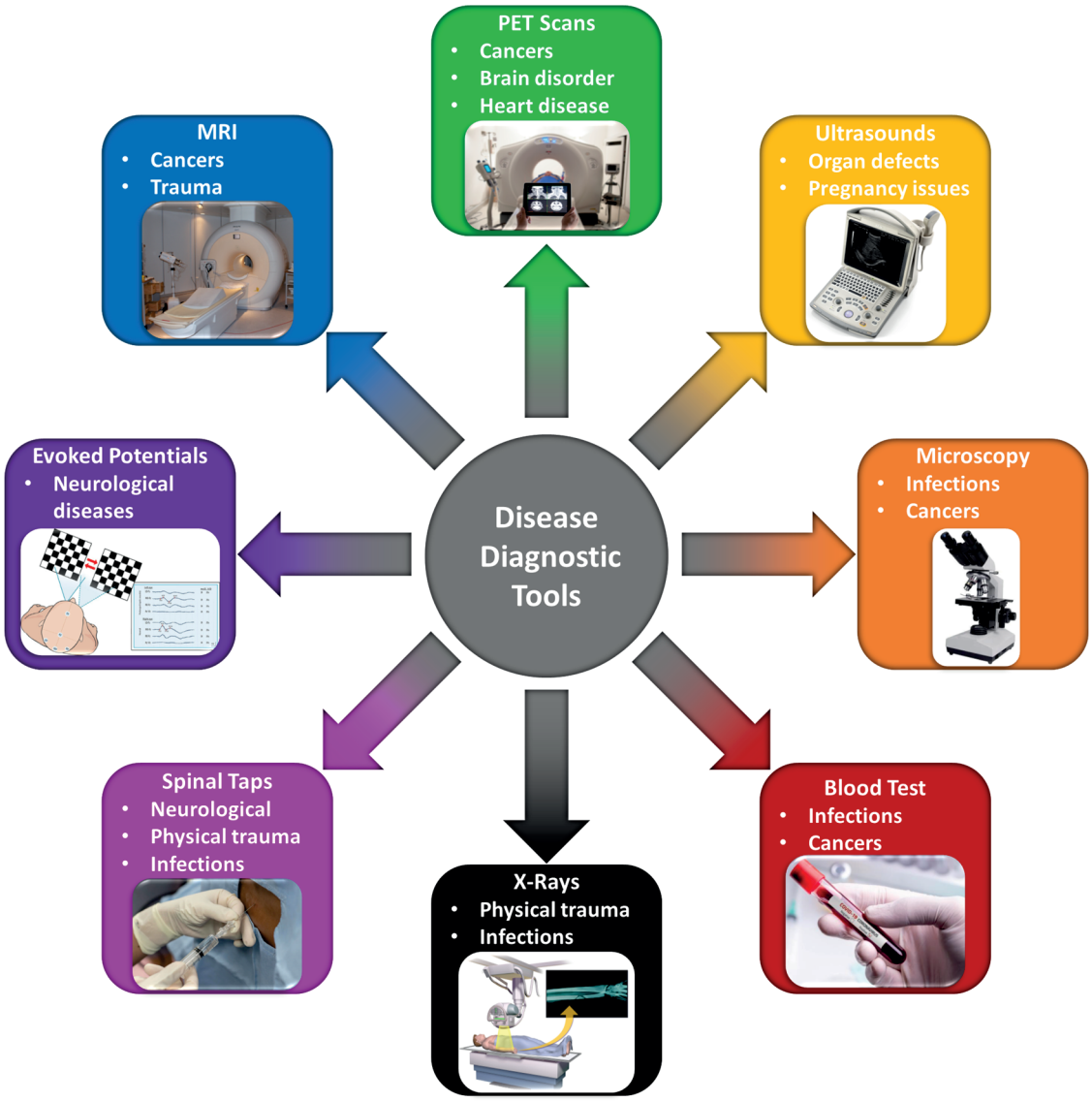


Figure 1.1. Overview of medical techniques and devices used for the detection and monitoring of diseases.

Investigating difficult to detect diseases that are normally not monitored for in screening test, tends to be a complicated and stressful process for both patients and health care professionals.[22] In most cases, an accurate diagnosis is only achieved after the combination of various expensive, invasive, and most advanced medical equipment and strategies available, some of which are shown in Figure 1.1.[23, 24] For the patient this process

is incredibly stressful, as an accurate disease diagnosis can range anywhere between a few minutes to several years depending on the disease.[25, 26] Some of the most well-known modern medical instrumentation and strategies to detect and monitor difficult to identify diseases include magnetic resonance imaging,[27] computerized tomography scans,[28] ultrasounds,[29] evoked potentials,[30] lumbar punctures,[31] and biopsies.[32] While these methods have allowed for the visualization and detection of substances that routine blood tests alone cannot detect, the majority of them are either highly invasive or indirect. For example, invasive methods like lumbar punctures, have a high likelihood of causing a great deal of pain or even severe complications, such as paralysis, for the patient.[33] Indirect methods, such as MRIs, only detect whether or not damage has occurred to the investigated area, but give no insights into what caused the observed damage.[34, 35] However, when these methods are combined with one another and blood tests, they can drastically increase the reliability of health care professionals ability to provide more accurate diagnosis. Nevertheless, there are still diseases where combining these techniques is still not enough.

Diseases that still prove to be a major challenge to detect even with the most modern medical equipment and strategies available, are usually related to the autoimmune system, such as neuromyelitis optica, multiple sclerosis, lupus, fibromyalgia, Huntington's disease, Parkinson's, and Alzheimer's.[35-42] While many of the advanced imaging methods used to detect and monitor these diseases are incorporating computing powers to drastically improve their reliability,[43] there remains a great deal of interest and need for the development of new ways to achieve early detection with disease specific biomarkers.[5, 17, 44] The need for this largely stems from the fact that for the majority of autoimmune diseases

there are no cures, and only therapies implemented during the earliest stages of the disease can improve or prevent further deterioration of the patient's quality of life.[45] This has led to a clear need for the development of new, early, and reliable disease detection methods, especially methods that detect and monitor diseases via blood, as they can easily be incorporated into the already established screening test.[36, 46-48]

To date, a great deal of effort has been invested into the development of new methodologies to detect and monitor multiple biomarkers associated with a diverse range of diseases simultaneously in blood.[16, 36, 48-52] This search for reliable and high throughput methods to detect a diverse range of disease biomarkers will be discussed extensively in Chapters 2 and 3 with an emphasis on multiple sclerosis (MS), which was used as a model disease.

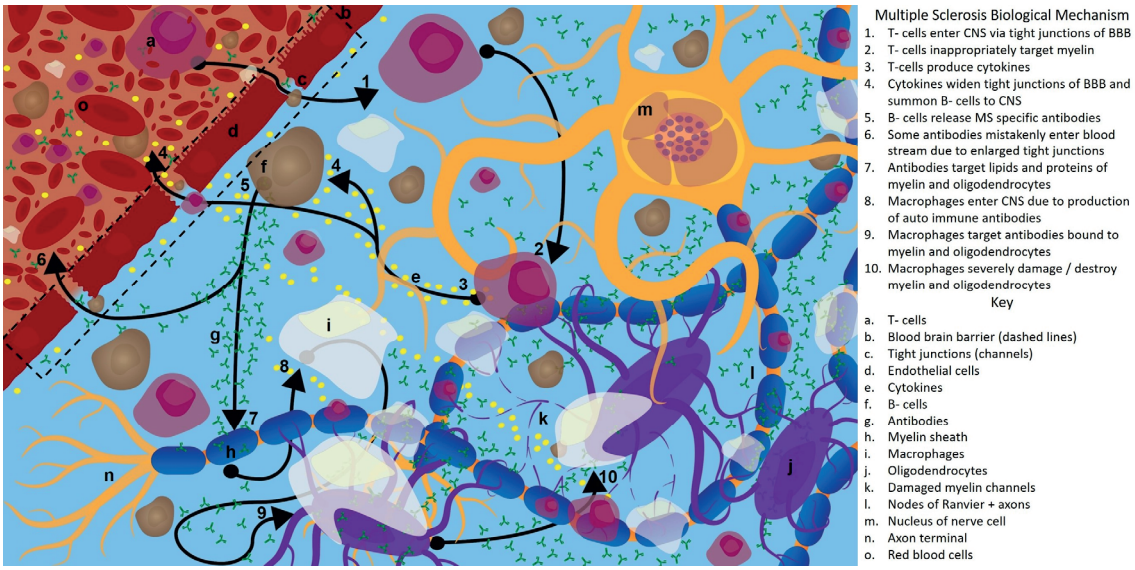


Figure 1.2. Graphic illustration of the complex biological process associated with multiple sclerosis and potential biomarkers that biosensors can use to aid in the early detection and monitoring of MS progression over time. Of note is that anti- myelin sheath antibodies, which are produced in the central nervous system by B cells, are able to enter the blood stream due to the tight junction of the blood brain barrier being enlarged due to several biochemical process affiliated with MS. The expanded tight junctions allow for MS specific biomarkers, like anti- ganglioside antibodies, to enter into the blood stream.

The major challenges associated with the detection of MS, as well as for other complex diseases, can be linked to the biomarkers themselves or lack of them. Specifically, most biomarkers associated with neurological diseases tend to have high cross reactivity with one another and other disease biomarkers, as well as being present at low concentrations in blood.[53-55] The majority of biological compounds associated with the demyelination process in MS, shown in Figure 1.2, are considered strong potential biomarkers to aid in the detection and monitoring the progression of MS.[56-58] Specific biological compounds of interest to act as biomarkers for MS include cytokines, like interleukin 1 and 2, varying types of anti-myelin sheath antibodies, such as anti-ganglioside antibodies, anti-aquaporin antibodies, and anti-myelin associated glycoprotein antibodies, as well as proteins that are not antibodies, such as chitinase-3-like-1 and

neurofilaments.[58-62] With the diverse range of potential biomarkers for MS, many of which overlap with other diseases, like neuromyelitis optica, a clear need for the development of new detection methods that can monitor many biomarkers simultaneously has arose.[63] To resolve this problem many scientific researchers and health care professionals have turned towards the development and application of biosensors for the detection and monitoring the progression of diseases.

1.3 Principles of Biosensors:

Biosensors have been used for a diverse range of applications such as early detection and monitoring of diseases,[19] dangerous compounds that threat national security/defense,[64] and ecosystems before, during, and after natural and/or human caused disasters.[65, 66] While the applications and types of biosensors are highly diverse and complex, the fundamental components by which they operate can be categorized into three specific parts, (1) a recognition element for the substance of interest, (2) a mechanism that can communicate the detection of that substance to, (3) a measurable redout that can be understood by the user. A cartoon representation of these components is illustrated in Figure 1.3.

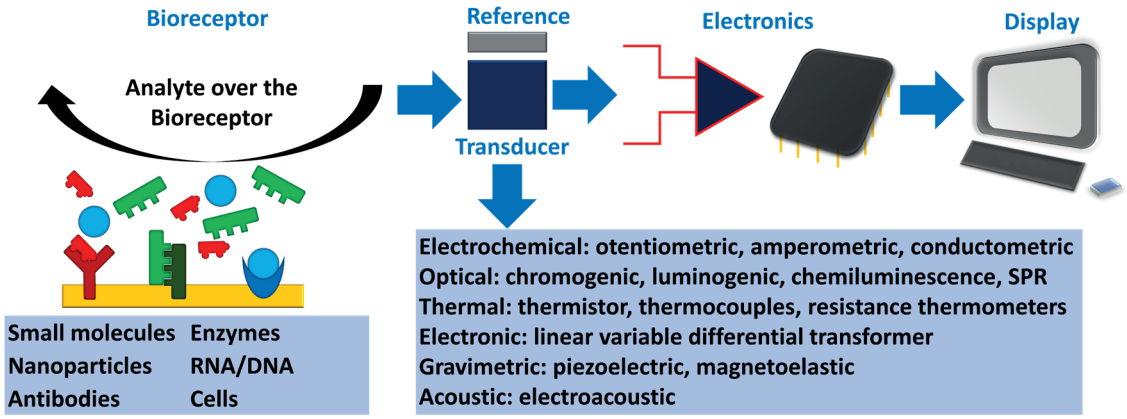


Figure 1.3. Summary of biosensors and their major components.

The recognition element, or antigen, predominantly relies on biological molecules that have high affinity for the analyte of interest, which is the substance being detected. Some of the most common recognition elements are antibodies, aptamers, DNA, RNA, enzymes, small molecules, and cells.[67] In biomedical applications analytes are commonly referred to as biomarkers. Detecting the interactions between an analyte and the recognition element is achieved through the second component, the transduction element.[68]

While there are many types of transduction elements, some of the most common types include fluorescent or luminescent emissions due to the release of energy as a result of the analyte and antigen binding,[69, 70] electrochemical methods that monitor perturbations to the movement of electrons thus indicating a chemical change,[71, 72] as well as methods that monitor changes in mass,[73] temperature,[74] and even acoustics.[75]

The final component of a biosensor is what allows for the detected interactions to be displayed in a readable format. How this information is presented can range drastically from very simple to highly complex outputs. Examples of simple presentations, include take home COVID-19 antigen tests and pregnancy test, both of which rely on lateral flow

assays, and operate by detecting whether or not a specific analyte is present in the sample.[76] For both pregnancy and COVID-19 antigen tests, the interaction of the analyte and antigen, either by a specific enzyme for pregnancy tests or functionalized nanoparticles for COVID-19 tests, causes the test channel to become visible to the human eye, thus allowing the user to tell if they tested positive or negative.[76]

Examples of complex readouts, which are normally processed by software on a computer and displayed on monitors, include sensorgrams and spectrograms.[77] These types of readouts are most commonly associated with biosensors that investigate complex biophysical interactions,[78] distribution of specific proteins, lipids, and metabolites in a sample,[79, 80] or the detection of specific compounds on and/or the physical properties of a surface.[81] Due to the complexity of these readouts deciphering their meaning normally requires the use of highly skilled laborers, complex post data acquisition statistical analysis tools, and in recent years, the increasing application of machine learning algorithms. As technology has advanced and the reliance on statistics and machine learning has grown, complex readouts have become more accessible and easier to read.[82]

Depending on the design of the biosensor the information provided can be either quantitative or qualitative. Quantitative biosensors such as glucose sensors, which normally rely on electrochemical assays, operate by detecting and accurately measuring the amount of glucose in a sample.[83] Qualitative biosensors, like the previously discussed examples for pregnancy and COVID-19 antigen tests,[76] usually rely on determining whether or not the analyte of interest is present at or above a specific concentration, but normally give no further information besides a binary positive or negative result.[76] In most

cases, quantitative biosensors are affiliated with complex readouts and qualitative biosensors are affiliated with simple readouts.

Due to the massive diversity of what biosensors can detect and monitor, through a wide range of complex modalities, they have been applied for the investigation and study of many diverse biological systems to aid in disease diagnosis, drug development, and the investigation of fundamental biophysical interactions.[16, 19] Many diseases, such as autoimmune,[18, 44, 84] cardiovascular,[51] and cancers,[85] can be linked to or associated with specific proteins. The link between proteins and diseases can largely be attributed to the fact that proteins tend to act as biological centers, thus when proteins go awry negative cascade events tend to occur. These cascade events normally cause a decline in health which results in the onset of disease specific symptoms.[19] Biosensors that detect and monitor diseases normally focus on quantifying the levels of a specific or several biomarkers, such as proteins like antibodies, in a specific matrix such as blood, which can then be linked to specific disease related symptoms.[16] How these biosensors detect and monitor these complex interactions can be categorized into two categories, labeled and unlabeled strategies.[86]

Labeled methods, such as enzyme-linked immunosorbent assay (ELISA), rely on non-native components attached to the recognition elements, which produces a detectable signal when the analyte of interest binds to a specific antigen.[87] While several subtypes of ELISA exist, one of the most common types is sandwich based ELISA, as it allows for high sensitivity and specificity through the use of antibodies.[88] The main process of this method is that a well is coated with a capture antibody, then non-specific binding cites are

blocked and a specific antigen is introduced. The antigen facilitates the detection of the analyte by capturing it, which allows for an enzyme-conjugated reagent to interact with the analyte, thus indirectly detecting the analyte antigen interaction.[87, 88] While labeled methods offer high sensitivity and good selectivity, the indirect approach can lead to loss of important information, especially in regards to biophysical studies, and can be challenging for high throughput screening analysis.[89]

Label-free techniques, or direct detection methods, rely on the principle of detecting the presence of an analyte directly with no secondary component.[86] Some of the most commonly used direct detection methods rely on monitoring electric or magnetic fields, which are highly sensitive to changes in mass or tension on a functionalized substrate.[86] Specific methods that rely on detecting these changes include: field-effect transistors (FETs),[90] magnetoelastic sensors, like quartz crystal microbalance with dissipation (QCM-D),[91] and optical biosensors like surface plasmons resonance (SPR).[49, 92] It is important to note that for all of these methods functionalized surfaces are heavily relied upon to monitor analyte-antigen interactions.

Once a substrate is known to reliably detect a specific analyte-antigen interaction, through surface functionalization techniques, the major benefits of label free methods become apparent. The most notable advantage is the ability to observe biological interactions in real time and to observe them directly without any modification.[86] The lack of a tag, which labeled methods rely on, removes the concern of having artifacts present that can disrupt, block, or alter the observed biophysical interactions, which can lead to false conclusions.[49] This is particularly important when investigating the development of

drugs, biomarkers, and understanding of disease mechanisms, which mainly rely on protein interactions with other biological compounds.[7, 49, 92] Because protein binding domains are notoriously sensitive to specific regions of the compounds they target, any modification may drastically reduce the binding affinity between a protein and its target.[93] This is significant as the vast majority of known biological interactions rely on protein binding domains to some extent.[94]

As the need to detect and understand diseases down to protein-protein and protein-membrane biophysical interactions has grown, so has the need for biosensors that allow for these investigations in highly controllable and reliable biomimetic environments.[19, 49] The need for this has largely stemmed from the desire to develop new and more effective disease detection and treatment strategies.[86,72] One of the most reliable label free investigation strategies for these complex biological interactions is Surface plasmon resonance (SPR) techniques.[49]

1.4 Surface Plasmon Resonance:

Surface plasmon resonance (SPR) is a label-free, real time, highly sensitive, reliable, and direct detection method. SPR has been applied for the investigation, detection, and monitoring of a diverse range of compounds and chemical interactions, but are most often employed for the study of biological interactions,[49] including disease biomarkers,[18, 84] drugs,[95] and complex biophysical interactions.[96] Normally the major components of SPR based instrumentation include a light source, an optical coupler, a goniometer, a sensor chip, a flow cell, a detector, and a computer with software that can process and display what the detector observes.[92]

1.4.1 Introduction to the Theoretical Basis of SPR:

The phenomenon known as SPR occurs when electrons on a thin, usually in the low nanometer (nm) thickness, of a conductive surface, such as gold (Au), silver (Ag), or aluminum (Al), become excited by light at a specific angle of incidence.[97] This nanoscale optical environment is highly sensitive to changes in refractive index (RI), and if observed over time allows for the measurement of binding kinetics in real time.[95] The limits to what SPR can detect, in simplified terms, can be summarized by two key properties. First, the mass captured onto the substrate should be no less than 0.01 pg/mm^2, as after this the implementation of amplification methods are required.[98] Second, the analyte's distance from the substrate should not exceed 300 nm, as beyond this point there will be an exponential loss in sensitivity.[99] While the concept by which SPR operates appears simple, more mass gives larger signal as long as the targets stay within the detection range, the physics behind this optical phenomena and the optics needed for it to occur are delicate and highly complex. Understanding the fundamentals of SPR is vital for the forthcoming Chapters, as each revolve around the advancement of SPR-methods for the detection, investigation, and understanding of biological interactions in the context of disease.

1.4.2 Total Internal Reflection and Evanescent Fields:

The most common process for generating plasmonic absorptions is through the use of total internal reflection (TIR) configurations.[99] Plasmonic absorptions can also occur with perpendicular light sources that rely on transmission-based configurations,[95] but will not be the focus of this Dissertation.

TIR configurations take advantage of refractions, which is the process by which light enters a material that will cause it to travel at a different speed. For this to occur there must be at least two materials with different RI,[100] and will be referred to as n_1 and n_2 for a binary system. n_2 will be assumed to represent the RI of a material with a lower RI, that is on top of a material with a higher RI, represented as n_1. As incident light travels from n_1 to n_2 the light bends away from the expected path it normally would travel due to a change in RI, as shown in Figure 1.4. This optical phenomenon occurs because the angle of refraction, θ_2, is greater than the angle of incident, θ_1. As the angle of incident increases more light is reflected and refracted, until a specific angle, known as the critical angle, θ_c, is reached. Once θ_c is achieved, light will travel along the boundary between n_1 and n_2 so that θ_2 is 90°. When the angle of incident is larger than the critical angle, $\theta_1 > \theta_2$, Snell's law occurs, as shown in Equation 1.1.[101] A visual representation of the different pathways that light travels under different conditions are shown in Figure 1.4.

$$\theta_c = sin^{-1}\left(\frac{n_2}{n_1}\right) \qquad \text{Equation 1.1}$$

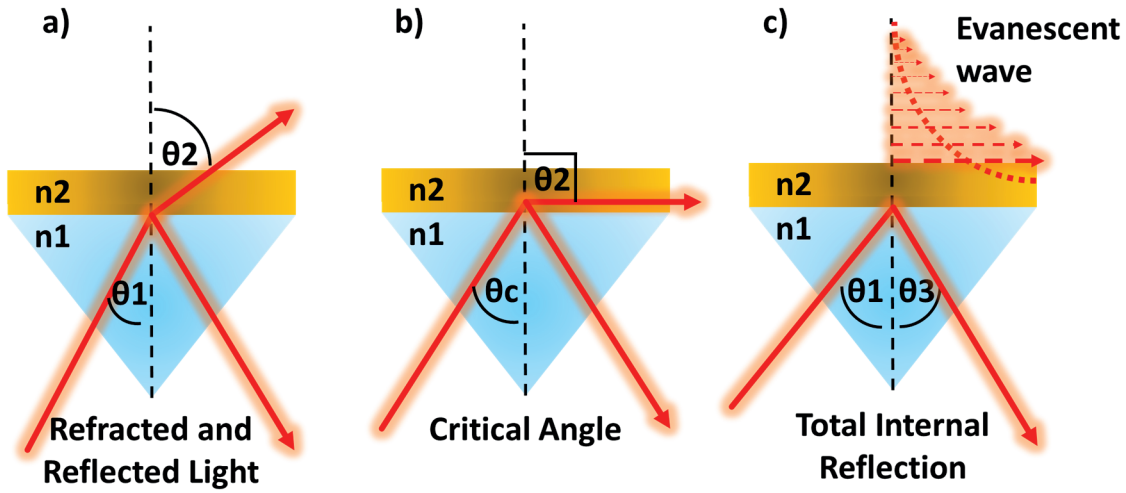


Figure 1.4. a) Refracted and reflected light ($\theta_1 < \theta_c$), b) critical angle ($\theta_1 = \theta_c$), and c) total internal reflection (TIR) ($\theta_1 > \theta_2$)

When using this configuration for analytical sensing purposes, n_1 is normally an optical component made of a material with an extremely high RI, like a quartz prism (1.458 RI). The TIR effect forms the basis of many advanced optical technologies that we rely on daily, such as cameras, telescopes, and optical fibers. For SPR based biosensors the most notable application of the TIR effect is for the creation of evanescent fields, or surface plasmons waves. Evanescent fields are composed of the localized oscillation of electrons on the top most layer of a conductive substance, n_2, and are only present within roughly 300 nm of n_2.[95] It is important to note that the amplitude of the evanescent wave decays exponentially as it travels away from n_2 and that the wave vector of the field, k_{ev}, travels parallel to the interface between n_1 and n_2[95], as shown in Equation 1.2:

$$k_{ev} = \left(\frac{\omega}{c}\right)\sqrt{\varepsilon_\rho}\sin\theta \qquad \text{Equation 1.2}$$

where c is the speed of light in a vacuum and ω is the angular frequency of the incident light. Because $2\pi c = \lambda\omega$, and the RI is the square root of the permittivity dielectric constant

ε, Equation 1.2 can be rewritten as Equation 1.3, which is more functional for the user and easier to read.[95]

$$k_{ev} = \left(\frac{2\pi}{\lambda}\right) n \sin\theta \qquad \text{Equation 1.3}$$

While most analytical methods that rely on measuring evanescent waves only account for the energy traveling in the upwards Z direction, these waves also penetrate the conductive material they are generated from. This penetration depth is defined by the wavelength of light being reflected at the n_1 and n_2 barrier which extends a distance roughly that of $\lambda/2$ nm.[100] The energy that is monitored for in the upwards Z direction, is dependent on θ_c, which is partly defined by n_2. Thus, any chemical and/or physical change to n_2, directly or indirectly, will be detectable through changes in the coupling conditions of θ_c.[95] Observing these changes is the corner stone of many types of biosensors that rely on electronically or optically active materials, and is fundamental to SPR biosensors.[99]

1.4.3 Surface Plasmons:

Surface plasmons are the collective oscillation of electrons in an excited state at the conduction band between a solid conductive material and a dielectric such as a liquid.[95] The simplest and most common way to generate surface plasmons is by achieving the TIR effect in conjunction with a thin smooth metallic surface, usually Au, on top of a prism.[97] Mathematically surface plasmons could have been predicted by the Maxwell–Heaviside equations, but were not theorized until 1957 by Dr. Rufus Ritchie,[102] and experimentally verified in 1959.[103] However the instrumentation necessary for surface plasmons to be used for practical applications, such as the investigation of biological interactions, were not readily available until the 1980s.[104] This achievement can largely be attributed to the work

conducted in Dr. Erich Kretschmann's[105] and Dr. Andreas Otto's[106] laboratories, both of whom have SPR configurations named after them.

The work done by these scientists proved that the free electrons of the topmost layer of a conductive material will resonate in an excited state when TIR occurs, due to the transfer of energy and momentum of photons.[102, 103, 105, 106] As the electrons in the conduction band absorb the energy from the photons a large disturbance in the outgoing lights intensity can be observed. This absorption of light is commonly referred to as a plasmonic dip, and occurs at angles higher than the θ_c.[99] Identifying when the plasmonic dip of a specific system will occur can be achieved mathematically through multilayer Fresnel equations.[97]

Fresnel equations rely on key information regarding a metal's behavior and the RI of the compounds that will be interacting with it, such as the prism, buffer solution, and biological material. While there are four classical Fresnel equations, which include light that is transmitted and reflected in both S- and P- polarized light, for TIR we are mainly concerned with those related to reflection. Equations 1.4 and 1.5 show the Fresnel equations for reflectance in S-polarized light, perpendicular to the plane of incidence, and P-polarized light, parallel to the plane of incidence, respectfully for a two layer system.

$$r_s = \frac{n_1 \cos\theta_i - n_2 \cos\theta_t}{n_1 \cos\theta_i + n_2 \cos\theta_t} \qquad \text{Equation 1.4}$$

$$r_p = \frac{n_2 \cos\theta_i - n_1 \cos\theta_t}{n_2 \cos\theta_i + n_1 \cos\theta_t} \qquad \text{Equation 1.5}$$

Here n_1 and n_2 are the RI values of two different components, as discussed in the previous section, θ_i is the angle of incidence, and θ_t is the angle of transmittance for both reflectance

in S-polarized (r_s) and in P-polarized (r_p) light. Most SPR studies conduct the entirety of the experiment in P-polarized light, as S-polarized light causes the maximum intensity possible to be absorbed. Monitoring the total change in RI, by switching from P- to S-polarized light is useful in several types of SPR-based techniques, such as SPR imaging, where detection is reported as a percentage of the total possible reflected intensity. However, to effectively use Fresnel equations one must know specific properties of the metal being used.

For Fresnel equations, a metal's n and k values are the most important features, which represent the real and imaginary portions respectively to that metal's RI. Identifying these values is typically achieved through the Lorentz-Drude model, as shown in Equation 1.6,[107] which explains the transport properties of electrons in a material.

$$\mathrm{R} = \frac{(n-1)^2 + k^2}{(n+1)^2 + k^2} \qquad \text{Equation 1.6}$$

Here R stands for the reflectivity coefficient of a thin metallic film independent of the incident angle. The values n and k are defined by the material's relative dielectric functions ϵ_r and ϵ_i, as well as its relative magnetic permeability, u_r, as shown in Equations 1.7 and 1.8 for the n and k values respectively.

$$n = \sqrt{\frac{u_r}{2}\left(\sqrt{(\epsilon_r^2 + \epsilon_i^2)} + \epsilon_r\right)} \qquad \text{Equation 1.7}$$

$$k = \sqrt{\frac{u_r}{2}\left(\sqrt{(\epsilon_r^2 + \epsilon_i^2)} - \epsilon_r\right)} \qquad \text{Equation 1.8}$$

While most SPR-based techniques use Au, other metals have been explored such as Ag, Al, copper (Cu), chromium (Cr), platinum (Pt), and titanium (Ti). The major property that SPR active materials share is that they are non-ferromagnetic as their magnetic permeability is nearly 1.[97] While many metals can produce surface plasmons, Au is the most commonly used.[95] There are three major reasons as to why most SPR biosensing studies use Au. First, Au's dielectric function is very well defined, due to the strong plasmonic properties of Au that are easily identified regardless of the quality of the surface or fabrication protocol used. Second, Au films are chemically stable across many experimental conditions. Third, Au is easy to functionalize through various strategies, such as with Au-thiol bonds.[100] However, in the context of surface plasmons, Au's dielectric function is the most significant property for its broad utilization.

Determining the dielectric functions of a metal is dependent on the frequency of the incident light ω, the metal's plasma frequency ω_p, and the metal's damping frequency Γ as shown in Equations 1.9, 1.10, and 1.11,[108] in context of Equations 1.7 and 1.8.

$$\epsilon_r = 1 - \frac{\omega_p^2}{\omega^2 + \Gamma^2} \qquad \text{Equation 1.9}$$

$$\epsilon_i = \frac{\omega_p^2 \Gamma}{\omega(\omega^2 + \Gamma^2)} \qquad \text{Equation 1.10}$$

$$\omega_p^2 = \frac{Ne^2}{m\epsilon_0} \qquad \text{Equation 1.11}$$

The plasma frequency, Equation 1.11, is widely considered the most significant property for determining a materials plasmonic capabilities.[108] Where N is the metal's free electron

density, e and m are the charge and mass of an electron respectively, and ϵ_0 is the permittivity of free space.

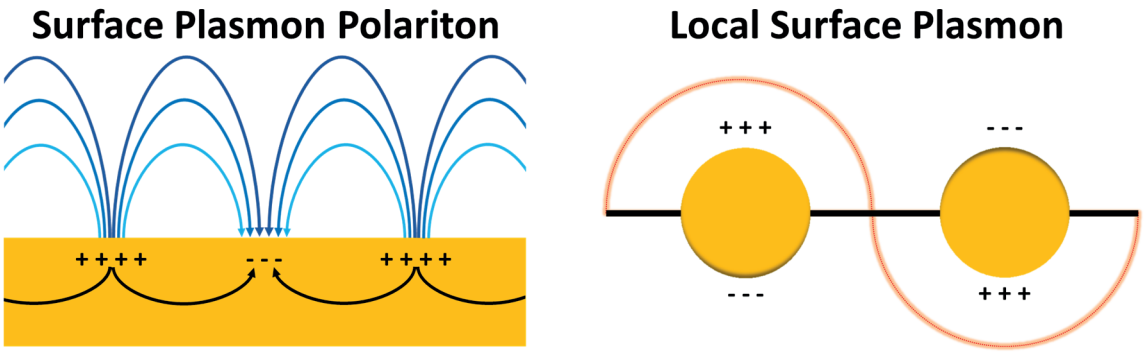


Figure 1.5. Cartoon illustrations for surface plasmon polariton and local surface plasmon.

The combination of surface plasmon excitations, evanescent fields, TIR, and a thin metallic films coupled together manifest into a propagating surface plasmon polariton (SPP),[95] a graphic illustration of which is shown in Figure 1.5. The wave vector of the SPP is shown in Equations 1.12 and an alternate version 1.13 that is rewritten to be more easily understood based upon the previously discussed equations.

$$k_{sp} = \frac{\omega}{c}\sqrt{\frac{\varepsilon_{mr}\varepsilon_s}{\varepsilon_{mr}+\varepsilon_s}} \qquad \text{Equation 1.12}$$

$$k_{sp} = \frac{2\pi}{\lambda}\sqrt{\frac{n_2^2 n_m^2}{n_2^2+n_m^2}} \qquad \text{Equation 1.13}$$

Where k_{sp} is the surface plasmon wave parallel to the metallic thin film with a RI of n_m.[95] Because surface plasmons are not inherently present on a surface several formats have been developed to generate them through the excitation of light. The two most common approaches are the Kretschmann and Otto configurations.[105, 106] The focus of the discussion

moving forward will be regarding to the Kretschmann configuration as it was used exclusively for data collection in the forthcoming Chapters.

For SPR to occur, the surface plasmons and photons must be coupled together. In the Kretschmann configuration, the prism acts as the coupler, as shown in Figure 1.6. The metallic surface, such as 50 nm of Au, is positioned directly onto the prism, thus facilitating the generation of surface plasmons at the metal-dielectric interface,[95] as shown in Figure 1.5 and 1.6. The wave vector coupling, more commonly referred to as SPR, takes place when the photon wave vector projected on the thin metal film, k_{ev}, equals the parallel vector component, k_{sp}, of the surface plasmon,[100] as shown in the process to achieve Equation 1.16.

$$k_{sp} = k_{ev} \qquad \text{Equation 1.14}$$

$$\frac{2\pi}{\lambda}\sqrt{\frac{n_2^2 n_m^2}{n_2^2 + n_m^2}} = \left(\frac{2\pi}{\lambda}\right) n_1 \sin\theta_{SPR} \qquad \text{Equation 1.15}$$

$$\theta_{SPR} = \sin^{-1}\left(\frac{1}{n_1}\sqrt{\frac{n_2^2 n_m^2}{n_2^2 + n_m^2}}\right) \qquad \text{Equation 1.16}$$

Now that θ_{SPR}, the angle required to achieve SPR, can be related to n_2 and n_m, which are known fixed values, the utilization of SPR as a biosensor can be pursued.

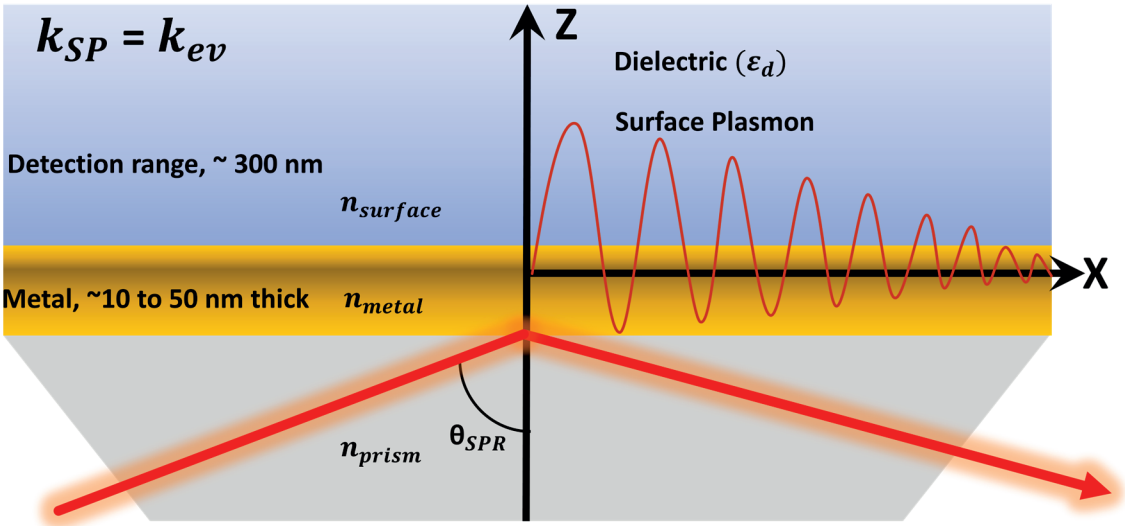


Figure 1.6. Schematic of SPR in the Kretschmann configuration with a dove prism showcasing the formation of the surface plasmon wave with a detection range of 300 nm propagating across the thin metallic surface.

While these equations computationally allow for the investigation of a materials plasmonic capabilities and even simulating biological interactions,[109] there is currently no way to reliably simulate them to date. However, as machine learning algorithms become more sophisticated and reliable it may be possible to simulate entire experiments before conducting them with biological materials. The benefits of this could save time in regards to troubleshooting and cost when using expensive material or low yield products.[110, 111] Until then, scientist using SPR-based biosensors to investigate various topics will need to continue relying on currently established instrumentation, such as SPR spectroscopy.

1.4.4 Surface Plasmon Resonance Spectroscopy:

Surface plasmon resonance (SPR) spectroscopy in the Kretschmann configuration was first proposed in 1968,[105] and is the basis of the vast majority of commercialized SPR-based biosensors. The majority of light sources used by SPR instruments rely on a red LED, with wavelengths ranging between 625 nm to 700 nm, which is optimal for Au based

SPR biochips.[112] The LED is then aimed at the desired angle, θ_{SPR}, which is identified either by angular or wavelength scanning of the Au SPR biochip.[97] Angular versions of SPR spectroscopy determine the incident angle either through physical rotation of the system or the system being held in a fixed position and a photodiode array generating angular displacement.[113] Wavelength scanning SPR spectroscopy relies on a white light source that targets a stationary system where spectroscopic gratings are present at the detector and separate the reflected intensities.[95] The SPR spectroscopy instrumentation used for data collection in this Dissertation relied on angle scanning with a red LED light source.

A cartoon illustration of a SPR spectroscopy readout during an experiment is shown in Figure 1.7. The observed binding signal depicted in the cartoon SPR sensorgram, is the result of tracking the detectable changes in the dielectric environment, which is monitored by the changes in RI of the conductive surface, $n_{surface}$. As this change occurs the position of the minimum reflected intensity, or θ_{SPR}, moves to a new angle. When this change is plotted against time the generation of a SPR sensorgram can be achieved. The SPR sensorgram allows for the label free observation of binding events between analytes and antigens.[95]

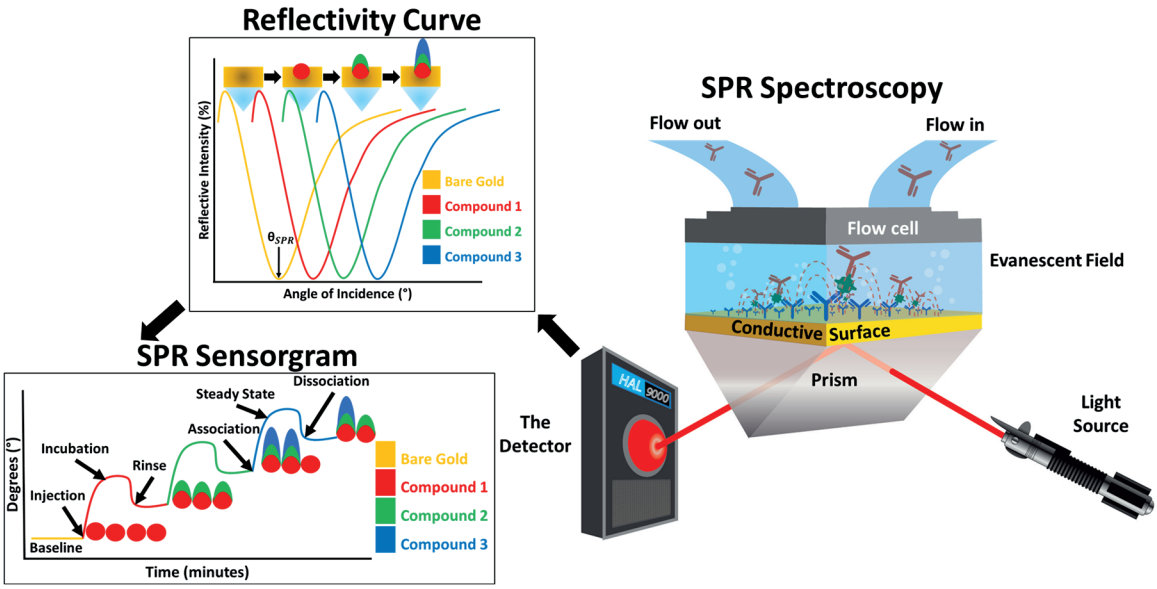


Figure 1.7. A traditional SPR spectroscopy configuration is shown with a light source, prism, thin conductive surface, a detector, and a flow cell with an inlet and outlet. A diagram showcasing the movement of the angle of incidence as compounds interact with the surface thus changing the RI. Finally a SPR sensorgram plotting the change of the SPR angle over time allowing for the observation of binding kinetic interactions in real time.

In most cases when multiple concentrations are compared, a positive linear correlation between RI and amount of material on the surface can be observed. However, some materials and unique biophysical interactions have been shown to cause negative changes.[114]. In addition to investigating various biophysical interactions, SPR spectroscopy has been implemented to determine the thickness of thin films and the mass present on a surface,[115] via Equation 1.17.

$$R = m(\eta_a - \eta_s)\left[1 - e^{-\frac{2d}{l_d}}\right] \qquad \text{Equation 1.17}$$

Where R is the reflected signal shift in terms of wavelength, angle, or intensity, m is the sensitivity based upon small RI increments, η_a is the RI of the adsorbed layer, η_s is the RI

of the overall solution, d is the thickness of the material, and l_d is the decay length of the evanescent field based upon the incident wavelength from the light source.[116]

In addition to SPR spectroscopy other SPR-based methods exists, such as fiber optic SPR, localized SPR (LSPR), and SPR imaging. For fiber optic SPR a waveguide is used to allow SPR to act as a probe which has many promising biomedical applications.[49] However, due to the potential of the light path within the cable to alter as the probe is moved, causing a drastic change in sensitivity, fiber optic SPR has not widely been adopted to date.[49] LSPR relies on the investigation of the evanescent field when coupled to a single plasmonically active particle, as shown in Figure 1.5. LSPR has been used extensively to investigate nanoparticles and nanostructures, and is the baseline for many amplification strategies for SPR-based strategies.[95] While many forms of SPR-based strategies exist, one of the most exciting and recent types is surface plasmon resonance imaging (SPRi). SPRi offers many advancements compared to other SPR-based platforms as it can easily incorporate microarrays, allowing for high-throughput and multiplex biomarker detection.[18, 84, 97]

1.4.5 Surface Plasmon Resonance Imaging:

Surface plasmon resonance imaging (SPRi) relies on the same underlying principles as SPR spectroscopy. While the majority of SPR spectroscopy instruments rely on angle scanning, most SPRi instruments operate under a fixed angle configuration. This configuration is ideal for SPRi as it facilitates the simultaneous monitoring of many changes in RI over an entire substrate.[117, 118] When SPRi is combined with microarrays the ability to conduct high throughput and multiplex detection is facilitated, as shown in Figure

1.8. This is achieved by the CCD camera capturing images throughout the entirety of the experiment that when coupled with specific software, allows for the investigation of changes in RI for each individual pixel of the captured image.[95, 96] In most cases specific regions, usually within a well of the microarray, are identified either by the user or an algorithm and are averaged together to give the reported changes in RI.[18, 84, 97] Due to SPRi's propensity for high throughput and multiplexed detection, instrument manufacturers such as Bruker and Horiba, have begun to offer SPRi instrumentation. This is significant as to date the vast majority of SPRi based technologies have been home built,[96] such as the one used in this Dissertation.

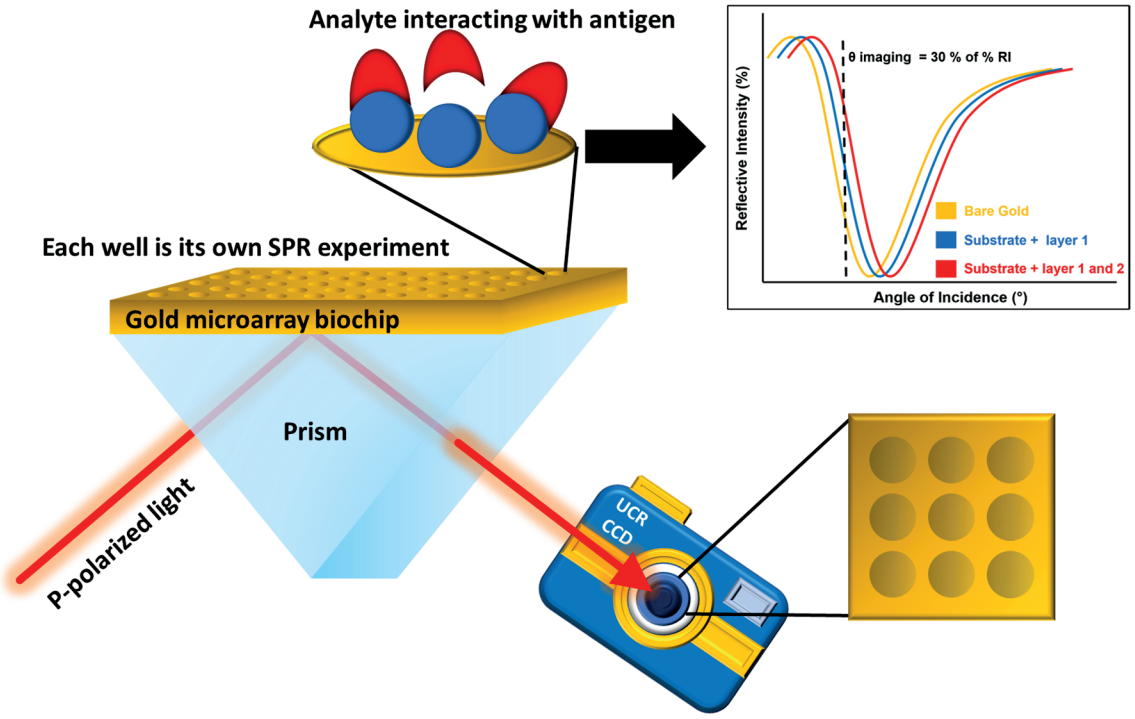


Figure 1.8. SPRi at a fixed angle detecting changes in reflected intensity based upon an analyte and antigen interacting through a UCR CCD camera. Wells on the surface allow for individual experiments to occur independent of one another and due to being inside the surface their SPPs are contained inside them allowing for higher intensity then compared to SPR spectroscopy.

Recently modified versions of SPRi, SPR microscopy (SPRM), have been used to investigate individual cells, in addition to the biomolecules SPRi is routinely used for.[119] The interests regarding SPRM is largely due to its capabilities for monitoring both the collective behavior of cells and individual cells responses to various stimuli.[119] While the utilization of SPRM may offer unique benefits for various microscopy studies, such as spatial biology, the most notable limitation is that the detection range must be within 300 nm from the SPR active substrate, and thus further development still needed to be done.[95, 96, 119]

SPRi also have notable challenges that need solutions. One issue is the potential for noise in individual wells on the microarrays to be amplified.[117] This concern arises as surface inconsistencies cannot be effectively averaged out within an individual well as easily as they can be over an entire thin film, as is the case with SPR spectroscopy.[18] It has been the focus of many to develop more reliable fabrication strategies for array based biochips.[18, 84, 97, 112, 120] In addition to improving manufacturing protocols, the variations observed in SPRi data can be negated by using large amounts of datasets and applying statistical analysis.[95, 96, 117] Control regions have also been utilized in SPRi experiments as a strategy to account for variations between wells and biochips. By combining control regions, large datasets, and statistical analysis tools, these issues are being addressed quickly and effectively. This is significant as there have been claims that SPRi's potential loss in sensitivity does not outweigh its array capabilities.[121] This Dissertation presents information that argues against these claims, and shows that with the correct investigation protocols a loss in sensitivity large enough to discourage the use of SPRi can be avoided. The significance of resolving these issues is key to implementing any SPR-based method for biosensing and bioanalysis studies.

1.5 Surface Plasmon Resonance Based Biosensing and Bioanalysis:

While there are many types of SPR-based strategies, each with pros and cons, they all have been used in various biosensing applications.[95] In this section, an overview of some of the most common approaches will be introduced with a focus on the utilization of biomimetic platforms such as self-assembled monolayers,[122] supported lipid bilayers,[123] and tethered supported lipid bilayers.[124]

1.5.1 Fundamentals of SPR-based Biosensing:

SPR-based techniques being label free and real time, have allowed for the detection, monitoring, and investigation of a diverse range of biological analytes.[118, 125-129] As briefly discussed earlier and shown in Figures 1.7, SPR sensorgrams can be segmented into three specific regions, association, steady state, and dissociation affinities.[98] The information gained from these regions and the total SPR sensorgram, allows for the analysis and calculation of affinity, K_D, defined and determined by the equations below.[98]

The first step in understanding binding kinetics is through observing the equilibrium constant and reaction rate interactions that can be assessed by the reaction:

$$A + B \leftrightarrow AB \qquad \text{Equation 1.18}$$

Where A is the analyte of interest and B is the antigen immobilized on the sensors surface. The rate at which complex AB forms is referred to as the association rate, k_a, and is determined by:

$$\frac{dAB}{dt} = AB_{max}[A]k_a \qquad \text{Equation 1.19}$$

Where AB_{max} is the max possible response from analyte binding, [A] is concentration of A, and t is time. With the knowledge of what k_a is, the identification of the dissociation constant, k_d, can be achieved:

$$AB_t = \left(AB_i - AB_f\right)e^{-k_d t} + AB_f \qquad \text{Equation 1.20}$$

Where AB_i is the beginning of dissociation and AB_f is the end of dissociation.

Once the formation of the AB complex reaches equilibrium, the equilibrium association and dissociation constants can be determined. To calculate the equilibrium

dissociation rate constant, K_D, we can use the above equations to calculate both k_a and k_d, then use that information in Equation 1.21 below, to solve for K_D.

$$K_D = \frac{k_d}{k_a} \qquad \text{Equation 1.21}$$

The reciprocal of which is the equilibrium association rate constant, K_A, calculated with the equation below.

$$K_A = \frac{k_a}{k_b} \qquad \text{Equation 1.22}$$

K_D can also be determined by the sensorgram itself with Equation 1.23.

$$AB_{eq} = AB_{max}\left(\frac{1}{\left(1 + \frac{K_D}{[A]}\right)}\right) \qquad \text{Equation 1.23}$$

Where AB_{eq} is the average response signal at the equilibrium for a specific [A]. While the discussed kinetic equations are not the only approaches to determine binding affinities and equilibrium constants, they are the most commonly used methods for SPR based detection methods.[98]

While the investigation of binding kinetics is currently the most common application of SPR-based strategies,[98] their use as biosensors is rapidly expanding. Recently SPR-based techniques have been demonstrated as effective ways to detect and differentiate gasses,[130] a diverse number of diseases,[49, 84, 96, 131] the detection of environmental pollutants,[132] and as a tool to investigate highly complex fundamental biophysical interactions.[78, 133, 134] For the majority of SPR-based investigations a great deal of effort has gone into how to functionalize the SPR active substrate, such as through the use of biomimetic cellular membranes.[71]

1.5.2 Biomimetic Cellular Membranes:

Cellular membranes form a fundamental component of cellular biology, as they protect and separate cells from the outside world.[135] The major components of cellular membranes are lipids, which include many diverse subcategories that play host to various types of proteins.[136] Similar to proteins, lipids participate in a wide variety of biological functions, including but not limited to ion-gating, cell to cell communication, adhesion, and immune responses.[135, 137] Due to lipids clear integral role to cellular membranes and many diverse biological functions, the development and application of lipid based membrane mimics has been the focus of many studies.[124, 138-142]

The creation of biomimetic lipid membranes is normally achieved by mixing desired amounts of membrane components together in a chloroform methanol mixture.[140] Though protein incorporation is possible,[143] it is very challenging and was not pursued in this Dissertation. After the lipid components have been mixed, dried under an inert gas, and resuspended in an aqueous solution the lipid mixture can be extruded into the desired vesicle sizes through thin polycarbonate films.[139, 144] The vesicles are then injected into the system where they can be made to unfurl, fuse, or rupture onto various surfaces through a plethora of chemical strategies.[145] Ensuring that the lipid vesicles rupture can be achieved by using Ca^{2+} to bridge the charge gap between the surface and the vesicle, polymer induced crowding effects, amphipathic peptides, and solvent exchange to name a few.[71, 92, 95, 141, 146] Regardless of the strategy used the final product should be a membrane mimicking platform that can be used to reliably investigate various biological interactions. The three most common of which are shown in Figure 1.9.

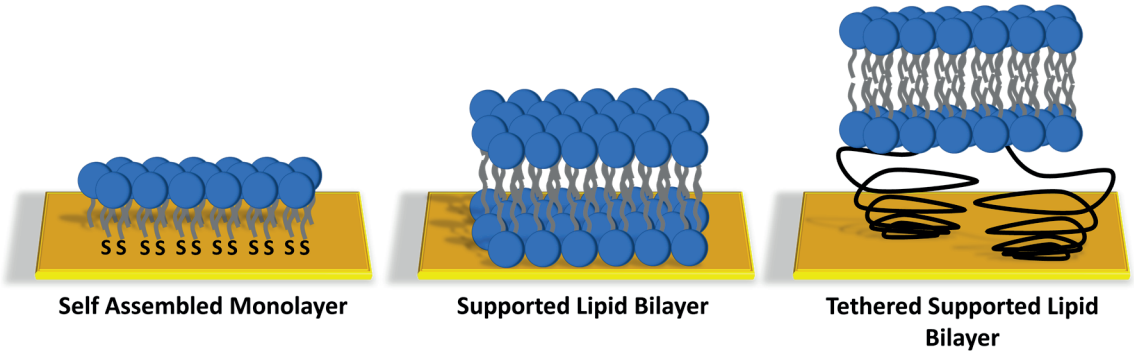


Figure 1.9. Illustrations of a self-assembled monolayer (SAM), supported lipid bilayer (SLB), and tethered SLB.

The most commonly used membrane mimics for the detection and investigation of analytes in SPR studies include self-assembled monolayers,[147] supported lipid bilayers,[148] and tethered supported lipid bilayers,[149] as shown in Figure 1.9. Recently a forth type of biomimetic membrane mimic, curved membranes, have started to be used for the investigation of biophysical interactions, though they are still arguably in their infancy.[146, 150-152] The advantage of curved membrane mimics over the three more traditional planar strategies, is that they offer curvature, thus are more accurate to the structure of naturally occurring cellular membranes.[153] This is significant as many protein-membrane interactions rely on curvature.[154] The majority of curved membrane mimics available to date have relied on solid support systems,[146] which is not ideal as it may alter the observed biophysical interactions. Chapters 4 will showcase the development of a strategy to generate curved membrane mimics without the use of solid support systems, which is then applied for SPR biosensing in Chapter 5.

One of the most frequently used base lipids to generate membrane mimics is palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC).[155] Other lipids will then be incorporated during the mixing phase at percentages normally ranging from 0.1 % to 5 %,

as higher percentages can lead to unstable membranes.[138] This process is highly flexible and allows for various types of membrane components to be included, giving one the ability to control charge, number of antigenic binding cites, and fluidity of the membrane mimicking platform.[156] To date the majority of investigations that rely on lipid membrane mimics have focused on their utilization for the detection of biomarkers in biological matrices, largely as a means to reduce nonspecific binding and/or to incorporate antigens.[139] However, as it has become more apparent that the fluidity and charge of a membranes plays a vital role in the observed biological interactions, these features have started to be routinely taken into account.[157]

Many strategies have been developed to increase or decrease the fluidity of membrane mimics, such as through the incorporation of cholesterol and other unique compounds.[142, 145, 158] Changing the charge of membranes can be accomplished through the inclusion of different lipids such as 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) ($POPG^-$), negative charge, and 1,2-dioleoyl-sn-glycero-3-ethylphosphocholine (EPC^+), positive charge. In addition to incorporating substances to control the charge and fluidity of membrane mimics, a great deal of work has focused on the incorporation of tissue-specific lipids, such as gangliosides, to make these models more biological relevant.[18, 141, 143] The combination of these components brings membrane mimics ever closer to becoming highly accurate representations of the membranes they attempt to mimic.[95] Further functionalization of membrane mimics can also be achieved by the incorporation of lipids with unique head groups, such as dioleoyl-sn-glycero-succinyl

(DGS-Ni-NTA),[159] which facilitates the inclusion of proteins into these studies while avoiding the challenges of incorporating them into the membrane mimics.[138, 143, 160]

Although membrane mimics are very simplified representations of the natural occurring cell membranes they attempt to mimic,[142] the controllability and simplification of these mimics facilitate the pin point investigation of specific biophysical interactions.[153] The ability to control what is in the membrane in controlled ratios is something that studies that rely on natural cellular membranes are incapable of achieving.[139] By investigating each component of a cell membrane individually and combined in controlled ratios, complex biophysical interactions with specific analytes can be understood in far greater detail. The investigation of these complex interactions in controlled environments has led to major advancements in drug development, drug delivery methods, and expanding our fundamental knowledge of biological systems.[128, 143, 158, 161, 162] However due to SPR-methods predominantly relying on Au films, a great deal of effort has been invested into the development of strategies to reliably functionalize membrane mimics onto Au surfaces.

1.5.3 Generation of Membranes Mimics on Gold Films:

While SPR-based strategies have relied heavily on the benefits of membrane mimics for biological investigations, Au which is the most common conductive material used in SPR studies, is not an optimal substrate for the direct immobilization of membrane mimics.[97] The inability to reliably generate lipid membrane mimics on top of bare Au surfaces is largely due to the slightly native hydrophobicity of Au surfaces, which repels the highly polar head groups of lipids.[118, 126, 163] However, multiple strategies to overcome

this challenge have been developed, and a brief overview of the two most common will be discussed.

First, Au surfaces can be modified with various functional groups that aid in the formation of membrane mimics, usually by making the surface highly hydrophobic so that the tails of the lipids self-insert and adhere to the surface through hydrophobic-hydrophobic interactions.[164] One of the most common strategies to achieve this is through the formation hydrophobic alkyl chains via Au thiolate bonds.[165] The presences of the alkyl chains promotes fusion with the ruptured lipid vesicles resulting in the creation of self-assembled monolayers (SAM). Additionally the vesicles themselves can be constructed to include PEG-ylated or thiolated lipids,[149] which allows for the lipids to tether themselves to the Au substrate directly, and can then be ruptured using the previously discussed chemical strategies to ensure the formation of a tethered supported lipid bilayer.

The second strategy to aid in the formation of lipids on Au surfaces is by depositing thin films, usually less than 5 nm, of various materials, such as silica (SiO_x), on top of the Au surface via dry or wet deposition techniques.[166] In the case of SiO_x, it acts as an adhesion material for the lipids and promotes the formation of supported lipid bilayers.[118] The silica layer is thin enough that it does not significantly affect the plasmonic absorption, and has even been reported to enhance its sensitivity, and can be functionalized with a diverse range of compounds such a perfluorinated carbon chains.[18]

By combining label free techniques, like SPR-based detection methods, and biomimetic membranes, the observation of unmodified biological interactions can occur in a highly controllable and reliable environment.[95] In addition to offering the ability to

investigate biophysical interactions, each one of the discussed membrane mimicking strategies have been found to provide anti-fouling properties to some extent.[18, 49, 84, 95, 138, 167] The utilization of membrane mimics in this fashion is of great significance as nonspecific and cross reactive interactions are the major challenges for SPR-based biosensors to date.[92]

1.6 Challenges for SPR Biosensing and Bioanalysis:

A great deal of the work that will be presented in this Dissertation has focused on the improvement of SPR-based detection methods for biomedical applications. Specifically, this was achieved through the development of protocols to account for and suppress nonspecific binding, as well as deciphering cross reactivity between multiple biomarkers in serum.

1.6.1 Nonspecific Binding:

Nonspecific binding (NSB) is when compounds interact with the surface and produce an unwanted detectable change that usually suppresses the detection of the analyte of interest.[49] NSB is a major challenge and concern in the development and optimization of biosensors.[168] As the level of NSB increases the reliability of the biosensor's sensitivity and selectivity becomes questionable, as the presence of the elevated background signal cannot be easily differentiated from the specific analyte-antigen interactions.[168] NSB is normally caused by substrate stickiness, electrostatic interactions, hydrophobic binding to the surface, the substrate not being fully functionalized, and/or not effectively blocking highly reactive regions of the surface.[92, 95, 96, 169, 170] Understanding NSB is integral for the clinical application of SPR, where biological samples, such as blood, are commonly used.[49]

Biological matrices contain large amounts of biological components in addition to the analyte of interest. These other compounds present are the cause of NSB, and have led to the implementation of a diverse range of antifouling strategies.[49]

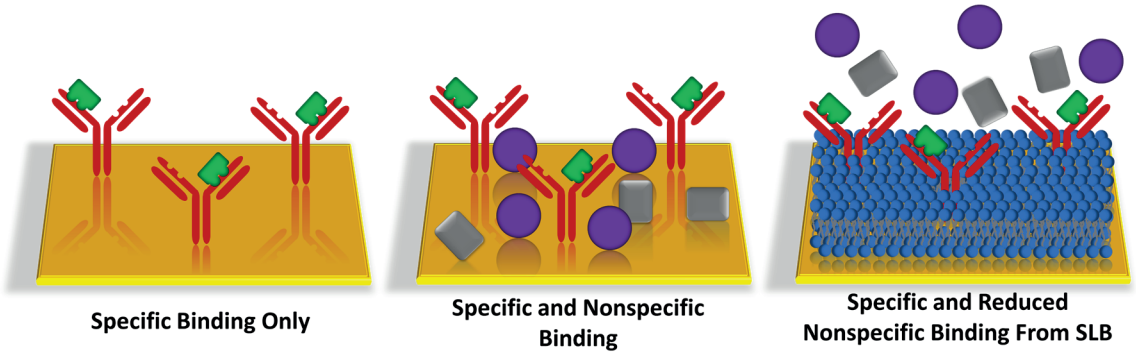


Figure 1.10. Cartoon representation illustrating the difference in specific interactions only, nonspecific and specific interactions without an antifouling supported lipid bilayer present, and the reduction of nonspecific binding due to the presence of an antifouling SLB.

Many strategies have been developed to reduce or even eliminate NSB under the right conditions, these methods predominately rely on either physical or chemical blocking strategies.[95, 168] An example of a chemical method can be seen in Figure 1.10, through the use of a SLB. Some of the most popular and effective methods at preventing NSB are also the simplest.

An example of a simple physical method to reduce NSB is through the use of blocker proteins. Blocker proteins will adsorb onto the surface without a substantial amount of disruption to the functionalized antigen on the surface.[171] Once the excess amount of blocker proteins have been removed from the surface, through a rinse cycle, the majority of the residual signal should be due to the desired analyte antigen interactions.[172] Other approaches rely on the complex biological matrix itself as the blocker. These strategies will use a version of the biological sample that either does not have the analyte present in it

and/or is a heavily diluted version to passivate the surface through saturation.[169] The saturation step works in an identical manner as when blocker proteins are used, but with less concern of unaccounted for interactions between the compounds in the analyte containing sample with the blocker proteins.[172] The goal for both of these processes is that only the analyte in a biological sample will bind with the antigens left available after being previously exposed to the biological components.

While using biological matrices to foul surfaces and reduce NSB is effective, it can lead to many potential complications.[168, 169, 171] First is the concern that too many potential antigen analyte binding locations will be blocked, thus reducing the potential observed analyte signal more drastically than using a different antifouling approach.[172] Second is that the captured proteins during the fouling step may lead to future nonspecific interactions that complicate the investigation and detection of the analyte.[169] Chemical methods such as SAM and SLB, offer much more reliable and stronger resistance to NSB, than relying purely on blocking steps.[173] However, the combination of these two strategies is ideal for addressing NSB.

While antifouling methods can be very reliable strategies to investigate biophysical interactions and aid in the prevention of NSB interaction, especially in complex media, they do not guarantee that the analyte antigen interactions of interest will be detectable. To help resolve this, signal amplification strategies can be utilized.

1.6.2 Weak Signal and Low Sensitivity:

As SPR based detection methods have been employed and characterized more for the detection of disease biomarkers, the ability to reliably detect biomarkers at very low

concentrations in biological fluids, as well as low molecular weight molecules, such as DNA and RNA, have increased exponentially.[119, 174, 175] For the detection of biomarkers at low concentrations or that have small molecular masses, signal amplification strategies have been employed routinely for SPR studies.[99, 150, 167] Antibodies are a common source of amplification as they offer high specificity for other proteins and even small molecules, such as lipids, and have large masses.[92] In addition to their high specificity and mass, antibodies can be easily conjugated in a variety of ways such as through biotin, histidine tags, tyrosine residues, and amine residues.[49, 95] The mechanism of functionalization is normally straightforward relying on established chemical reactions such as EDC/NHS coupling.[176]

For SPR based amplification, the goal can simply be put as the desire to increase the mass present on the surface to further shift the binding signals to allow for easier quantification. Further amplification beyond the increase in mass can be achieved by perturbations to the surface plasmons from optically active materials, such as Au and silica nanoparticles to enhance the overall SPR signal.[150] This process is referred to as plasmon-excitation coupling.[177] The dramatic increase in signal has made the use of Au and other plasmonically active nanoparticles a common tool for signal amplification strategies on a diverse range of analytes.[178]

Signal amplification capabilities have allowed for SPR-based techniques to dramatically increase the type and concentration range of biomarkers that are detectable.[49, 71, 96, 173] This has allowed SPR to be applied for the detection of many diverse disease related biomarkers, which is essential for accurate disease diagnosis and evaluations.[96, 117]

However, reliable detecting, monitoring, and quantifying biomarkers at disease relevant concentrations with or without amplification methods is not enough, due to the possibility of cross reactivity.[78, 111]

1.6.3 Cross Reactivity:

In general, cross reactivity is when multiple analytes bind to or interact with the same antigen,[179] a cartoon representation of this is shown in Figure 1.11. For the vast majority of biomedical applications cross reactivity is a major concern and challenge, as it can lead to false positives or even misdiagnosis.[111] When multiple analytes are attempting to be monitored simultaneously that have cross reactivity with one another, the conclusions of the investigations can be greatly hindered if not properly taken into account.[49] This is due to the fact that the cross reactivity between the analytes will cause a reduction in the desired analyte-antigen interactions.[180] In addition to being a major concern for biomarker detection, drug development studies also have to take into account cross reactivity when investigating the effects and properties of potential drugs.[181] Drugs that have cross reactivity beyond that of the designated target may lead to unwanted downstream biological interactions, which can have highly adverse effects to patients.[181]

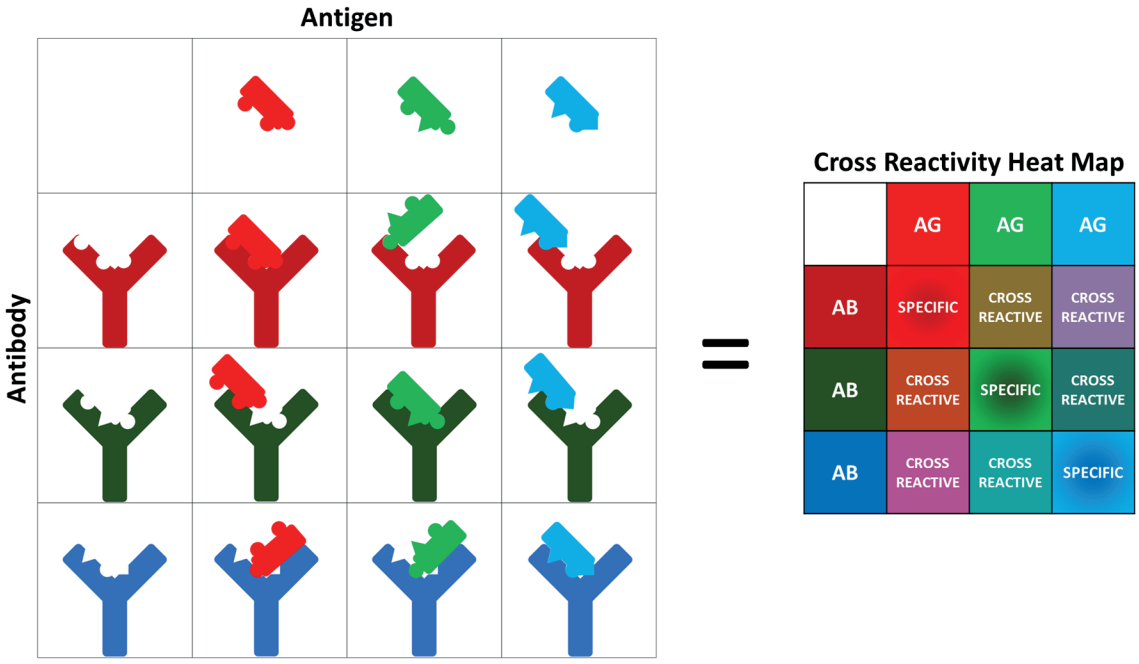


Figure 1.11. Graphic illustration showcasing specific analyte antigen interactions and cross reactive interactions.

Because cross reactivity in general is something that is considered undesirable, especially in the context of SPR biosensing studies,[182] a great deal of effort has been put into the development of strategies to account for it.[168] Most of these studies have relied on various antifouling techniques and combining them with post data acquisition analysis strategies.[138, 171, 173] One of the most useful approaches to account for cross reactivity is by monitoring each potential cross reactive analyte interactions with the selected antigen under the same conditions individually.[84] After identifying the known binding signal that occurs between each analyte-antigen interaction, the cross reactive interactions can be taken into account from the bulk change, normally through subtraction. This leaves only the observed interaction for the desired analyte-antigen interaction.[49, 169] While this approach has been integral in accounting for cross reactivity, as well as for nonspecific

binding,[169] other approaches are being developed and implemented to investigate and account for cross reactivity, such as through statistics and machine learning.

1.7 Machine Learning and Statistics:

Statistics in its purest form is the use of mathematical equations to resolve uncertainty.[183] However, the usefulness of statistics goes far beyond this, as it can drastically improve the analysis, validation, interpretation, and presentation of data.[184, 185] Many consider that the next evolution of statistics for data analysis and representation is through the use of machine learning algorithms.[186] Machine learning (ML) essentially is the training of an algorithm with a specific dataset and then determining the accuracy of the algorithm with a test dataset.[187] If the algorithm passes a determined threshold, usually presented in terms of accuracy, it can then be applied to solve various problems through the identification of trends and patterns in datasets similar to what was used to train and test the model.[188] This sections aims to provide an entry level introduction to the statistical analysis methods and ML algorithms utilized in the forthcoming Chapters. Example text for the code used in R programing language for all of discussed statistical and ML methods in this Dissertation, except for analysis of variance which was obtained via a plugin in excel, are shown in the Appendix section of this Dissertation.

1.7.1 Analysis of Variance:

Analysis of variance (ANOVA) is a collection of statistical methods that are used to estimate and identify variance, outliers, and other basic statistical information in a dataset and is centered around the law of total variance.[189] The law of total variance is based upon a mathematical equation, shown in Equation 1.24, that is founded on the idea,

that if there are two random variables that occupy a space there is a finite difference between them, where one term is explained and the other is unexplained variance.[190]

$$Var\,(Y) \;=\; E\,[\,Var\,(Y \mid X)] \;+\; Var\,(E\,[\,Y \mid X])$$ Equation 1.24

Here X and Y are random variable on the same probability space where the variance (Y) is finite, E represent explained variance in terms of the changes observed in X.

In many studies, ANOVA, which is the accumulation of several equations, is the first step in post data acquisition analysis.[184] The equations used in ANOVA to determine if the collected data is statistically relevant include the equations for sum of squares, degrees of freedom, variance, F value, mean of squares for between groups, error, and total as shown below.[189, 191] The application of these equations allows for the determination of whether or not the data passes the null hypothesis.

$$sum\ of\ squares\ between\ groups\ (SSB) \;=\; \sum_{i}^{n} (\overline{y}_{i} - \bar{y})^2$$ Equation 1.25

Where n is the number of observations, $\overline{y}_{i}$ is the mean of the i^{th} group, $\bar{y}$ is the mean value of the sample, and n is the sample size of that group to i.

$$sum\ of\ squares\ of\ errors(SSE) \;=\; \sum_{i} (y - \overline{y}_{i})^2$$ Equation 1.26

Where y is the data point in question and $\overline{y}_{i}$ is the average in the dataset i. To calculate the total sum of squares one simply needs to add Equations 1.25 and 1.26 together. The degrees of freedom between groups, error, and total are shown in Equations 1.27, 1.28, and 1.29 respectively.

$$degrees\ of\ freedom_{between\ groups} = k - 1$$ Equation 1.27

$$degrees\ of\ freedom_{error} = N - k \qquad \text{Equation 1.28}$$

$$degrees\ of\ freedom_{total} = N - 1 \qquad \text{Equation 1.29}$$

Where k is the number of groups and N is the total number of observations in group k. The mean of squares between groups can be calculated by dividing SSB by the degrees of freedom between groups, and the mean of squares for error is calculated by dividing SSE by degrees of freedom for error. By calculating the values of mean of squares for error and mean of squares between groups one can identify the F value with Equation 1.30.

$$F\ value\ = \frac{\left(\sum_i^n (\bar{y}_i - \bar{y})^2 \Big/ k - 1\right)}{\left(\sum_i (y - \bar{y}_i)^2 \Big/ N - k\right)} \qquad \text{Equation 1.30}$$

However, if one wanted to calculate the variance of a sample they would need to use Equation 1.31 below.

$$V = \frac{\sum_i^n (y_i - \bar{y})^2}{n - 1} \qquad \text{Equation 1.31}$$

Where V is variance, y_i is the value at i, $\bar{y}$ is the mean, and n is the total number of observations for that sample. To get the standard deviation of the sample can simply take the squares root of the variances, as shown in Equation 1.32.

$$Standard\ Deviation_{sample}\ (STD) = \sqrt{\frac{\sum_i^n (y_i - \bar{y})^2}{n - 1}} \qquad \text{Equation 1.32}$$

While not a part of ANOVA, Student's t-Test, are commonly used in combination as a way to rule out outliers.[192] The equation for Student's t-Test is shown in Equation 1.33.

$$Student's\ t-Test = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1} - \frac{s_2^2}{n_2}}} \qquad \text{Equation 1.33}$$

Where $\bar{x}$ is the mean of that sample, s is the sample's variance, and n is the sample size.

The combination of these equations allows for identification of the variation between groups, observed error, and total variance in a dataset. Interpretation is normally achieved through the P-value, which is widely considered the most significant value in the ANOVA test.[189] The P-value represents the probability that the F-value is larger than the observed value, or more simply put, the P-value is determined by the deviation between the observed value and a reference value.[191] If the P-value is less than 0.05, 5 %, the null hypotheses is rejected and the tested data is considered statistically relevant, meaning there is at least or above a 95 % confidence that the investigated values are different.[191] However, if the dataset is not large enough a low P-value can simply be due to chance,[193] which is why replicates and high quality data are integral for analytical studies. The clear benefits of ANOVA for data analysis has made it a common tool in various analytical chemistry applications.[194] Once a dataset is identified to be statistical relevant by rejecting the null hypothesis more complex statistical analysis tools can be employed. One of the most common statistical tools for visualization and analysis is through principal components analysis.

1.7.2 Principal Component Analysis:

Principal component analysis (PCA) is commonly used for the analysis of datasets that contain features that are suspected to be highly dependent or interact with one another,[195] as well as datasets that have high levels of dimensionality.[196] Processing high

dimensional data to be represented in a lower dimension while maintaining the original datasets information is very useful for deciphering complex interactions and visualizing the collected data.[197] This is achieved through the use of hyperplanes, where each plane is a principal component orthogonal to each other.[198] In the upcoming Chapters, PCA will be used extensively to aid in the identification and a means to measure the relationship between multiple principle components, or variables, by determining the significance that each one has towards one another and to the overall dataset.

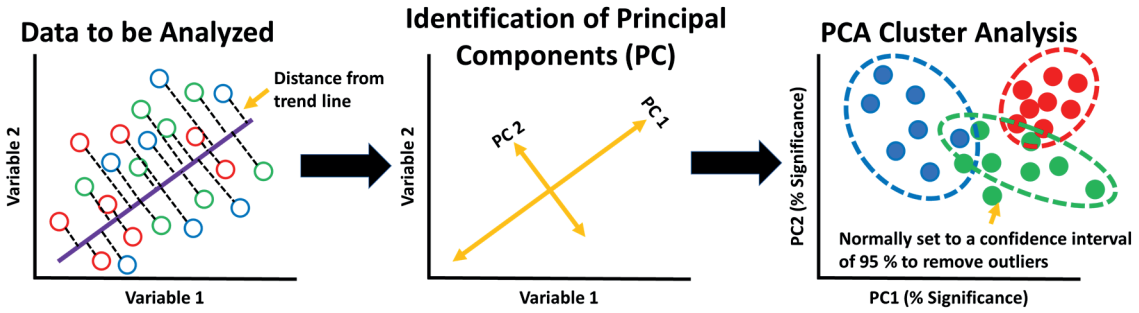


Figure 1.12. Graphic representation of PCA process by which the major principal components in a dataset are identified and is then reconstructed to show the groupings and differentiations between the input datasets in an more readable format.

A graphic representation by which PCA operates is shown in Figure 1.12, which is achieved through a large number of mathematical equations.[199] The purple line is the mathematical optimization to maximize the variance between each data point, which is therefore distinguished as principle component (PC) 1. A second PC, that is orthogonal to PC 1, is deemed PC 2 which is the dashed black lines. [199] Because the new PC is orthogonal and each PC is treated as a vector, mathematically each component is independent from the other, allowing for the representation of data in a more compact and easier to decipher format.[200] Each component has a percent significance related to it, which is the weight that component holds at differentiating each clusters from one another.[201] This is useful as it

visually shows that the data either has high or low differentiability from each other based upon specific features in the dataset.[199]

The final step in PCA is the visualization of the interactions; in most cases the confidence interval is set to 95 % to ensure that notable outliers are removed from the dataset.[195] In the imaginary example dataset of Figure 1.12, it can be seen that red and blue data points have high differentiability from one another. In contrast, the green data points have some notable confusion in regards to both red and blue data points, as well as having a clear outlier outside the 95 % confidence interval. This outlier could have easily been removed via a Student's t-test.[202]

It is key that when utilizing PCA, and other statistical models, that standardization of the data series is implemented when possible.[203] The most common strategies to achieve this, are either by finding ways to remove units or make all input variable have the same units. By standardizing the dataset, one can avoid the misrepresentation of the significance that each variable has.[200] While there are several strategies to achieve and avoid bias in one's dataset when using PCA, the most common approach is to subtract the mean from the data and divide by the standard deviation.[204] This process allows for the mean of that particular dataset to be 0, or the center of the PCs regardless of dimensionality. However, this approach does not work for every dataset, and other strategies may need to be employed, PCA still remains a powerful tool.[196, 205]

While PCA has many benefits that have been used for a diverse range of studies,[206] it can experience a large amount of difficulty when interpreting datasets where both X, Y, and Z or higher dimensional variables have categorical values.[196, 197, 200, 205] An example of

this, is that PCA is ideal for the analysis of SPR end point data, which is the value when subtracting the average of the pre-injection degree change or % RIU values from the average post rinse incubation degree change or % RIU values, but has major complications when analyzing SPRi sensorgram data.[84] To investigate datasets such as this other statistical methods must be employed, one of the most common of which are partial least squares discriminant analysis.

1.7.3 Partial Least Squares Discriminant Analysis:

Partial least squares (PLS) regression works under very similar principals as PCA does.[207] However, where PCA finds the maximum variables and represents them as hyperplanes,[200] PLS identifies the linear regression by analyzing the data series in a bilinear fashion.[208] Partial least squares discriminant analysis (PLS-DA) enhances PLS in many ways,[209] but most notably by allowing for two or more matrices, i.e. X and Y or multiple variables, to be categorical in determining the significance of each component.[210] This is something PCA cannot effectively do.[211] A simple cartoon representation comparing the capabilities of a linear (PCA) and nonlinear (PLS-DA) regression, is shown in Figure 1.13. This differentiation is significant as PCA will attempt to identify the dominant component in one direction,[199] whereas PLS-DA identifies the maximum multidimensional variance in both the X and Y space.[212] This is very useful for the analysis of data with more variables than observations, such as when analyzing SPR sensorgram data.[72] Sensorgram data holds host to various binding kinetics information that can be considered variables, but are not directly observable without implementing the previously mentioned equations.[18, 84] This strategy of analysis allows for the classification of interactions that PCA would be

incapable of achieving without modifying the data series extensively, which in most cases is undesirable.[196]

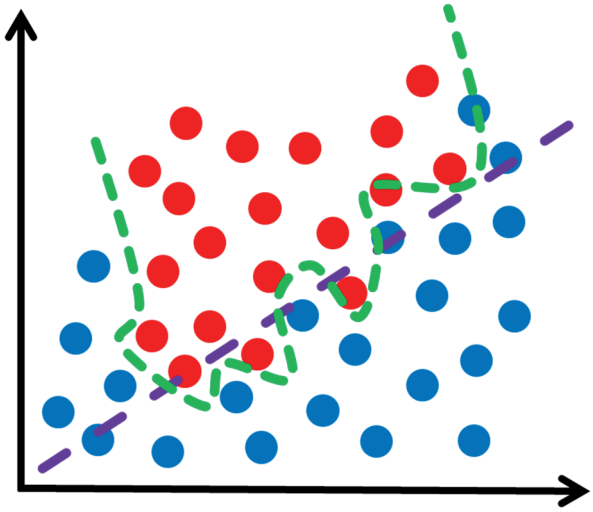


Figure 1.13. Shows a simple comparison between the general concept of data separation in a linear (purple) and nonlinear (green) regression model. Clearly in the hypothetical linear relation several red and blue data points would get misidentified or deemed as outliers compared to the analysis with a nonlinear regression approach.

While statistical algorithms such as ANOVA, PLS-DA, and PCA offer powerful data analysis capabilities, they are limited in that they can only identify trends that the user feeds into the model. Machine learning algorithms, however, allow the user to feed an entire data series and potentially find trends and correlations that the user was not aware of prior to the ML implementation.[213] While there are various types of ML algorithms available that have rigorously been investigated and applied for post data acquisition analysis,[186] this Dissertation will only discuss K-nearest neighbor (k-NN) and neural networks (nnet) in depth.

1.7.4 K-Nearest Neighbor:

K-NN is a type of supervised ML algorithm, meaning that the algorithm uses labeled datasets in its analysis, as opposed to unsupervised ML which uses unlabeled datasets.[214] K-NN is most often used as a classification model, where classification operates by creating defined boundaries that separate the data into specific categories.[215] One of k-NN models strongest capabilities is in its ability to recognize patterns in large datasets.[187] This ability to identify patterns relies on Euclidean distance matrices to locate the nearest neighbor in a dataset. Equation 1.34 is viewed by many as the major component for k-NN's ability to measure Euclidian distances in metrics of d(x, y).[187]

$$d(c,y) = \sum_{i}^{N} \sqrt{x_i^2 - y_i^2} \qquad \text{Equation 1.34}$$

Where N is the number of features for x_i to x_N and y_i to y_N. This process allows for the estimation of the conditional distribution of y and x, which can then be used to identify which classification the value has the highest probability of being associated with.[187] The k-NN classification will first identify the K points in a training dataset, N_0, that are closest to the test observation, x_0, which allows for the estimation of the probability that the data point is associated with class j based upon its relation with N_0,[187] as shown in equation 1.35.

$$Pr(Y = j|X = x_0) = \frac{1}{k}\sum_{i\in N_0} I(y_i = j) \qquad \text{Equation 1.35}$$

Where $I(y_i = j)$ is an indicator variable that is equal to 1 if $y_i = j$ or 0 if $y_i \neq j$, N_0 is the set of k-nearest observation in terms of the member class j, and Pr is the probability.[187]

While Equation 1.34 and 1.35 are just part of the many equations that make up k-NN models they are core to its classification model system.[216-219] For many modifying already established k-NN models in various programing languages is the most efficient way of utilizing ML models,[216] as was done for the work presented in this Dissertation. A simple example showcasing how k-NN models perform classification through the discussed equations is illustrated in Figure 1.14.

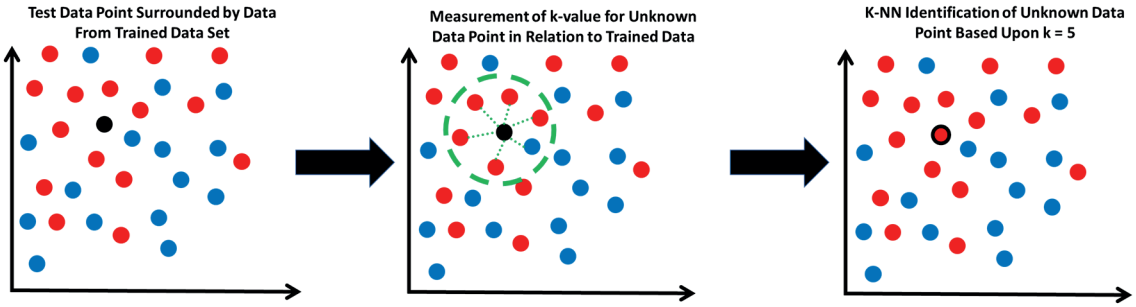


Figure 1.14. k-NN model determining the identity of a single unknown data point, test data, based upon its relation to training datasets that the model knows the identity of. The model determines how to classify the unknown values based upon the k value, which in the example is set to 5.

K-NN models are well known for being robust against noisy datasets and effective for large datasets, such as SPR sensorgram data, where there can be hundreds of thousands of data points.[215] When it comes to identifying what group a specific data point belongs to when using k-NN models the k value is the determinator, as discussed in the equations above.[217] The k value essentially measures the distance between a set number of neighbors and the value of interest that will be checked to determine the classification of that data point.[220] The more data points from the training set that the unknown value is affiliated with, will result in more weight being put towards that specific classification,[221] once it surpasses the k value the unknown value is identified as part of that classification.[222] This

process is normally achieved through regression which allows for the model to predict trends within the training dataset, and can then be applied to forecast the outcome with the test dataset.[223] K-NN models ability to identify unique trends in data series can be modified by changing parameters,[187] which control the learning process of the algorithm. Modifying these parameters can drastically improve the reliability and precision of the model, as well as aiding in the ML algorithm's ability at finding unique trends and relations in a dataset that linear models or human users would likely miss.[223]

The general benefits of k-NN algorithms compared to other ML models is less training time, easy and simple to train, and are widely considered to be more user friendly than other ML models.[217] However, k-NN models have some notable limitations, such as every new iteration of the model has to be retrained with new training datasets, thus to accurately represent the models reliability random iterations are commonly employed to measures the models average success rate over all potential iterations.[217] While k-NN is a powerful and robust ML algorithm there are other models that offer unique data analysis strategies, such as neural networks that may help uncover new trends due to the difference in how it analyzes the training and testing datasets.

1.7.5 Neural Networks:

Nnet's operate in a fashion that attempts to mimic the data processing capabilities of the neurons in the central nervous system.[224] This is very fitting as nnet were applied in this Dissertation to aid in the detection, differentiation, and characterization of several disease biomarkers associated with multiple sclerosis, which causes the demyelination of neurons.

In brief a nnet is composed of multiple node layers that are separated into an input layer, where the data is fed into the algorithm, one or more hidden layers, where the data is processed, and an output layer, which is how the data is visualized.[187] A graphic illustration of this process is shown in Figure 1.15.

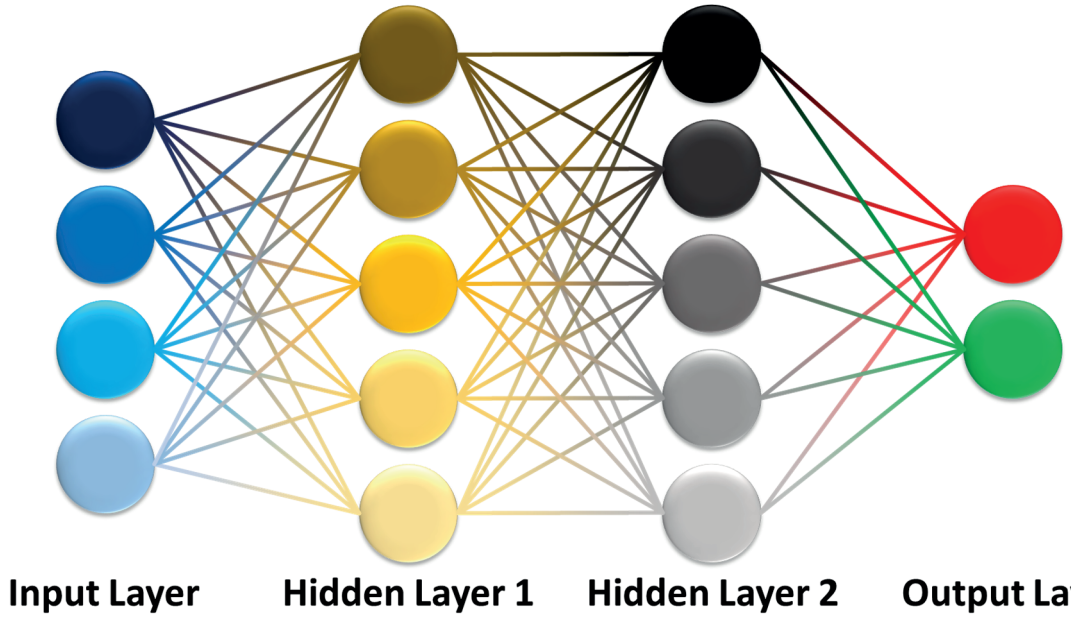


Figure 1.15. Graphical representation of a simple neural network that has 4 different input variables present in the input layer, has 2 hidden layers for computational analysis each with 5 weights to determine the significance of all data fed into the nnet, and a binary output layer for identification.

Because each feature that is investigated does not need to be a defined number, such as the use of an image's specific pixels, nnet can be applied in many diverse investigations.[206] Identifying the classification of a value through a nnet is routinely determined by the accuracy, precision, sensitivity, specificity, and F1 value that aid in the nnet classification based upon the input variables from the training dataset.[206] The equations for these are shown below in the context of a binary system.[206]

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{Equation 1.36}$$

$$precision = \frac{TP}{TP + FP} \quad \text{Equation 1.37}$$

$$sensitivity = \frac{TP}{TP + FN} \quad \text{Equation 1.38}$$

$$specificity = \frac{TN}{TN + FP} \quad \text{Equation 1.39}$$

$$F1 = \frac{2\left(\frac{TP}{TP + FP}\right)\left(\frac{TP}{TP + FN}\right)}{\left(\frac{TP}{TP + FP}\right) + \left(\frac{TP}{TP + FN}\right)} \quad \text{Equation 1.40}$$

Where TP is the true positive, TN is the true negative, FP is false positive, and FN is false negative.[206] All of these values will be identified using the training dataset through pattern recognition which allows the model to come to a conclusion when the test dataset is presented.

Each individual node can be viewed as a PLS-DA or PCA, that connects to another node that has an associated significance and tolerance in relation to each of the nodes it is connected with.[225] It is important to note that nnet's are nonlinear models even though they rely on linear regression strategies.[226] When the threshold is passed the node is activated and sends the data to the next layer of the network.[227] Identifying if a threshold has passed or not is dependent on the training dataset.[228] The more data that the nnet is trained with will inherently increase its capability of identifying if the inputted data will be passed on to the next node or not. This increase is due to the nnet having a better understanding of that node's threshold, which in turn leads to providing better analysis and predictions.[229]

While there are many benefits to nnet's, including nnet output generation having a high tolerance to error,[230] suitability for both classification and clustering analysis,[231] and relatively easy to train.[232] The most well-known and main limitation to nnet's is the black box problem, which stems from the user not understanding how the hidden layer operates.[233] This problem can be avoided if the features being analyzed have interpretable conclusions that can be cross validated by the user, and understanding how the data is processed in the hidden layer.[230] Another disadvantage of nnet's is the duration of time for the development, training, and testing of the algorithm compared to other ML models, such as k-NN models.[234] The time to build and train a nnet can be avoided by modifying already established nnet algorithms, which are readily available as plugins in various programing languages, such as R.[235] Modifying already established nnet code was done in the forthcoming Chapters and examples of these modifications can be seen in the Appendix of this Dissertation. Using already established and characterized nnet also aids in the user avoiding the black box problem. Some of the most well-known and used nnet packages are neuralnet, nnet, and RSNNS, all of which are available as plugins for various programing languages such as R.[235]

With the clear benefits of ML and statistics it is not surprising that their use is becoming more common in various scientific studies. The aims and scopes of this Dissertation would not have been achievable without using the statistical algorithms and ML models discussed in this introduction.

1.8 Aims and Scopes of Dissertation

The aims of this dissertation are to aid in the development of new strategies to detect, differentiate, and interpret various complex biological interactions with SPR based strategies. To accomplish this several biomimetic platforms were created for the investigation of a diverse range of biophysical interactions, as well as for the detection of disease specific biomarkers in complex matrices, such as human serum and urine, at biologically relevant concentrations. The collected data for these observed interactions were investigated with robust statistical analysis and ML algorithms in a manner that had not previously been explored before with SPR. These strategies were further implemented to account for cross reactivity, nonspecific binding, and the identification of the major components influencing the observed biophysical interactions. The forthcoming Chapters will showcase how these advancements were accomplished, present an in depth analysis of the collected data to support these claims, and a clear description of the methodologies developed to achieve these advancements.

Chapter 2 showcases the design, characterization, and proof of concept of a platform to allow for SPRi to be utilized for the detection of multiple disease biomarkers simultaneously. For this study several multiple sclerosis specific antibodies were selected as representatives for disease detection. The platform relied on a ganglioside array that was fabricated with a plasmonically tuned, background-free biochip, and coated with a perfluorodecyltrichlorosilane (PFDTS) layer for antigen attachment as a self-assembled pseudo-myelin sheath. The chip was characterized with AFM and matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS), demonstrating effective

functionalization of the surface. SPRi measurements of patients' mimicking blood samples, 10 % serum, were conducted. A multiplexed detection of antibodies for anti-GT_{1b}, anti-GM_1, and anti-GA_1 in 10 % serum was demonstrated, with a working range of 1 to 100 ng/mL, suggesting that it is well suited for clinical assessment of antibody abnormality in MS patients. Statistical analyses, including PLS-DA and PCA showed that the array allowed for the comprehensive characterization of cross reactivity patterns between the MS specific antibodies and can generate a wide range of information compared to traditional end point assays. This work used a PFDTS surface functionalization strategy that enabled direct MS biomarker detection in 10 % serum, offering a powerful alternative for MS assessment and potentially improved patient care.

Chapter 3 focuses on improving and using the characterized platform in Chapter 2 for the detection of multiple sclerosis biomarkers in undiluted human serum. A working range of 1–100 ng/mL was demonstrated with the limit of detection (LODs) below 7 ng/mL for each of the investigated biomarkers. The investigated biomarkers in this study are known to range between 3 to 25 ng/mL in blood samples, showcasing that this method was able to detect them at disease relevant concentrations without the need for complex amplification methods. This is of great significance as it shows this method can effectively be incorporated into currently established blood test with ease, drastically reducing the complexity of detecting and monitoring multiple sclerosis. In addition, machine learning (ML) algorithms were applied to the carbohydrate array/SPRi data analysis to understand and characterize the cross reactivities observed between the antibodies. Both endpoint results and SPRi sensorgrams were analyzed with statistical models for the evaluation of

binding events that include kinetic and steady state components. In addition, K-nearest neighbor (k-NN) and neural networks (nnet) were utilized to examine specific and cross-reactive binding, yielding higher accuracy than what traditional methods can achieve. The combination of ML models and microarray data provides a comprehensive understanding of complex interactions and could be used to differentiate and identify closely behaving biomarkers in a clinical setting.

Chapter 4 focuses on the development and characterization of a technique that can reliably generate curved membrane mimics with controllable curvatures and the ability to incorporate various cellular membrane components with ease. This strategy addresses the need for a reliable platform to generate reproducible and modifiable curved membrane mimics for biophysical studies between proteins and curved membranes. This is significant as to date, the vast majority of biomimetic membranes used for various biological studies have relied on planar membrane mimics, such as supported lipid bilayer (SLB) and self-assembled monolayers (SAM). While they have allowed for a great deal of information to be collected, the lack of curvature makes these models ineffective for the investigation of curvature dependent biophysical interactions. The developed curved membrane platform was generated through the use of an underlayer SLB composed of DGS-Ni-NTA-lipids on a silicated Au substrate. The interaction with histidine tagged cholera toxin (his-CT) provided the necessary orientation to ensure binding to lipid vesicles containing controlled percentage of gangliosides, including GA_1, GM_1, GT_{1b}, and GQ_{1b}. Characterization of the curved membrane platform was accomplished with surface plasmon resonance (SPR) spectroscopy, fluorescence recovery after photobleaching (FRAP), and nano tracking

analysis (NTA). Further investigation of the platform was achieved through principal component analysis (PCA) to measure the significance of ganglioside type, percent ganglioside, and vesicle size in regards to the observed biophysical interactions. In addition, Monte Carlo simulations were used to predict how curvature impacts the distribution of the gangliosides, as well as to determine if monovalent or multivalent interactions were occurring with his-CT. This work demonstrated a simple and reliable method to generate curved membrane mimics that can be easily modified in terms of size and composition to investigate various protein and lipid membrane biophysical interactions more accurately than what traditional SAMs and SLBs can achieve.

Chapter 5 focuses on the application of the platform characterized in Chapter 4. Specifically, the platform was used for the investigation of the curvature dependent biophysical interactions of the Bin-amphiphysin-Rvs (BAR) domain containing protein, bridging integrator 1 (BIN1), with several curved membrane mimics. Further evaluation was achieved by the detection of BIN1 in urine as a model system for the detection of several diseases, such as muscular dystrophy. In addition to BIN1's relation to muscular dystrophy in urine, it is a key component of membrane reshaping as it plays an essential role in vesicle generation. Recently, it was shown that BIN1 and several proteins that target BIN1 and its auxiliary domains, have been linked to various other diseases, such as lung, skin, and breast cancers, various cardiovascular diseases, skeletal muscle disorders, and skin diseases, like psoriasis. While BIN1 is known to play key roles in all of these diseases, the extent and how are currently not well understood. This is largely due to the fact that the vast majority of investigations regarding BIN1 to date have relied on cell based studies,

which lack the ability to control cellular membrane composition and curvature size. This Chapter aims to decipher the importance of curvature and composition of a membrane on the biophysical interactions with BIN1, as well as to detect BIN1 spiked into urine to showcase that the developed platform can be used as a detection method for muscular dystrophy.

1.9 References:

1. Nguyen, T.; Duong Bang, D.; Wolff, A., 2019 novel coronavirus disease (COVID-19): paving the road for rapid detection and point-of-care diagnostics. *Micromachines* **2020,** *11* (3), 306.

2. Wang, C.; Liu, M.; Wang, Z.; Li, S.; Deng, Y.; He, N., Point-of-care diagnostics for infectious diseases: From methods to devices. *Nano Today* **2021,** *37*, 101092.

3. Piret, J.; Boivin, G., Pandemics throughout history. *Frontiers in microbiology* **2021,** *11*, 631736.

4. Quintana-Murci, L., Understanding rare and common diseases in the context of human evolution. *Genome biology* **2016,** *17* (1), 1-14.

5. Baldacci, F.; Lista, S.; Vergallo, A.; Palermo, G.; Giorgi, F. S.; Hampel, H., A frontline defense against neurodegenerative diseases: the development of early disease detection methods. *Expert Review of Molecular Diagnostics* **2019,** *19* (7), 559-563.

6. Naresh, V.; Lee, N., A review on biosensors and recent development of nanostructured materials-enabled biosensors. *Sensors* **2021,** *21* (4), 1109.

7. Myszka, D. G.; Rich, R. L., Implementing surface plasmon resonance biosensors in drug discovery. *Pharmaceutical science technology today* **2000,** *3* (9), 310-317.

8. Ridker, P. M., Clinical application of C-reactive protein for cardiovascular disease detection and prevention. *Circulation* **2003,** *107* (3), 363-369.

9. Meklin, J.; SyrjÄnen, K.; Eskelinen, M., Fecal occult blood tests in colorectal cancer screening: systematic review and meta-analysis of traditional and new-generation fecal immunochemical tests. *Anticancer Research* **2020,** *40* (7), 3591-3604.

10. Sirovich, B. E.; Welch, H. G., The frequency of Pap smear screening in the United States. *Journal of general internal medicine* **2004,** *19*, 243-250.

11. Healy, D. A.; Hayes, C. J.; Leonard, P.; McKenna, L.; O'Kennedy, R., Biosensor developments: application to prostate-specific antigen detection. *TRENDS in Biotechnology* **2007,** *25* (3), 125-131.

12. Sogani, J.; Mango, V. L.; Keating, D.; Sung, J. S.; Jochelson, M. S., Contrast-enhanced mammography: past, present, and future. *Clinical imaging* **2021,** *69*, 269-279.

13. Ogunwobi, O. O.; Mahmood, F.; Akingboye, A., Biomarkers in colorectal cancer: current research and future prospects. *International journal of molecular sciences* **2020,** *21* (15), 5311.

14. Tang, L.; Chang, S. J.; Chen, C.-J.; Liu, J.-T., Non-invasive blood glucose monitoring technology: a review. *Sensors* **2020,** *20* (23), 6925.

15. Broza, Y. Y.; Zhou, X.; Yuan, M.; Qu, D.; Zheng, Y.; Vishinkin, R.; Khatib, M.; Wu, W.; Haick, H., Disease detection with molecular biomarkers: from chemistry of body fluids to nature-inspired chemical sensors. *Chemical reviews* **2019,** *119* (22), 11761-11817.

16. Saylan, Y.; Akgönüllü, S.; Denizli, A., Nanosensors for medical diagnosis. In *Nanotechnology for Hematology, Blood Transfusion, and Artificial Blood*, Elsevier: **2022**; pp 195-213.

17. Ghorbani, F.; Abbaszadeh, H.; Mehdizadeh, A.; Ebrahimi-Warkiani, M.; Rashidi, M.-R.; Yousefi, M., Biosensors and nanobiosensors for rapid detection of autoimmune diseases: a review. *Microchimica Acta* **2019,** *186*, 1-11.

18. Malinick, A. S.; Lambert, A. S.; Stuart, D. D.; Li, B.; Puente, E.; Cheng, Q., Detection of multiple sclerosis biomarkers in serum by ganglioside microarrays and surface plasmon resonance imaging. *ACS sensors* **2020,** *5* (11), 3617-3626.

19. Haleem, A.; Javaid, M.; Singh, R. P.; Suman, R.; Rab, S., Biosensors applications in medical field: A brief review. *Sensors International* **2021,** *2*, 100100.

20. Sin, M. L.; Mach, K. E.; Wong, P. K.; Liao, J. C., Advances and challenges in biosensor-based diagnosis of infectious diseases. *Expert review of molecular diagnostics* **2014,** *14* (2), 225-244.

21. Mayeux, R., Biomarkers: potential uses and limitations. *NeuroRx* **2004,** *1*, 182-188.

22. Balogh, E. P.; Miller, B. T.; Ball, J. R., Improving diagnosis in health care. **2015**.

23. Caliendo, A. M.; Gilbert, D. N.; Ginocchio, C. C.; Hanson, K. E.; May, L.; Quinn, T. C.; Tenover, F. C.; Alland, D.; Blaschke, A. J.; Bonomo, R. A., Better tests, better care: improved diagnostics for infectious diseases. *Clinical Infectious Diseases* **2013,** *57* (suppl_3), S139-S170.

24. Olson, A. P.; Graber, M. L.; Singh, H., Tracking progress in improving diagnosis: a framework for defining undesirable diagnostic events. *Journal of general internal medicine* **2018,** *33*, 1187-1191.

25. Solomon, A. J.; Corboy, J. R., The tension between early diagnosis and misdiagnosis of multiple sclerosis. *Nature Reviews Neurology* **2017,** *13* (9), 567-572.

26. Surinova, S.; Schiess, R.; Hüttenhain, R.; Cerciello, F.; Wollscheid, B.; Aebersold, R., On the development of plasma protein biomarkers. *Journal of proteome research* **2011,** *10* (1), 5-16.

27. Filippi, M.; Preziosa, P.; Banwell, B. L.; Barkhof, F.; Ciccarelli, O.; De Stefano, N.; Geurts, J. J.; Paul, F.; Reich, D. S.; Toosy, A. T., Assessment of lesions on magnetic resonance imaging in multiple sclerosis: practical guidelines. *Brain* **2019,** *142* (7), 1858-1875.

28. Rubin, G. D., Computed tomography: revolutionizing the practice of medicine for 40 years. *Radiology* **2014,** *273* (2S), S45-S74.

29. Shen, Y.-T.; Chen, L.; Yue, W.-W.; Xu, H.-X., Artificial intelligence in ultrasound. *European Journal of Radiology* **2021,** *139*, 109717.

30. Ciccarelli, O.; Cohen, J. A.; Reingold, S. C.; Weinshenker, B. G.; Amato, M. P.; Banwell, B.; Barkhof, F.; Bebo, B.; Becher, B.; Bethoux, F., Spinal cord involvement in multiple sclerosis and neuromyelitis optica spectrum disorders. *The Lancet Neurology* **2019,** *18* (2), 185-197.

31. Schwenkenbecher, P.; Wurster, U.; Konen, F. F.; Gingele, S.; Sühs, K.-W.; Wattjes, M. P.; Stangel, M.; Skripuletz, T., Impact of the McDonald criteria 2017 on early diagnosis of relapsing-remitting multiple sclerosis. *Frontiers in neurology* **2019,** *10*, 188.

32. Chen, M.; Zhao, H., Next-generation sequencing in liquid biopsy: cancer screening and early detection. *Human genomics* **2019,** *13*, 1-10.

33. Evans, R. W., Complications of lumbar puncture. *Neurologic clinics* **1998,** *16* (1), 83-105.

34. Filippi, M.; Rocca, M. A.; Ciccarelli, O.; De Stefano, N.; Evangelou, N.; Kappos, L.; Rovira, A.; Sastre-Garriga, J.; Tintorè, M.; Frederiksen, J. L., MRI criteria for the diagnosis of multiple sclerosis: MAGNIMS consensus guidelines. *The Lancet Neurology* **2016,** *15* (3), 292-303.

35. McDonald, W. I.; Compston, A.; Edan, G.; Goodkin, D.; Hartung, H. P.; Lublin, F. D.; McFarland, H. F.; Paty, D. W.; Polman, C. H.; Reingold, S. C., Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Annals of Neurology: Official Journal of the American Neurological Association the Child Neurology Society* **2001,** *50* (1), 121-127.

36. Parnetti, L.; Gaetani, L.; Eusebi, P.; Paciotti, S.; Hansson, O.; El-Agnaf, O.; Mollenhauer, B.; Blennow, K.; Calabresi, P., CSF and blood biomarkers for Parkinson's disease. *The Lancet Neurology* **2019,** *18* (6), 573-586.

37. Bach, J.-F., Infections and autoimmune diseases. *Journal of autoimmunity* **2005,** *25*, 74-80.

38. Lennon, V. A.; Wingerchuk, D. M.; Kryzer, T. J.; Pittock, S. J.; Lucchinetti, C. F.; Fujihara, K.; Nakashima, I.; Weinshenker, B. G., A serum autoantibody marker of neuromyelitis optica: distinction from multiple sclerosis. *The Lancet* **2004,** *364* (9451), 2106-2112.

39. Fava, A.; Petri, M., Systemic lupus erythematosus: diagnosis and clinical management. *Journal of autoimmunity* **2019,** *96*, 1-13.

40. Maffei, M. E., Fibromyalgia: recent advances in diagnosis, classification, pharmacotherapy and alternative remedies. *International journal of molecular sciences* **2020,** *21* (21), 7877.

41. Stoker, T. B.; Mason, S. L.; Greenland, J. C.; Holden, S. T.; Santini, H.; Barker, R. A., Huntington's disease: Diagnosis and management. *Practical Neurology* **2022,** *22* (1), 32-41.

42. Dubois, B.; Villain, N.; Frisoni, G. B.; Rabinovici, G. D.; Sabbagh, M.; Cappa, S.; Bejanin, A.; Bombois, S.; Epelbaum, S.; Teichmann, M., Clinical diagnosis of Alzheimer's disease: recommendations of the International Working Group. *The Lancet Neurology* **2021,** *20* (6), 484-496.

43. Stafford, I.; Kellermann, M.; Mossotto, E.; Beattie, R. M.; MacArthur, B. D.; Ennis, S., A systematic review of the applications of artificial intelligence and machine learning in autoimmune diseases. *NPJ digital medicine* **2020,** *3* (1), 30.

44. Florea, A.; Melinte, G.; Simon, I.; Cristea, C., Electrochemical biosensors as potential diagnostic devices for autoimmune diseases. *Biosensors* **2019,** *9* (1), 38.

45. Weiner, H. L., The challenge of multiple sclerosis: how do we cure a chronic heterogeneous disease? *Annals of Neurology: Official Journal of the American Neurological Association the Child Neurology Society* **2009,** *65* (3), 239-248.

46. Dobson, R.; Giovannoni, G., Multiple sclerosis–a review. *European journal of neurology* **2019,** *26* (1), 27-40.

47. Hansson, O., Biomarkers for neurodegenerative diseases. *Nature medicine* **2021,** *27* (6), 954-963.

48. Ashton, N. J.; Hye, A.; Rajkumar, A. P.; Leuzy, A.; Snowden, S.; Suárez-Calvet, M.; Karikari, T. K.; Schöll, M.; La Joie, R.; Rabinovici, G. D., An update on blood-based biomarkers for non-Alzheimer neurodegenerative disorders. *Nature Reviews Neurology* **2020,** *16* (5), 265-284.

49. Masson, J.-F., Surface plasmon resonance clinical biosensors for medical diagnostics. *ACS sensors* **2017,** *2* (1), 16-30.

50. Muinao, T.; Boruah, H. P. D.; Pal, M., Multi-biomarker panel signature as the key to diagnosis of ovarian cancer. *Heliyon* **2019,** *5* (12), e02826.

51. Ouyang, M.; Tu, D.; Tong, L.; Sarwar, M.; Bhimaraj, A.; Li, C.; Cote, G. L.; Di Carlo, D., A review of biosensor technologies for blood biomarkers toward monitoring cardiovascular diseases at the point-of-care. *Biosensors Bioelectronics* **2021,** *171*, 112621.

52. Cathcart, N.; Chen, J. I., Sensing biomarkers with plasmonics. *Analytical chemistry* **2020,** *92* (11), 7373-7381.

53. Martinsen, V.; Kursula, P., Multiple sclerosis and myelin basic protein: Insights into protein disorder and disease. *Amino Acids* **2022,** *54* (1), 99-109.

54. Baecher-Allan, C.; Kaskow, B. J.; Weiner, H. L., Multiple sclerosis: mechanisms and immunotherapy. *Neuron* **2018,** *97* (4), 742-768.

55. Quintana, F. J.; Yeste, A.; Weiner, H. L.; Covacu, R., Lipids and lipid-reactive antibodies as biomarkers for multiple sclerosis. *Journal of neuroimmunology* **2012,** *248* (1-2), 53-57.

56. Ziemssen, T.; Akgün, K.; Brück, W., Molecular biomarkers in multiple sclerosis. *Journal of neuroinflammation* **2019,** *16* (1), 272.

57. Cutillo, G.; Saariaho, A.-H.; Meri, S., Physiology of gangliosides and the role of antiganglioside antibodies in human diseases. *Cellular molecular immunology* **2020,** *17* (4), 313-322.

58. Wanleenuwat, P.; Iwanowski, P.; Kozubski, W., Antiganglioside antibodies in neurological diseases. *Journal of the neurological sciences* **2020,** *408*, 116576.

59. Serin, M.; Kara, P., Biosensing strategies (approaches) for diagnosis and monitoring of multiple sclerosis. *Talanta* **2022**, 123794.

60. Podbielska, M.; Ariga, T.; Pokryszko-Dragan, A., Sphingolipid Players in Multiple Sclerosis: Their Influence on the Initiation and Course of the Disease. *International Journal of Molecular Sciences* **2022,** *23* (10), 5330.

61. Höftberger, R.; Lassmann, H.; Berger, T.; Reindl, M., Pathogenic autoantibodies in multiple sclerosis—from a simple idea to a complex concept. *Nature Reviews Neurology* **2022**, 1-8.

62. Sapko, K.; Jamroz-Wiśniewska, A.; Marciniec, M.; Kulczyński, M.; Szczepańska-Szerej, A.; Rejdak, K., Biomarkers in Multiple Sclerosis: a review of diagnostic and prognostic factors. *Neurologia i neurochirurgia polska* **2020,** *54* (3), 252-258.

63. Chen, J. J.; Pittock, S. J.; Flanagan, E. P.; Lennon, V. A.; Bhatti, M. T., Optic neuritis in the era of biomarkers. *survey of ophthalmology* **2020,** *65* (1), 12-17.

64. Pohanka, M., Current trends in the biosensors for biological warfare agents assay. *Materials* **2019,** *12* (14), 2303.

65. Shanta, P. V.; Li, B.; Stuart, D. D.; Cheng, Q., Lipidomic profiling of algae with microarray MALDI-MS toward ecotoxicological monitoring of herbicide exposure. *Environmental Science Technology* **2021,** *55* (15), 10558-10568.

66. Gavrilaș, S.; Ursachi, C. Ș.; Perța-Crișan, S.; Munteanu, F.-D., Recent trends in biosensors for environmental quality monitoring. *Sensors* **2022,** *22* (4), 1513.

67. Chambers, J. P.; Arulanandam, B. P.; Matta, L. L.; Weis, A.; Valdes, J. J., Biosensor recognition elements. *Current issues in molecular biology* **2008,** *10* (1-2), 1-12.

68. Coulet, P. R., What is a Biosensor? *Biosensor principles applications* **2019**, 1-6.

69. Su, Y.; Hammond, M. C., RNA-based fluorescent biosensors for live cell imaging of small molecules and RNAs. *Current opinion in biotechnology* **2020,** *63*, 157-166.

70. Zhang, R.; Yuan, J., Responsive metal complex probes for time-gated luminescence biosensing and imaging. *Accounts of chemical research* **2020,** *53* (7), 1316-1329.

71. Hinman, S. S.; Cheng, Q., Bioinspired assemblies and plasmonic interfaces for electrochemical biosensing. *Journal of Electroanalytical Chemistry* **2016,** *781*, 136-146.

72. Singh, A.; Sharma, A.; Ahmed, A.; Sundramoorthy, A. K.; Furukawa, H.; Arya, S.; Khosla, A., Recent advances in electrochemical biosensors: Applications, challenges, and future scope. *Biosensors* **2021,** *11* (9), 336.

73. Hu, J.; Liu, F.; Chen, Y.; Shangguan, G.; Ju, H., Mass Spectrometric Biosensing: A Powerful Approach for Multiplexed Analysis of Clinical Biomolecules. *ACS sensors* **2021,** *6* (10), 3517-3535.

74. Song, L.; Chen, J.; Xu, B. B.; Huang, Y., Flexible plasmonic biosensors for healthcare monitoring: Progress and prospects. *ACS nano* **2021,** *15* (12), 18822-18847.

75. Zhang, J.; Zhang, X.; Wei, X.; Xue, Y.; Wan, H.; Wang, P., Recent advances in acoustic wave biosensors for the detection of disease-related biomarkers: A review. *Analytica Chimica Acta* **2021,** *1164*, 338321.

76. Andryukov, B. G., Six decades of lateral flow immunoassay: From determining metabolic markers to diagnosing COVID-19. *AIMS microbiology* **2020,** *6* (3), 280.

77. Jung, Y. H.; Park, B.; Kim, J. U.; Kim, T. i., Bioinspired electronics for artificial sensory systems. *Advanced Materials* **2019,** *31* (34), 1803637.

78. Tan, J.; Wen, Y.; Li, M., Emerging biosensing platforms for quantitative detection of exosomes as diagnostic biomarkers. *Coordination Chemistry Reviews* **2021,** *446*, 214111.

79. Holzlechner, M.; Eugenin, E.; Prideaux, B., Mass spectrometry imaging to detect lipid biomarkers and disease signatures in cancer. *Cancer Reports* **2019,** *2* (6), e1229.

80. Palla, G.; Fischer, D. S.; Regev, A.; Theis, F. J., Spatial components of molecular tissue biology. *Nature Biotechnology* **2022,** *40* (3), 308-318.

81. Soler, M.; Huertas, C. S.; Lechuga, L. M., Label-free plasmonic biosensors for point-of-care diagnostics: A review. *Expert review of molecular diagnostics* **2019,** *19* (1), 71-81.

82. Echle, A.; Rindtorff, N. T.; Brinker, T. J.; Luedde, T.; Pearson, A. T.; Kather, J. N., Deep learning in cancer pathology: a new generation of clinical biomarkers. *British journal of cancer* **2021,** *124* (4), 686-696.

83. Lisi, F.; Peterson, J. R.; Gooding, J. J., The application of personal glucose meters as universal point-of-care diagnostic tools. *Biosensors Bioelectronics* **2020,** *148*, 111835.

84. Malinick, A. S.; Stuart, D. D.; Lambert, A. S.; Cheng, Q., Surface plasmon resonance imaging (SPRi) in combination with machine learning for microarray analysis of multiple sclerosis biomarkers in whole serum. *Biosensors Bioelectronics: X* **2022,** *10*, 100127.

85. Kaur, B.; Kumar, S.; Kaushik, B. K., Recent advancements in optical biosensors for cancer detection. *Biosensors Bioelectronics* **2022,** *197*, 113805.

86. Samuel, V. R.; Rao, K. J., A review on label free biosensors. *Biosensors Bioelectronics: X* **2022**, 100216.

87. Iha, K.; Inada, M.; Kawada, N.; Nakaishi, K.; Watabe, S.; Tan, Y. H.; Shen, C.; Ke, L.-Y.; Yoshimura, T.; Ito, E., Ultrasensitive ELISA developed for diagnosis. *Diagnostics* **2019,** *9* (3), 78.

88. Hayrapetyan, H.; Tran, T.; Tellez-Corrales, E.; Madiraju, C., Enzyme-Linked Immunosorbent Assay: Types and Applications. *ELISA: Methods Protocols* **2023**, 1-17.

89. Wu, L.; Li, G.; Xu, X.; Zhu, L.; Huang, R.; Chen, X., Application of nano-ELISA in food analysis: Recent advances and challenges. *TrAC Trends in Analytical Chemistry* **2019,** *113*, 140-156.

90. Wadhera, T.; Kakkar, D.; Wadhwa, G.; Raj, B., Recent advances and progress in development of the field effect transistor biosensor: A review. *Journal of Electronic Materials* **2019,** *48*, 7635-7646.

91. Lim, H. J.; Saha, T.; Tey, B. T.; Tan, W. S.; Ooi, C. W., Quartz crystal microbalance-based biosensors as rapid diagnostic devices for infectious diseases. *Biosensors Bioelectronics* **2020,** *168*, 112513.

92. Phillips, K. S.; Cheng, Q., Recent advances in surface plasmon resonance based techniques for bioanalysis. *Analytical bioanalytical chemistry* **2007,** *387*, 1831-1840.

93. Abouhajar, F.; Chaudhuri, R.; Valiulis, S. N.; Stuart, D. D.; Malinick, A. S.; Xue, M.; Cheng, Q., Label-Free Analysis of Binding and Inhibition of SARS-Cov-19 Spike Proteins to ACE2 Receptor with ACE2-Derived Peptides by Surface Plasmon Resonance. *ACS Applied Bio Materials* **2022**.

94. Kuhlman, B.; Bradley, P., Advances in protein structure prediction and design. *Nature Reviews Molecular Cell Biology* **2019,** *20* (11), 681-697.

95. Hinman, S. S.; McKeating, K. S.; Cheng, Q., Surface plasmon resonance: material and interface design for universal accessibility. *Analytical chemistry* **2018,** *90* (1), 19.

96. Bocková, M.; Slabý, J.; Špringer, T.; Homola, J., Advances in surface plasmon resonance imaging and microscopy and their biological applications. *Annual Review of Analytical Chemistry* **2019,** *12*, 151-176.

97. Lambert, A. S.; Valiulis, S. N.; Malinick, A. S.; Tanabe, I.; Cheng, Q., Plasmonic biosensing with aluminum thin films under the Kretschmann configuration. *Analytical chemistry* **2020,** *92* (13), 8654-8659.

98. Schasfoort, R. B. M., Introduction to Surface Plasmon Resonance. In *Handbook of Surface Plasmon Resonance*, Schasfoort, R. B. M., Ed. The Royal Society of Chemistry: 2017; p 0.

99. Nguyen, H. H.; Park, J.; Kang, S.; Kim, M., Surface plasmon resonance: a versatile technique for biosensor applications. *Sensors* **2015,** *15* (5), 10481-10510.

100. Nivedha, S.; Babu, P. R.; Senthilnathan, K., Surface plasmon resonance. *Current Science* **2018,** *115* (1), 56-63.

101. Wu, L.; Guo, J.; Wang, Q.; Lu, S.; Dai, X.; Xiang, Y.; Fan, D., Sensitivity enhancement by using few-layer black phosphorus-graphene/TMDCs heterostructure in surface plasmon resonance biochemical sensor. *Sensors Actuators B: Chemical* **2017,** *249*, 542-548.

102. Ritchie, R. H., Plasma losses by fast electrons in thin films. *Physical review* **1957,** *106* (5), 874.

103. Powell, C.; Swan, J., Origin of the characteristic electron energy losses in aluminum. *Physical Review* **1959,** *115* (4), 869.

104. Liedberg, B.; Nylander, C.; Lundström, I., Biosensing with surface plasmon resonance—how it all started. *Biosensors Bioelectronics* **1995,** *10* (8), i-ix.

105. Kretschmann, E.; Raether, H., Radiative decay of non radiative surface plasmons excited by light. *Zeitschrift für Naturforschung A* **1968,** *23* (12), 2135-2136.

106. Otto, A., Excitation of nonradiative surface plasma waves in silver by the method of frustrated total reflection. *Zeitschrift für Physik A Hadrons nuclei* **1968,** *216* (4), 398-410.

107. Hoddeson, L. H.; Baym, G., The development of the quantum mechanical electron theory of metals: 1900-28. *Proceedings of the Royal Society of London. A. Mathematical Physical Sciences* **1980,** *371* (1744), 8-23.

108. Olmon, R. L.; Slovick, B.; Johnson, T. W.; Shelton, D.; Oh, S.-H.; Boreman, G. D.; Raschke, M. B., Optical dielectric function of gold. *Physical Review B* **2012,** *86* (23), 235147.

109. Xinglong, Y.; Dingxin, W.; Zibo, Y., Simulation and analysis of surface plasmon resonance biosensor based on phase detection. *Sensors Actuators B: Chemical* **2003,** *91* (1-3), 285-290.

110. Moon, G.; Lee, J.; Lee, H.; Yoo, H.; Ko, K.; Im, S.; Kim, D., Machine learning and its applications for plasmonics in biology. *Cell Reports Physical Science* **2022,** *3* (9).

111. Zain, H.; Batumalay, M.; Harith, Z.; Rahim, H.; Harun, S. In *Machine learning algorithms for surface plasmon resonance bio-detection applications, A short review*, Journal of Physics: Conference Series, IOP Publishing: **2022**; p 012013.

112. Wilkop, T.; Wang, Z.; Cheng, Q., Analysis of μ-contact printed protein patterns by SPR imaging with a LED light source. *Langmuir* **2004,** *20* (25), 11141-11148.

113. Gwon, H. R.; Lee, S. H., Spectral and angular responses of surface plasmon resonance based on the Kretschmann prism configuration. *Materials transactions* **2010,** *51* (6), 1150-1155.

114. Bonnet, H.; Coche-Guérente, L.; Defrancq, E.; Spinelli, N.; Van der Heyden, A.; Dejeu, J., Negative SPR signals during low molecular weight analyte recognition. *Analytical Chemistry* **2021,** *93* (8), 4134-4140.

115. Altug, H.; Oh, S.-H.; Maier, S. A.; Homola, J., Advances and applications of nanophotonic biosensors. *Nature nanotechnology* **2022,** *17* (1), 5-16.

116. Jung, L. S.; Campbell, C. T.; Chinowsky, T. M.; Mar, M. N.; Yee, S. S., Quantitative interpretation of the response of surface plasmon resonance sensors to adsorbed films. *Langmuir* **1998,** *14* (19), 5636-5648.

117. Wang, D.; Loo, J. F. C.; Chen, J.; Yam, Y.; Chen, S.-C.; He, H.; Kong, S. K.; Ho, H. P., Recent advances in surface plasmon resonance imaging sensors. *Sensors* **2019,** *19* (6), 1266.

118. Phillips, K. S.; Wilkop, T.; Wu, J.-J.; Al-Kaysi, R. O.; Cheng, Q., Surface plasmon resonance imaging analysis of protein-receptor binding in supported membrane arrays on gold substrates with calcinated silicate films. *Journal of the American Chemical Society* **2006,** *128* (30), 9590-9591.

119. Zhou, X. L.; Yang, Y.; Wang, S.; Liu, X. W., Surface plasmon resonance microscopy: From single-molecule sensing to single-cell imaging. *Angewandte Chemie International Edition* **2020,** *59* (5), 1776-1785.

120. Shanta, P. V.; Li, B.; Stuart, D. D.; Cheng, Q., Plasmonic gold templates enhancing single cell lipidomic analysis of microorganisms. *Analytical chemistry* **2020,** *92* (9), 6213-6217.

121. Liu, C.; Hu, F.; Yang, W.; Xu, J.; Chen, Y., A critical review of advances in surface plasmon resonance imaging sensitivity. *TrAC Trends in Analytical Chemistry* **2017,** *97*, 354-362.

122. Wink, T.; Van Zuilen, S.; Bult, A.; Van Bennekom, W., Self-assembled monolayers for biosensors. *Analyst* **1997,** *122* (4), 43R-50R.

123. Tabaei, S. R.; Choi, J.-H.; Haw Zan, G.; Zhdanov, V. P.; Cho, N.-J., Solvent-assisted lipid bilayer formation on silicon dioxide and gold. *Langmuir* **2014,** *30* (34), 10363-10373.

124. Jackman, J. A.; Knoll, W.; Cho, N.-J., Biotechnology applications of tethered lipid bilayer membranes. *Materials* **2012,** *5* (12), 2637-2657.

125. Wilkop, T.; Xu, D.; Cheng, Q., Surface plasmon resonance spectroscopic study on pore-forming behavior of streptolysin O on supported phospholipid bilayers. *MRS Online Proceedings Library* **2003,** *774*.

126. Phillips, K. S.; Han, J.-H.; Martinez, M.; Wang, Z.; Carter, D.; Cheng, Q., Nanoscale glassification of gold substrates for surface plasmon resonance analysis of protein toxins with supported lipid membranes. *Analytical chemistry* **2006,** *78* (2), 596-603.

127. Dong, Y.; Phillips, K. S.; Cheng, Q., Immunosensing of Staphylococcus enterotoxin B (SEB) in milk with PDMS microfluidic systems using reinforced supported bilayer membranes (r-SBMs). *Lab on a Chip* **2006,** *6* (5), 675-681.

128. Liu, Y.; Liao, P.; Cheng, Q.; Hooley, R. J., Protein and small molecule recognition properties of deep cavitands in a supported lipid membrane determined by calcination-enhanced SPR spectroscopy. *Journal of the American Chemical Society* **2010,** *132* (30), 10383-10390.

129. Perez, L.; Ghang, Y.-J.; Williams, P. B.; Wang, Y.; Cheng, Q.; Hooley, R. J., Cell and Protein Recognition at a Supported Bilayer Interface via In Situ Cavitand-Mediated Functional Polymer Growth. *Langmuir* **2015,** *31* (41), 11152-11157.

130. Davis, B.; Stuart, D.; Nieves Maldonado, M.; Kamavaram, V.; Cheng, Q.; Veedu, V. In *Intelligent Hydrogen Gas Monitoring in Natural Gas/Hydrogen Blending*, Offshore Technology Conference, OnePetro: **2022**.

131. Stuart, D. D.; Ebel, C. P.; Cheng, Q. J. S.; Reports, A., Biosensing empowered by molecular identification: Advances in surface plasmon resonance techniques coupled with mass spectrometry and Raman spectroscopy. **2022,** *4*, 100129.

132. Herrera-Domínguez, M.; Morales-Luna, G.; Mahlknecht, J.; Cheng, Q.; Aguilar-Hernández, I.; Ornelas-Soto, N., Optical Biosensors and Their Applications for the Detection of Water Pollutants. *Biosensors* **2023,** *13* (3), 370.

133. Couture, M.; Zhao, S. S.; Masson, J.-F., Modern surface plasmon resonance for bioanalytics and biophysics. *Physical Chemistry Chemical Physics* **2013,** *15* (27), 11190-11216.

134. Holdgate, G.; Embrey, K.; Milbradt, A.; Davies, G., Biophysical methods in early drug discovery. *ADMET DMPK* **2019,** *7* (4), 222-241.

135. Cheng, X.; Smith, J. C., Biological membrane organization and cellular signaling. *Chemical reviews* **2019,** *119* (9), 5849-5880.

136. Casares, D.; Escribá, P. V.; Rosselló, C. A., Membrane lipid composition: effect on membrane and organelle structure, function and compartmentalization and therapeutic avenues. *International journal of molecular sciences* **2019,** *20* (9), 2167.

137. Mukhopadhyay, R., Surface plasmon resonance instruments diversify. ACS Publications: **2005**.

138. McKeating, K. S.; Hinman, S. S.; Rais, N. A.; Zhou, Z.; Cheng, Q., Antifouling lipid membranes over protein A for orientation-controlled immunosensing in undiluted serum and plasma. *ACS sensors* **2019,** *4* (7), 1774-1782.

139. Siontorou, C. G.; Nikoleli, G.-P.; Nikolelis, D. P.; Karapetis, S. K., Artificial lipid membranes: Past, present, and future. *Membranes* **2017,** *7* (3), 38.

140. Hinman, S. S.; Ruiz, C. J.; Drakakaki, G.; Wilkop, T. E.; Cheng, Q., On-demand formation of supported lipid membrane arrays by trehalose-assisted vesicle delivery for SPR imaging. *ACS applied materials interfaces* **2015,** *7* (31), 17122-17130.

141. Richter, R. P.; Him, J. L. K.; Brisson, A., Supported lipid membranes. *Materials today* **2003,** *6* (11), 32-37.

142. Luchini, A.; Vitiello, G., Mimicking the mammalian plasma membrane: An overview of lipid membrane models for biophysical studies. *Biomimetics* **2020,** *6* (1), 3.

143. Khan, M. S.; Dosoky, N. S.; Williams, J. D., Engineering lipid bilayer membranes for protein studies. *International journal of molecular sciences* **2013,** *14* (11), 21561-21597.

144. Szoka Jr, F.; Papahadjopoulos, D., Comparative properties and methods of preparation of lipid vesicles (liposomes). *Annual review of biophysics bioengineering* **1980,** *9* (1), 467-508.

145. Richter, R. P.; Bérat, R.; Brisson, A. R., Formation of solid-supported lipid bilayers: an integrated view. *Langmuir* **2006,** *22* (8), 3497-3505.

146. Sundh, M.; Svedhem, S.; Sutherland, D. S., Formation of supported lipid bilayers at surfaces with controlled curvatures: Influence of lipid charge. *The Journal of Physical Chemistry B* **2011,** *115* (24), 7838-7848.

147. Singh, M.; Kaur, N.; Comini, E., The role of self-assembled monolayers in electronic devices. *Journal of Materials Chemistry C* **2020,** *8* (12), 3938-3955.

148. Mazur, F.; Bally, M.; Städler, B.; Chandrawati, R., Liposomes and lipid bilayers in biosensors. *Advances in colloid interface science* **2017,** *249*, 88-99.

149. Taylor, J. D.; Linman, M. J.; Wilkop, T.; Cheng, Q., Regenerable tethered bilayer lipid membrane arrays for multiplexed label-free analysis of lipid− protein interactions on poly (dimethylsiloxane) microchips using SPR imaging. *Analytical chemistry* **2009,** *81* (3), 1146-1153.

150. Yang, Z.; Malinick, A. S.; Yang, T.; Cheng, W.; Cheng, Q., Gold nanoparticle-coupled liposomes for enhanced plasmonic biosensing. *Sensors Actuators Reports* **2020,** *2* (1), 100023.

151. Belkilani, M.; Farre, C.; Chevalier, Y.; Minot, S.; Bessueille, F.; Abdelghani, A.; Jaffrezic-Renault, N.; Chaix, C., Mechanisms of Influenza Virus HA2 Peptide Interaction with Liposomes Studied by Dual-Wavelength MP-SPR. *ACS Applied Materials Interfaces* **2022,** *14* (29), 32970-32981.

152. Cawley, J. L.; Jordan, L. R.; Wittenberg, N. J., Detection and characterization of vesicular gangliosides binding to myelin-associated glycoprotein on supported lipid bilayers. *Analytical Chemistry* **2020,** *93* (2), 1185-1192.

153. Salehi-Reyhani, A.; Ces, O.; Elani, Y., Artificial cell mimics as simplified models for the study of cell biology. *Experimental Biology Medicine* **2017,** *242* (13), 1309-1317.

154. Peter, B. J.; Kent, H. M.; Mills, I. G.; Vallis, Y.; Butler, P. J. G.; Evans, P. R.; McMahon, H. T., BAR domains as sensors of membrane curvature: the amphiphysin BAR structure. *Science* **2004,** *303* (5657), 495-499.

155. Boyd, M. A.; Kamat, N. P., Designing artificial cells towards a new generation of biosensors. *Trends in biotechnology* **2021,** *39* (9), 927-939.

156. Jõemetsa, S.; Spustova, K.; Kustanovich, K.; Ainla, A.; Schindler, S.; Eigler, S.; Lobovkina, T.; Lara-Avila, S.; Jesorka, A.; Gözen, I., Molecular lipid films on microengineering materials. *Langmuir* **2019,** *35* (32), 10286-10298.

157. Los, D. A.; Murata, N., Membrane fluidity and its roles in the perception of environmental signals. *Biochimica et Biophysica Acta -Biomembranes* **2004,** *1666* (1-2), 142-157.

158. Eeman, M.; Deleu, M., From biological membranes to biomimetic model membranes. *Biotechnologie, Agronomie, Société et Environnement* **2010,** *14* (4).

159. Jenkins, E.; Santos, A. M.; O'Brien-Ball, C.; Felce, J. H.; Wilcock, M. J.; Hatherley, D.; Dustin, M. L.; Davis, S. J.; Eggeling, C.; Sezgin, E., Reconstitution of

immune cell interactions in free-standing membranes. *Journal of cell science* **2019,** *132* (4), jcs219709.

160. Has, C.; Sivadas, P.; Das, S. L., Insights into membrane curvature sensing and membrane remodeling by intrinsically disordered proteins and protein regions. *The Journal of Membrane Biology* **2022,** *255* (2-3), 237-259.

161. Raza, F.; Zafar, H.; Zhang, S.; Kamal, Z.; Su, J.; Yuan, W. E.; Mingfeng, Q., Recent advances in cell membrane-derived biomimetic nanotechnology for cancer immunotherapy. *Advanced healthcare materials* **2021,** *10* (6), 2002081.

162. Patching, S. G., Surface plasmon resonance spectroscopy for characterisation of membrane protein–ligand interactions and its potential for drug discovery. *Biochimica et Biophysica Acta -Biomembranes* **2014,** *1838* (1), 43-55.

163. Smith, T., The hydrophilic nature of a clean gold surface. *Journal of Colloid Interface Science* **1980,** *75* (1), 51-55.

164. Plant, A. L., Self-assembled phospholipid/alkanethiol biomimetic bilayers on gold. *Langmuir* **1993,** *9* (11), 2764-2767.

165. Bain, C. D.; Whitesides, G. M., Formation of monolayers by the coadsorption of thiols on gold: variation in the length of the alkyl chain. *Journal of the American Chemical Society* **1989,** *111* (18), 7164-7175.

166. Wijaya, E.; Lenaerts, C.; Maricot, S.; Hastanin, J.; Habraken, S.; Vilcot, J.-P.; Boukherroub, R.; Szunerits, S., Surface plasmon resonance-based biosensors: From the development of different SPR structures to novel surface functionalization strategies. *Current Opinion in Solid State Materials Science* **2011,** *15* (5), 208-224.

167. Lambert, A.; Yang, Z.; Cheng, W.; Lu, Z.; Liu, Y.; Cheng, Q., Ultrasensitive detection of bacterial protein toxins on patterned microarray via surface plasmon resonance imaging with signal amplification by conjugate nanoparticle clusters. *ACS sensors* **2018,** *3* (9), 1639-1646.

168. Frutiger, A.; Tanno, A.; Hwu, S.; Tiefenauer, R. F.; Voros, J.; Nakatsuka, N., Nonspecific binding—fundamental concepts and consequences for biosensing applications. *Chemical Reviews* **2021,** *121* (13), 8095-8160.

169. Masson, J.-F.; Battaglia, T. M.; Cramer, J.; Beaudoin, S.; Sierks, M.; Booksh, K. S., Reduction of nonspecific protein binding on surface plasmon resonance biosensors. *Analytical bioanalytical chemistry* **2006,** *386*, 1951-1959.

170. Myszka, D. G., Improving biosensor analysis. *Journal of molecular recognition* **1999,** *12* (5), 279-284.

171. Jiang, C.; Wang, G.; Hein, R.; Liu, N.; Luo, X.; Davis, J. J., Antifouling strategies for selective in vitro and in vivo sensing. *Chemical reviews* **2020,** *120* (8), 3852-3889.

172. Lichtenberg, J. Y.; Ling, Y.; Kim, S., Non-specific adsorption reduction methods in biosensing. *Sensors* **2019,** *19* (11), 2488.

173. D'Agata, R.; Bellassai, N.; Jungbluth, V.; Spoto, G., Recent advances in antifouling materials for surface plasmon resonance biosensing in clinical diagnostics and food safety. *Polymers* **2021,** *13* (12), 1929.

174. Ding, X.; Cheng, W.; Li, Y.; Wu, J.; Li, X.; Cheng, Q.; Ding, S., An enzyme-free surface plasmon resonance biosensing strategy for detection of DNA and small molecule based on nonlinear hybridization chain reaction. *Biosensors Bioelectronics* **2017,** *87*, 345-351.

175. Li, J.; Lei, P.; Ding, S.; Zhang, Y.; Yang, J.; Cheng, Q.; Yan, Y., An enzyme-free surface plasmon resonance biosensor for real-time detecting microRNA based on allosteric effect of mismatched catalytic hairpin assembly. *Biosensors Bioelectronics* **2016,** *77*, 435-441.

176. Jazayeri, M. H.; Amani, H.; Pourfatollah, A. A.; Pazoki-Toroudi, H.; Sedighimoghaddam, B., Various methods of gold nanoparticles (GNPs) conjugation to antibodies. *Sensing bio-sensing research* **2016,** *9*, 17-22.

177. Hutter, E.; Fendler, J. H., Exploitation of localized surface plasmon resonance. *Advanced materials* **2004,** *16* (19), 1685-1706.

178. Szunerits, S.; Spadavecchia, J.; Boukherroub, R., Surface plasmon resonance: Signal amplification using colloidal gold nanoparticles for enhanced sensitivity. *Reviews in Analytical Chemistry* **2014,** *33* (3), 153-164.

179. Albert, K. J.; Lewis, N. S.; Schauer, C. L.; Sotzing, G. A.; Stitzel, S. E.; Vaid, T. P.; Walt, D. R., Cross-reactive chemical sensor arrays. *Chemical reviews* **2000,** *100* (7), 2595-2626.

180. Juncker, D.; Bergeron, S.; Laforte, V.; Li, H., Cross-reactivity in antibody microarrays and multiplexed sandwich assays: shedding light on the dark side of multiplexing. *Current opinion in chemical biology* **2014,** *18*, 29-37.

181. Romano, A.; Guéant-Rodriguez, R.-M.; Viola, M.; Gaeta, F.; Caruso, C.; Guéant, J.-L., Cross-reactivity among drugs: clinical problems. *Toxicology* **2005,** *209* (2), 169-179.

182. Ledsgaard, L.; Jenkins, T. P.; Davidsen, K.; Krause, K. E.; Martos-Esteban, A.; Engmark, M.; Rørdam Andersen, M.; Lund, O.; Laustsen, A. H., Antibody cross-reactivity in antivenom research. *Toxins* **2018,** *10* (10), 393.

183. González, A. G.; Herrador, M. Á., A practical guide to analytical method validation, including measurement uncertainty and accuracy profiles. *TrAC Trends in Analytical Chemistry* **2007,** *26* (3), 227-238.

184. Miller, J. C.; Miller, J. N., Basic statistical methods for analytical chemistry. Part I. Statistics of repeated measurements. A review. *Analyst* **1988,** *113* (9), 1351-1356.

185. Miller, J. N., Basic statistical methods for analytical chemistry. Part 2. Calibration and regression methods. A review. *Analyst* **1991,** *116* (1), 3-14.

186. Mahesh, B., Machine learning algorithms-a review. *International Journal of Science Research* **2020,** *9*, 381-386.

187. Boateng, E. Y.; Otoo, J.; Abaye, D. A., Basic tenets of classification algorithms K-nearest-neighbor, support vector machine, random forest and neural network: a review. *Journal of Data Analysis Information Processing* **2020,** *8* (4), 341-357.

188. Debus, B.; Parastar, H.; Harrington, P.; Kirsanov, D., Deep learning in analytical chemistry. *TrAC Trends in Analytical Chemistry* **2021,** *145*, 116459.

189. Larson, M. G., Analysis of variance. *Circulation* **2008,** *117* (1), 115-121.

190. Bertsekas, D.; Tsitsiklis, J. N., *Introduction to probability*. Athena Scientific: **2008**; Vol. 1.

191. Scheffe, H., *The analysis of variance*. John Wiley & Sons: **1999**; Vol. 72.

192. Limentani, G. B.; Ringo, M. C.; Ye, F.; Bergquist, M. L.; McSorley, E. O., Beyond the t-test: statistical equivalence testing. ACS Publications: 2005.

193. Gamage, J.; Weerahandi, S., Size performance of some tests in one-way ANOVA. *Communications in Statistics-Simulation Computation* **1998,** *27* (3), 625-640.

194. Hirsch, R. F., Analysis of variance in analytical chemistry. *Analytical Chemistry* **1977,** *49* (8), 691A-700A.

195. Abdi, H.; Williams, L. J., Principal component analysis. *Wiley interdisciplinary reviews: computational statistics* **2010,** *2* (4), 433-459.

196. Bro, R.; Smilde, A. K., Principal component analysis. *Analytical methods* **2014,** *6* (9), 2812-2831.

197. Ivosev, G.; Burton, L.; Bonner, R., Dimensionality reduction and visualization in principal component analysis. *Analytical chemistry* **2008,** *80* (13), 4933-4944.

198. Dunteman, G. H., *Principal components analysis*. Sage: **1989**.

199. Shlens, J., A tutorial on principal component analysis. *arXiv preprint arXiv:*. **2014**.

200. Wold, S.; Esbensen, K.; Geladi, P., Principal component analysis. *Chemometrics intelligent laboratory systems* **1987,** *2* (1-3), 37-52.

201. Greenacre, M.; Groenen, P. J.; Hastie, T.; d'Enza, A. I.; Markos, A.; Tuzhilina, E., Principal component analysis. *Nature Reviews Methods Primers* **2022,** *2* (1), 100.

202. Livingston, E. H. J. J. o. S. R., Who was student and why do we care so much about his t-test? 1. **2004,** *118* (1), 58-65.

203. Gewers, F. L.; Ferreira, G. R.; Arruda, H. F. D.; Silva, F. N.; Comin, C. H.; Amancio, D. R.; Costa, L. d. F., Principal component analysis: A natural approach to data exploration. *ACM Computing Surveys* **2021,** *54* (4), 1-34.

204. Sauro, J.; Kindlund, E. In *A method to standardize usability metrics into a single score*, Proceedings of the SIGCHI conference on Human factors in computing systems, **2005**; pp 401-409.

205. Ilin, A.; Raiko, T., Practical approaches to principal component analysis in the presence of missing values. *The Journal of Machine Learning Research* **2010,** *11*, 1957-2000.

206. Schackart III, K. E.; Yoon, J.-Y., Machine learning enhances the performance of bioreceptor-free biosensors. *Sensors* **2021,** *21* (16), 5519.

207. Worley, B.; Halouska, S.; Powers, R., Utilities for quantifying separation in PCA/PLS-DA scores plots. *Analytical biochemistry* **2013,** *433* (2), 102-104.

208. Geladi, P.; Kowalski, B. R., Partial least-squares regression: a tutorial. *Analytica chimica acta* **1986,** *185*, 1-17.

209. Brereton, R. G.; Lloyd, G. R., Partial least squares discriminant analysis: taking the magic away. *Journal of Chemometrics* **2014,** *28* (4), 213-225.

210. Gromski, P. S.; Muhamadali, H.; Ellis, D. I.; Xu, Y.; Correa, E.; Turner, M. L.; Goodacre, R., A tutorial review: Metabolomics and partial least squares-discriminant analysis–a marriage of convenience or a shotgun wedding. *Analytica chimica acta* **2015,** *879*, 10-23.

211. Rosipal, R.; Krämer, N. In *Overview and recent advances in partial least squares*, Subspace, Latent Structure and Feature Selection: Statistical and Optimization Perspectives Workshop, SLSFS 2005, Bohinj, Slovenia, February 23-25, 2005, Revised Selected Papers, Springer: **2006**; pp 34-51.

212. Chevallier, S.; Bertrand, D.; Kohler, A.; Courcoux, P., Application of PLS-DA in multivariate image analysis. *Journal of Chemometrics: A Journal of the Chemometrics Society* **2006,** *20* (5), 221-229.

213. Ij, H., Statistics versus machine learning. *Nat Methods* **2018,** *15* (4), 233.

214. Alloghani, M.; Al-Jumeily, D.; Mustafina, J.; Hussain, A.; Aljaaf, A., A systematic review on supervised and unsupervised machine learning algorithms for data science. *Supervised unsupervised learning for data science* **2020**, 3-21.

215. Ayres, L. B.; Gomez, F. J.; Linton, J. R.; Silva, M. F.; Garcia, C. D., Taking the leap between analytical chemistry and artificial intelligence: A tutorial review. *Analytica Chimica Acta* **2021,** *1161*, 338403.

216. Zhang, Z., Introduction to machine learning: k-nearest neighbors. *Annals of translational medicine* **2016,** *4* (11).

217. Cunningham, P.; Delany, S. J., k-Nearest neighbour classifiers-A Tutorial. *ACM computing surveys* **2021,** *54* (6), 1-25.

218. Peterson, L. E., K-nearest neighbor. *Scholarpedia* **2009,** *4* (2), 1883.

219. Ali, N.; Neagu, D.; Trundle, P., Evaluation of k-nearest neighbour classifier performance for heterogeneous datasets. *Applied Sciences* **2019,** *1*, 1-15.

220. Abu Alfeilat, H. A.; Hassanat, A. B.; Lasassmeh, O.; Tarawneh, A. S.; Alhasanat, M. B.; Eyal Salman, H. S.; Prasath, V. S., Effects of distance measure choice on k-nearest neighbor classifier performance: a review. *Big data* **2019,** *7* (4), 221-248.

221. Kataria, A.; Singh, M., A review of data classification using k-nearest neighbour algorithm. *International Journal of Emerging Technology Advanced Engineering* **2013,** *3* (6), 354-360.

222. Cui, F.; Yue, Y.; Zhang, Y.; Zhang, Z.; Zhou, H. S., Advancing biosensors with machine learning. *ACS sensors* **2020,** *5* (11), 3346-3364.

223. Taunk, K.; De, S.; Verma, S.; Swetapadma, A. In *A brief review of nearest neighbor algorithm for learning and classification*, 2019 International Conference on Intelligent Computing and Control Systems (ICCS), IEEE: 2019; pp 1255-1260.

224. Bishop, C. M., Neural networks and their applications. *Review of scientific instruments* **1994,** *65* (6), 1803-1832.

225. Ehlers, R. In *Formal verification of piece-wise linear feed-forward neural networks*, Automated Technology for Verification and Analysis: 15th International Symposium, ATVA 2017, Pune, India, October 3–6, 2017, Proceedings 15, Springer: **2017**; pp 269-286.

226. Samek, W.; Montavon, G.; Lapuschkin, S.; Anders, C. J.; Müller, K.-R., Explaining deep neural networks and beyond: A review of methods and applications. *Proceedings of the IEEE* **2021,** *109* (3), 247-278.

227. Krogh, A., What are artificial neural networks? *Nature biotechnology* **2008,** *26* (2), 195-197.

228. Simen, P.; Cohen, J. D.; Holmes, P., Rapid decision threshold modulation by reward rate in a neural network. *Neural networks* **2006,** *19* (8), 1013-1026.

229. Günther, F.; Fritsch, S., Neuralnet: training of neural networks. *R J.* **2010,** *2* (1), 30.

230. Thakur, A.; Konde, A., Fundamentals of neural networks. *International Journal for Research in Applied Science Engineering Technology* **2021,** *9*, 407-26.

231. Liao, S.-H.; Wen, C.-H., Artificial neural networks classification and clustering of methodologies and applications–literature analysis from 1995 to 2005. *Expert Systems with applications* **2007,** *32* (1), 1-11.

232. Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J., Recent advances in convolutional neural networks. *Pattern recognition* **2018,** *77*, 354-377.

233. Dayhoff, J. E.; DeLeo, J. M., Artificial neural networks: opening the black box. *Cancer: Interdisciplinary International Journal of the American Cancer Society* **2001,** *91* (S8), 1615-1635.

234. Tu, J. V., Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of clinical epidemiology* **1996,** *49* (11), 1225-1231.

235. Beck, M. W., NeuralNetTools: Visualization and analysis tools for neural networks. *Journal of statistical software* **2018,** *85* (11), 1.

Chapter 2: Detection of Multiple Sclerosis Biomarkers in Serum by Ganglioside Microarrays and Surface Plasmon Resonance Imaging

2.1 Introduction:

It is estimated that almost two million people are affected by multiple sclerosis (MS) worldwide, but the true number of MS patients remains unknown due to the difficulty of obtaining an accurate diagnosis.[1] Although there have been attempts to link MS to genetics, environmental conditions, dietary restrictions, or viruses as the underlying factors that lead to its development, many of these studies have proved to be inconclusive.[1, 2] While the underlying causes of MS remain elusive, the process by which the symptoms associated with MS are displayed is well understood.[3] The most common symptoms observed in MS patients are numbness, slurred speech, paralysis, vertigo, impotence, tremors, loss of muscle control, and change in or loss of vision.[4] These symptoms have been directly associated with damage to the myelin sheath in the nervous system. The myelin sheath is a lipid rich material: 0.3 % of which are gangliosides, 15 to 30 % are various proteins, and the remaining 70 to 85 % are various lipids.[5] Gangliosides are sialic acid-containing glycosphingolipids and are essential for cell-cell recognition, adhesion, and signal transduction throughout the CNS.[6] In MS patients, the concentrations of anti- ganglioside antibodies in serum range between 3 ng/mL to 20 ng/mL.[7] These antibodies may initiate the attack on myelin sheath as well as oligodendrocytes, which leads to demyelination. A number of symptoms in MS have been suggested to be associated with specific anti-ganglioside antibodies.[8] For instance, anti-GA_1 and anti-GM_1 have been linked to optic nerve damage resulting in symptoms associated with eye movement and inflammation,

while anti-GT_{1b} is believed to be associated with loss of muscle control, specifically in the upper and lower limbs.[9-12]

Current diagnostic techniques for MS are considered to be inconsistent and unreliable.[13] This is primarily due to the variation of symptoms expressed from patient to patient. Magnetic resonance imaging (MRI) and electrophysiologic recordings are often used in tandem to assess whether an autoimmune attack has occurred and caused damage to the CNS.[14-16] To further confirm the diagnosis, blood tests and spinal taps are then conducted to rule out other diseases that show similar symptoms.[16, 17] Spinal taps or lumbar punctures are employed to collect cerebrospinal fluid (CSF) for elevated levels of autoimmune antibodies, oligoclonal bands, and white blood cells.[18, 19] While lumbar punctures facilitate the detection of MS, it is an unattractive avenue for disease progression monitoring due to limited sample size that can be extracted, difficulty in performing the procedure, and the inability to routinely collect CSF.[19] The last issue proves to be a major hurdle since accurate diagnosis through these methods can range anywhere between a few months to several years.[20] Blood tests thus have gained considerable interest in recent years as a way to detect and monitor the progression of MS due to their quick and minimally invasive collection procedure. Unlike lumbar punctures, blood samples can be easily collected in large quantities and at different timepoints. However, common biomarkers often have much lower concentrations in blood than in CSF, which is the reason why highly sensitive detection is urgently needed.[17, 21-24] While there currently is no definitive blood test that allows unambiguous MS diagnosis, these tests are routinely conducted to determine if other diseases may be present that have similar symptoms as MS. It should be

noted that the common blood biomarker tests intended for MS diagnosis also include those markers for Lyme disease, rare hereditary disorders, syphilis, HIV/AIDS, lupus, and fibromyalgia.[17, 21-23] ELISA has been the method of choice for antibody detection and serologic diagnosis. The CDC has called for the development of new tests of these markers in an advanced format that can be more informative. Microarray technology is capable of detecting antibodies at biologically relevant concentrations and under identical testing conditions, which would give new insights into which antibodies are linked to specific inflammation and damage to the CNS. The microarray test of anti- ganglioside antibodies and a quantitative assessment of their cross reactivity could also simplify the diagnostic procedure, combining the benefits of blood tests and spinal taps for effective detection, characterization, and evaluation of MS.

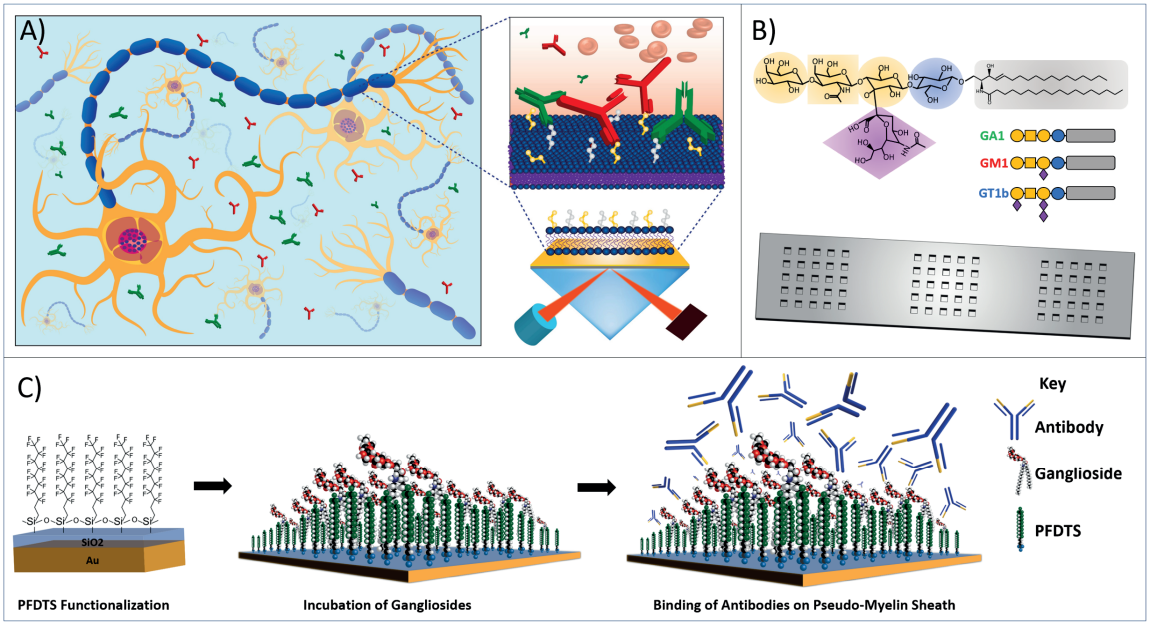


Figure 2.1. (A) Cartoon illustration of MS antibody attack on the myelin sheath of nerve cells, their circulation in the blood stream and detection by an SPR sensor with a membrane-mimicking interface. (B) Structure of GM_1 ganglioside and its headgroup carbohydrates: blue circle for glucose (Glc), orange circle for galactose (Gal), orange box for N-acetylgalactosamine (GalNAC), purple diamond for N-acteylneuraminic (NeuAc); along with structures for GA_1 (green) and GT_{1b} (blue). (C) Surface functionalization steps and the detection scheme for MS specific antibodies in serum.

In this work, we report a multiplexed detection for a series of MS biomarkers in serum by using surface plasmon resonance imaging (SPRi) in combination with ganglioside microarrays (Figure 2.1.). SPRi is a label free, real time, sensitive and direct detection method,[25] and has found broad applications in drug delivery,[26] cell based analysis,[27] biomarker profiling,[28] and biotechnology.[29] The ganglioside array used in this work was built with plasmonically tuned, background-free biochips and a coating of a perfluorodecyl-trichlorosilane (PFDTS) monolayer. PFDTS is a near superhydrophobic agent that forms compact monolayers on silica coatings on our SPR chips, which exhibits antifouling properties[30] and allows simple functionalization with gangliosides (Figure

2.1C). Three gangliosides (GA_1, GM_1, and GT_{1b}), structures of which are shown in Figure 2.1B, were chosen for this work and incorporated into the PFDTS microarray that mimic the myelin sheath. These gangliosides were chosen for initial test because antibodies for these antigens are pathologically relevant and are associated with common MS symptoms such as rapid eye movement, loss or change of vision, and loss of upper and lower muscle control.[8-12] Once verified, this microarray method can expand and include additional gangliosides for a more comprehensive testing.

2.2 Materials and Methods:

Materials and Reagents: *1H,1H,2H,2H*-Perfluorodecyltrichlorosilane (PFDTS) was obtained from Fisher Scientific (Pittsburgh, PA). Monoganglioside GM_1 was obtained from Matreya (Pleasant Gap, PA). Trisialoganglioside GT_{1b} was purchased from Biosynth (Itsaca, Il). Anti monoganglioside GM_1 rabbit polyclonal antibody, anti asialoganglioside GA_1 human anti mouse monoclonal antibody were purchased from Abcam (Cambridge, UK). Anti Trisialoganglioside GT_{1b} ganglioside mouse monoclonal antibody was purchased from Millipore Sigma (Billerica, MA). Asialoganglioside GA_1, anti Trisialoganglioside anti-GT_{1b} antibody heavy chain specific human anti mouse monoclonal antibody, α-cyano-4- hydroxycinnamic acid (CHCA) were from Sigma-Aldrich (St. Louis, MO). Human serum was purchased from Innovative Research (Upper Marlboro, MD).

Fabrication of SPR and SPRi Substrates: Fabrication of SPR and SPRi chips was performed following a previous procedure published by our group.[31] In brief, 2 nm of titanium (0.5 Å/s) followed by 48 nm of gold (2.0 Å/s), were deposited on slides via electron beam physical vapor deposition (EBPVD) (Temescal, Berkeley, CA). This is

followed by depositing roughly 1-3 nm of SiO_2 onto the gold layer via plasma enhanced chemical vapor deposition (PECVD) using a Unaxis Plasmatherm 790 system (Santa Clara, CA). For SPRi arrays, the slides were spin-coated with hexamethyldisilazane (HMDS) to promote adhesion with AZ5214E at 4000 rpm for 45 s. After baking at 110 °C for approximately 1 minute, UV exposure via a Karl- Suss MA- 6 system was used to create an array pattern on the photoresist, which was followed by standard photoresist development protocols. The well walls were formed by a 2 nm layer of titanium as an adhesion layer followed by 200 nm of gold both deposited via EBPVD. The remaining photoresist was lifted from the surface with acetone. Once all the wells were removed, another layer of 2 nm of titanium was deposited followed by 48 nm of gold both deposited via EBPVD to form the sensing surface inside of the wells. PECVD was then used to deposit 1-3 nm of SiO_2 on the microarray chips. The SPRi microarray chips consisted of 10 x 10 well arrays that were 600 μm in diameter and 200 nm deep.

Surface Functionalization and Preparation: For functionalization, the chips were submerged in 1 mM PFDTS that was diluted in toluene. After 30 minutes, the chip was removed from the solution and rinsed with toluene, ethanol, and deionized water and dried with nitrogen gas. Once the chips were dry, 20 μL of 100 μg/mL GA_1, GM_1, or GT_{1b} gangliosides in chloroform were pipetted onto the chips and covered with a glass cover slip to allow the gangliosides to evenly cover the SPR chips surface. The chips were then allowed to dry for 5 minutes under ambient conditions. For SPRi chips, 1.5 μL of the 100 μg/mL stock solution for each ganglioside were incubated into each well of their respective row to create four different working channels (eight working wells per channel) when

attached to the PDMS flow cell. Each chip was fabricated using the same configuration: the first channel is left non-functionalized as an internal standard to determine chip to chip variation, while 1 μL of 100 μg/mL stock solutions of GA_1, GM_1, and GT_{1b} gangliosides are pipetted onto each well and left to dry.

Microarray Surface Characterization: The biochips were further characterized at various fabrication steps via AFM. The SPR biochips were examined during each functionalization step: first silica gold (Au), then PFDTS silica Au, and finally GT_{1b} ganglioside functionalized onto the PFDTS silica Au biochip. GT_{1b} ganglioside was used as a representative of the gangliosides to confirm that they form evenly onto the surface. Contact angle measurements were used to determine the hydrophobicity of the surfaces at various fabrication steps.[32] All images were taken at ambient temperature and pressure. MALDI-MS instrument was used to confirm the presence and identity of gangliosides on the array after functionalization. A MALDI matrix of 1 mL CHCA was prepared with 50 % acetonitrile, 49 % DI H_2O, and 1 % trifluoroacetic acid. For the SPRi multiplexed biochips, 1 μL of matrix solution was pipetted into each well and allowed to dry. Chips were mounted on a steel MALDI plate and analyzed with an AB SCIEX TOF/TOF 5800 spectrometer operating in negative ion reflector mode at an intensity of 5000 A.U.

SPR and SPRi analysis: A NanoSPR5-321 (NanoSPR, Chicago, IL), a dual-channel SPR spectrometer with a GaAs semiconductor laser light source set at a wavelength of 670 nm, was used for all spectroscopic measurements for conventional SPR biosensing. The device utilizes a prism with refractive index of n=1.61 and a 30 μL flow

cell. PBS (phosphate buffered saline) running buffer at a pH of 7.4 was used in all experiments with a flow speed of 5 mL/h. Solutions of anti-GA_1, anti-GM_1, and anti-GT_{1b} were diluted with 1 x PBS to various concentrations for SPR experiments. After injection into the devices, solutions were incubated for 30 minutes before rinsing. In serum experiments anti-GT_{1b} was spiked into 10 % human serum diluted with PBS.

SPRi measurement was conducted on a home-built setup and a detailed description of which can be found in our previous work.[33] The functionalized substrates were mounted onto an optical stage that houses a PDMS flow cell. The array was placed in contact with an equilateral SF2 prism (n = 1.65) with a layer of refractive index matching fluid (Cargill Laboratories, Cedar Grove, NJ). A 648 nm light emitting diode (LED) was used as the light source for SPR excitation. Reflected images of the microarray were captured by a cooled 12- bit CCD camera (QImaging Retiga 1300) and data acquisition was controlled via a home built LabView program. Intensity data was normalized by using the intensity from the *p*-polarized light over the *s*-polarized beam and described as a percentage.

Statistical analysis: Analysis of variance (ANOVA) was conducted in Excel with the Analysis ToolPak add-in and used the end point data obtained with SPRi. Partial least squares discrimination analysis (PLS-DA) plots were generated with MetaboAnalyst platform and used the raw data from the SPRi sensorgrams for analysis. Principal component analysis (PCA) was completed with the prcomp function in R and graphed with the ggbiplot package with an ellipse probability set to 95 % using the end point data from the SPRi ganglioside microarray data.

2.3 Results and Discussion:

Biochip Surface Functionalization and Characterization: Perfluorinated hydrocarbon monolayers have been reported as an attractive surface for detection of analytes in complex media due to their near superhydrophobic and antifouling properties.[34] The monolayer can interact strongly with hydrophobic tags or moieties that are attached to the probes, presenting the antigenic sensing site in a well-organized and easy to access manner.[34] This works ideally for ganglioside immobilization and formation of ganglioside microarrays (Figure 2.2). We generated the PFDTS monolayer on a silica coating of the gold SPR substrate, where the silica nanofilm can also improve plasmonic properties for binding interactions.[35] Confirmation of ganglioside immobilization on the biochips was achieved by contact angle measurement, AFM, and MALDI-MS. Contact angle measurement showed that the pristine silica surface had a contact angle of 64.0 ± 2.0 degrees, indicating that the surface is rather hydrophilic. After PFDTS coating, the contact angle drastically increased to 131.1 ± 2.5 degrees (Figure 2.2); this value is close to what is accepted as being a superhydrophobic surface.[33] Three different gangliosides were used to make the array as shown in Figure 2.2: GA_1, GM_1, and GT_{1b}. After incubating, the contact angles for the ganglioside functionalized surfaces were 125.9 ± 3.1, 125.9 ± 2.3, and 124.9 ± 3.9 degrees, respectively. Images of the contact angles for the four surfaces can be found in the (Figure 2.2) The results also showed that the functionalized PFDTS substrate maintained hydrophobicity after the ganglioside deposition. This is important to the sensing in complex media as highly hydrophobic surfaces have been shown to play significant roles in antifouling techniques.[30, 36]

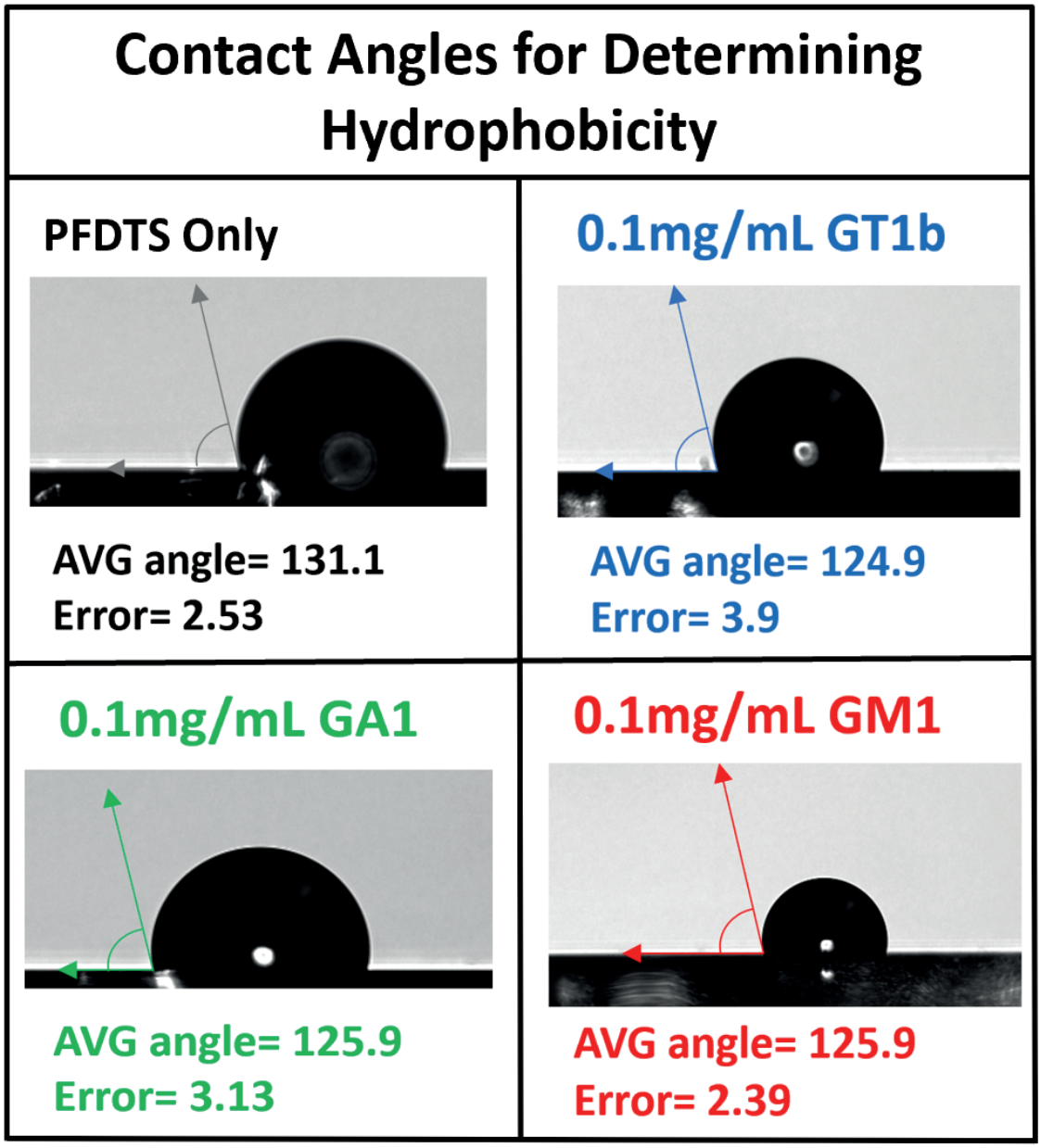


Figure 2.2. Contact angles for PFDTS surface and 0.1 mg/mL ganglioside surface. Shows near super hydrophobic angle of PFDTS and the surface remains hydrophobic after functionalization with the three gangliosides. It can be seen that even after the gangliosides are functionalized on the PFDTS surface the hydrophobicity is only slightly decreased and there is very little deviation between the three gangliosides.

MALDI-MS was conducted to confirm that ganglioside functionalization onto the PFDTS surface was successful and remained after washing under microfluidic conditions.

The MS spectra for GA_1, GM_1, and GT_{1b} gangliosides are shown in Figure 3.3A-C respectively. Peaks used for the confirmation that GT_{1b} ganglioside was on the surface were 2158.91 m/z, 2129.98 m/z, and 2108.92 m/z.[37, 38] During the ionization process it is possible for $GT1_b$ ganglioside to fragment and lose its sialic acid (SA) groups.[37] The fragmentation of GT_{1b} splits it into one SA and either $GD1_a$ or $GD1_b$, which are both associated with the peaks at 1863.82 m/z and 1835.77 m/z.[37, 38] Because $GD1_a$ and $GD1_b$ gangliosides both have the same mass and the difference between the two is the location of the SA groups, they cannot be differentiated by the presented spectra. Fragmentation continues with the loss of a second SA groups resulting in GM_1 associated peaks 1572.77 m/z and 1544.75 m/z being present.[37, 38] The loss of the third and final SA group from the fragmented GT_{1b} ganglioside results in the peaks associated with GA_1 ganglioside being present which are 1270.30 m/z and 1252.29 m/z.[38] The collected results agree well with the spectra found in literature. GT_{1b} associated peaks range from 2150 to 2250 m/z, $GD1_a$ and $GD1_b$ range between 1800 to 1900 m/z, GM_1 confirmed peaks are from 1500 to 1600 m/z, and GA_1 associated peaks range between 1200 to 1300 m/z. [37-41]

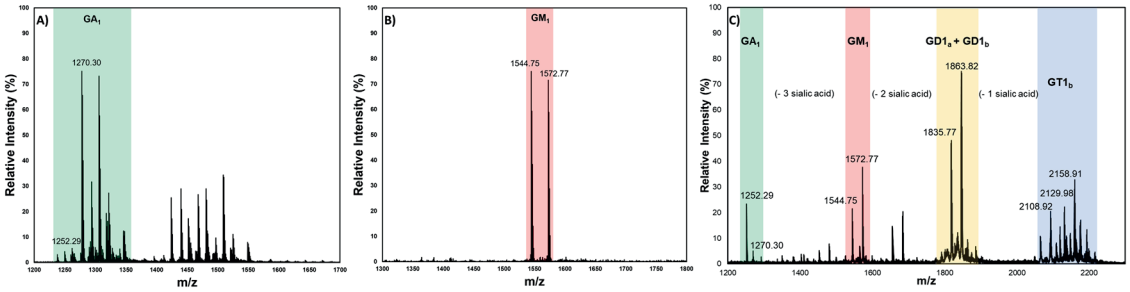


Figure 2.3. MALDI-TOF-MS spectra of A) 0.1 mg/mL GA_1 ganglioside green, B) 0.1 mg/mL GM_1 ganglioside red, and C) 0.1 mg/mL GT_{1b} ganglioside blue on the PFDTS Silica Au SPRi microarray chip after being exposed to microfluidic conditions. The peaks associated with GA_1 are present 1252.29 and 1270.30 m/z in both A) for GA_1 and the fragmented GT_{1b} ganglioside in C). The associated peaks for GM_1 are 1544.75 and 1572.77 m/z which are both present in B) for GM_1 and C) for GT_{1b}. The fragmentation of GT_{1b} in C) also gave rise to peaks associated with GD_{1a} and GD_{1b} as is highlighted in yellow.

AFM was also utilized to characterize the surface at various fabrication steps: silica gold chip, PFDTS monolayer on silica gold, and ganglioside on PFDTS on the silica gold chip (Figure 2.4A-F). Surface roughness was measured for each fabrication step to ensure no clumping had occurred, which could potentially compromise the sensitivity on both SPR and SPRi. GT_{1b}, the largest ganglioside in the group, was used as a representative ganglioside for array fabrication characterization. The RMS for the silica gold surface, PFDTS monolayer on the silica gold surface, and GT_{1b} ganglioside functionalized PFDTS monolayer were 1.63 nm, 10.47 nm, and 9.34 nm, respectively. The deposited layers did not appear to agglomerate in specific regions on the chip at any of the fabrication steps, and that the surfaces were in fact smooth. While there is a moderate increase in surface roughness for the functionalized surfaces, the values still fall well below the decay length of the SPR evanescent field, suggesting that the PFDTS functionalization has little effect on sensing performance.[42]

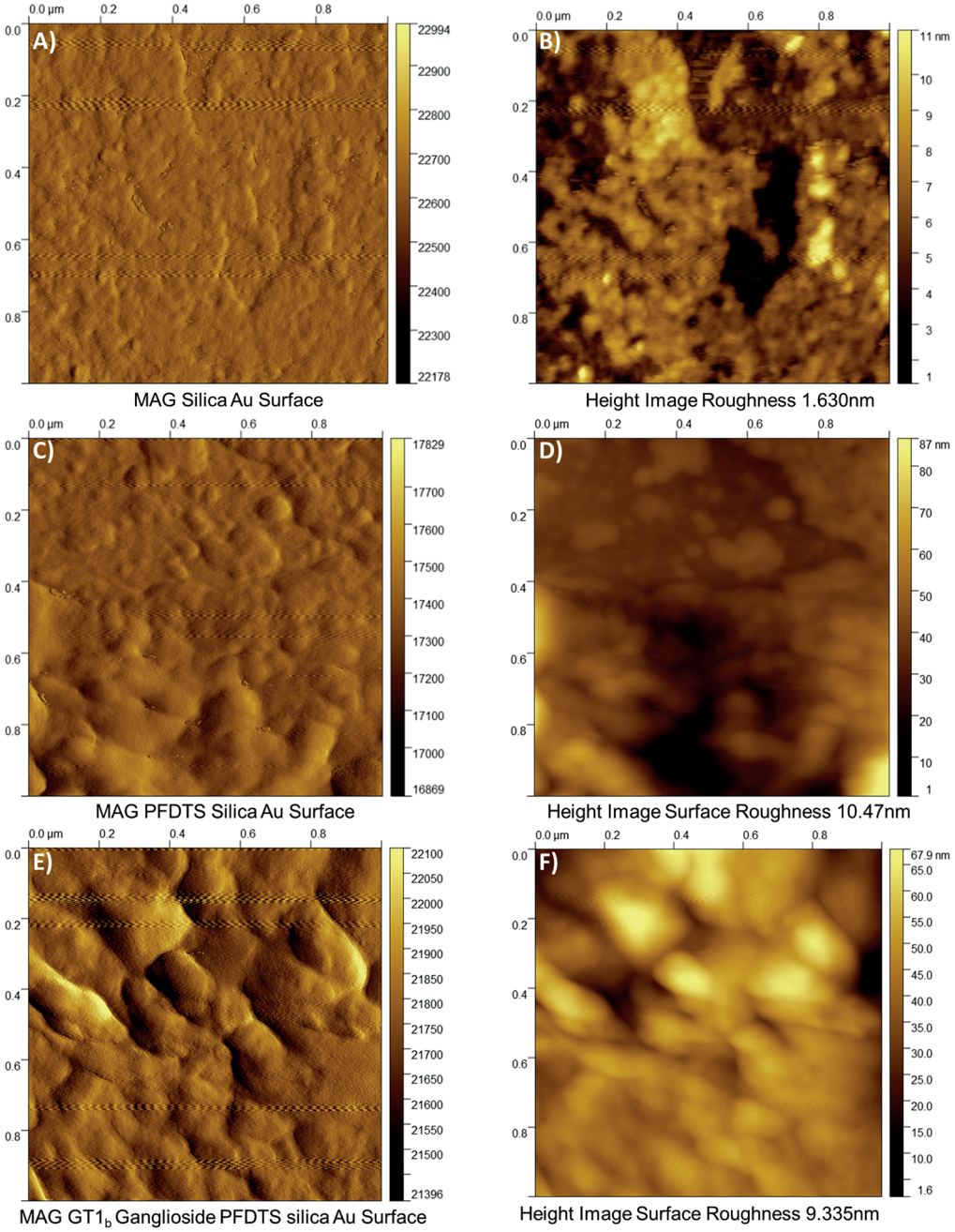


Figure 2.4. A) AFM image and B) height map to show surface roughness of silica gold surface, C) AFM image and D) height map to show surface roughness of PFDTS silica gold surface, E) AFM image and F) heigh map to show surface roughness of for 0.1 mg/mL GT_{1b} ganglioside functionalized onto the PFDTS silica Au surface. While the surfaces did get slightly rougher after each functionalization step they were still very smooth and there appears to be no agglomeration occurring as the roughness did not decrease the sensitivity of either conventional SPR or SPRi biosensing experiments.

Capture of Antibodies on Ganglioside Functional Surfaces by Spectroscopic SPR: Conventional SPR biosensing was initially used to characterize the antibody binding to the functionalized biochips before attempting multiplexed detection on an SPRi instrument. All experiments were conducted with antibodies diluted in PBS to measure specific binding on their respective surfaces, as well as nonspecific binding and cross reactivity between various antibodies. To account for cross reactivity between the ganglioside surfaces and anti- ganglioside specific antibodies, the antibody that showed the highest cross reactive binding was used as a control. Figure 2.5A shows the sensorgrams where 50 ng/mL of anti-GM_1 and anti-GA_1 were injected onto a 0.1 mg/mL GA_1 ganglioside surface. This process was also performed on the GT_{1b} ganglioside surface with 10 ng/mL of anti-GT_{1b} and 10 ng/mL of anti-GA_1 (Figure 2.5B). In both cases, a baseline was quickly established, and the control antibody had little to no binding to the surface. This is in direct contrast to the specific antibody/ganglioside pair (Figure 2.5 A and B lower curves), which shows a clear angular shift, even after the rinse cycle.

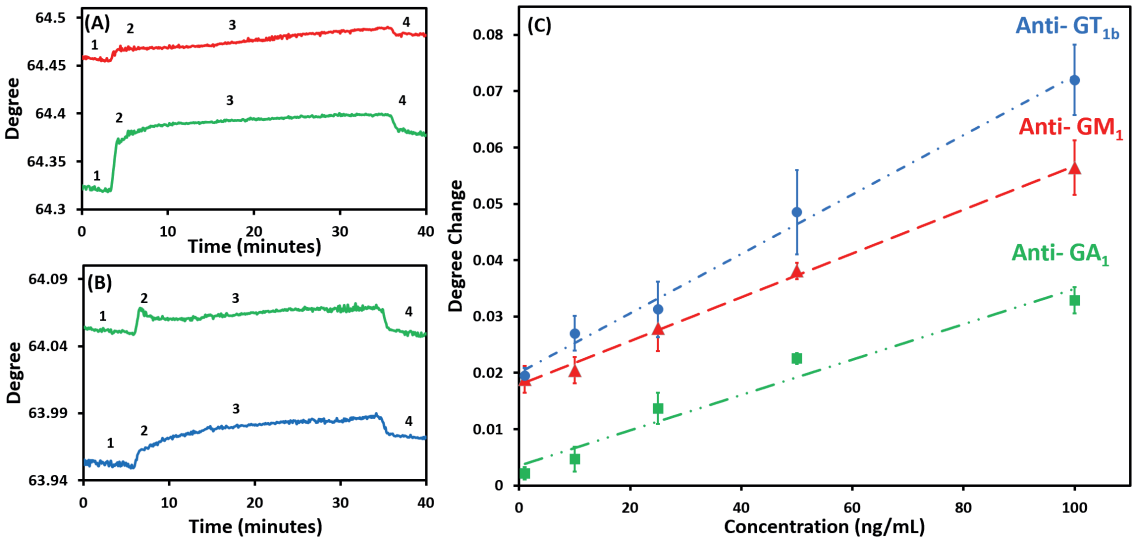


Figure 2.5. Sensorgrams for (A) 50 ng/mL anti-GM_1 (Red) and 50 ng/mL anti-GA_1 (Green) in PBS on a 0.1 mg/mL GA_1 ganglioside PFDTS functionalized surface, and (B) 10 ng/mL anti-GT_{1b} (Blue) and 10 ng/mL anti-GA_1 (Green) in PBS on a GT_{1b} ganglioside PFDTS functionalized surface. Numbers indicate experimental conditions/actions for 1) Baseline, 2) Injection of antibody, 3) Incubation, and 4) Rinse. (C) calibration curves of anti-GT_{1b} (Blue), anti-GM_1 (Red), and anti-GA_1 (Green) in PBS.

A calibration curve for the three antibodies was generated at concentrations ranging from 1-100 ng/mL, as shown in Figure 2.5C. From the sensorgrams, the specific binding of 10 ng/mL of anti-GT_{1b} on a GT_{1b} ganglioside surface has roughly the same specific binding as a 25 ng/mL anti-GM_1 on a GM_1 ganglioside surface. The reported LOD, calculated by using the 3σ method, were 17.6 ng/mL for anti-GA_1, 11.3 ng/mL for anti-GM_1, and 8.2 ng/mL for anti-GT_{1b}, respectively. The stronger binding interactions and lower LOD observed for anti-GT_{1b} can be attributed to the large headgroup (more sialic acids) of GT_{1b} ganglioside.[43] We thus chose anti-GT_{1b} as the model biomarker for the assessment of MS specific antibodies in serum at biologically relevant concentrations using SPRi.

SPR Imaging Detection of Anti-GT_{1b} in Serum

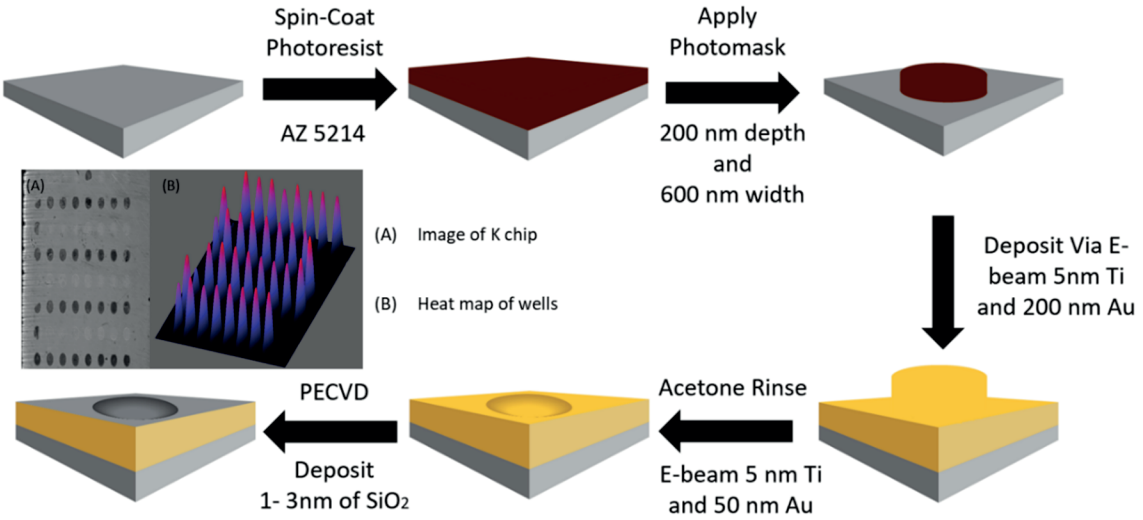


Figure 2.6. A cartoon illustration of the updated fabrication procedure for SPRi microarray chips. The first step is to apply a photoresist onto clean glass microchip slides after which a pattern is created via photolithographic procedures. The surface is then coated with a thin layer of titanium to help the gold adhere to the surface, the titanium gold on top of the photomask are than washed away with acetone creating the desired wells. Once all the wells are washed off another layer of titanium and gold is deposited onto the surface allowing the creation of the microwells that make up the microarray. The last step is to coat the surface with a thin layer of silica oxide. (A) Image of final K chip product and (B) heat map showing each well is uniformed except for minor fluctuations where the S shaped flow cell changes directions.

SPRi allows interrogation of multiple elements simultaneously, providing real time analysis with the additional benefit of high throughput and multiplexed detection.[28] A slightly modified fabrication procedure for the microarrays developed in our lab,[31, 44] is shown in Figure 2.6. These modifications aimed to improve the reproducibility of the arrays and attenuating background plasmonics. The benefits of the gold silicated microwells include unique tunable plasmonic properties,[45] and facilitated background free image analysis for enhanced sensitivity and contrast.[28, 31, 46] Through the use of

alkyltrichlorosilane-based procedure for surface modification,[34] a monolayer of PFDTS and gangliosides were generated, allowing for SPRi detection of MS related antibodies in 10 % serum (Figure 2.7). The PFDTS coating yielded a relatively ordered surface to enable strong hydrophobic-hydrophobic interactions, and the coating does not compromise surface smoothness and thus the performance, as discussed previously.

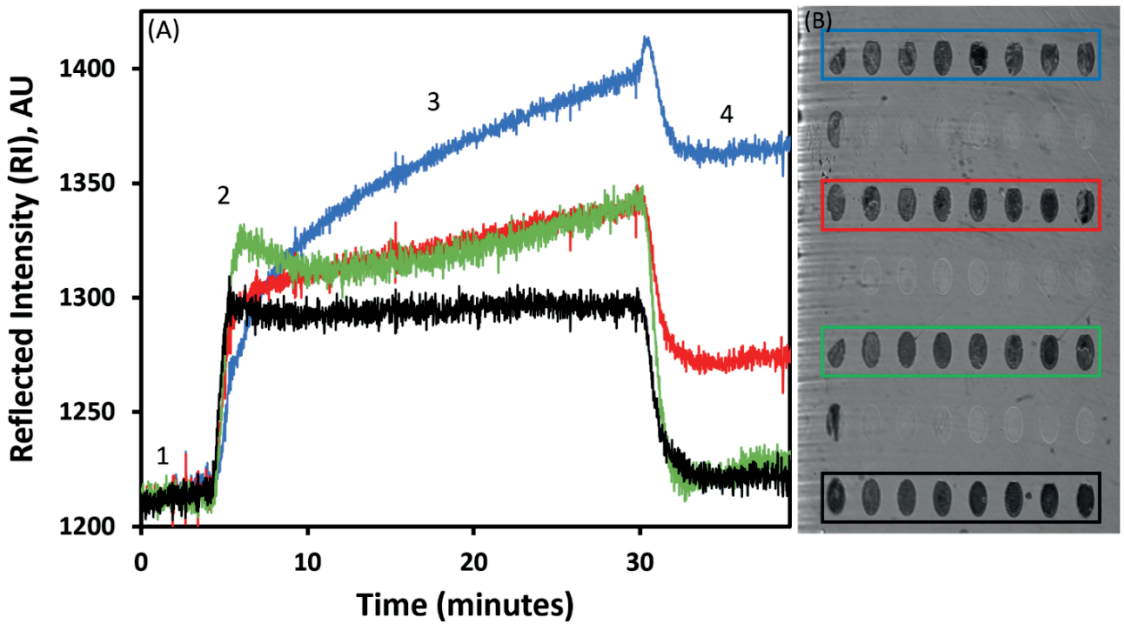


Figure 2.7. (A) SPRi sensorgrams showing shifts caused by injecting 100 ng/mL anti-GT_{1b} in 10 % serum on $GT1_b$ ganglioside surface (Blue), GM_1 ganglioside surface (Red), GA_1 ganglioside surface (Green), and PFDTS surface (Black). Number indicate experimental conditions/procedure: 1) Rinse after 10 % serum, 2) Inject 100 ng/mL anti-GT_{1b} in 10 % serum, 3) Incubation, and 4) Rinse. (B) SPR image of the ganglioside arrays. Color indicates functionalized surface in each row.

Using a loop flow cell, 4 working channels were defined in the microarray analysis; each channel having eight active wells for statistical analysis, as shown in Figure 2.7. There was an unfunctionalized channel (framed in black) that was used as a reference point to account for chip to chip variation. The response of anti-GT_{1b} on the ganglioside array, when specifically detecting anti-GT_{1b}, showed a low signal of nonspecific binding in the diluted

serum. The collected results show excellent agreement with the collected data on spectroscopic SPR (Figure 2.5) and SPRi biosensing in other buffering conditions such as PBS and 2.0 mg/mL BSA (Figure 2.9).

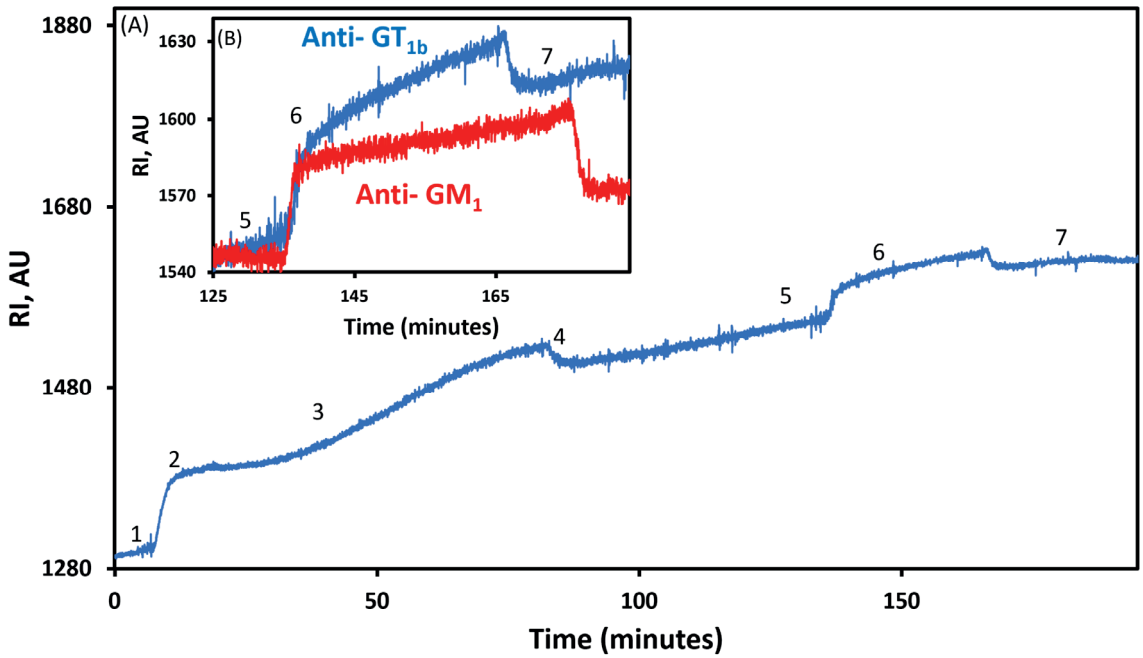


Figure 2.8. (A) SPRi spectrogram of 100 ng/mL anti-GT_{1b} in 10 % serum on a PFDTS ganglioside surface (1. Baseline, 2. Injection of 10 % serum, 3. Incubation, 4. Rinse, 5. Injection of 100 ng/mL anti-GT_{1b} (Blue) and 100 ng/mL anti-GM_1 (Red) in 10 % Serum, 6. Incubation, and 7. Rinse.) (B) Cross reactivity characterized by 100 ng/mL anti-GM_1 (Red) on a GT_{1b} ganglioside functionalized surface compared to shift caused by anti-GT_{1b} (Blue).

Nonspecific binding from complex media onto the sensing interface for SPR and SPRi is arguably the major factor that hinders wide utilization of SPR detection methods as a routine diagnostic tool.[46, 47] To further characterize the nonspecific binding on the PFDTS layer, we saturated the surface with 10 % serum before antibody injection and evaluated the sensing performance (Figure 2.8A). The serum acts as an antifouling agent by blocking nonspecific binding locations that may become occupied when the anti-GT_{1b} spiked serum samples are introduced. 10 % serum was selected to saturate the surface as

this was the same dilution that the antibodies were spiked into and 10 % serum has been used to represent complex media in literature.[48] The effectiveness of using the same concentration to saturate the surface as the concentration the antibodies were spiked into was first observed when investigating the spiked 2.0 mg/mL BSA samples. To characterize specific detection of anti-GT_{1b} and monitor cross reactivity, we tested antibodies at various concentrations (1 ng/mL to 100 ng/mL) and compared the response. A typical sensorgram is shown in Figure 2.8B, which shows low cross reactivity of anti-GM_1 on a GT_{1b} ganglioside surface. The shift caused by 100 ng/mL anti-GT_{1b} was much larger than that by 100 ng/mL anti-GM_1 on a $GT1_{1b}$ ganglioside surface, indicating specific binding and also some level of cross reactivity. After rinsing the surface with PBS, the signal for anti-GT_{1b} was again higher than for anti-GM_1, which agrees with the results on spectroscopic SPR.

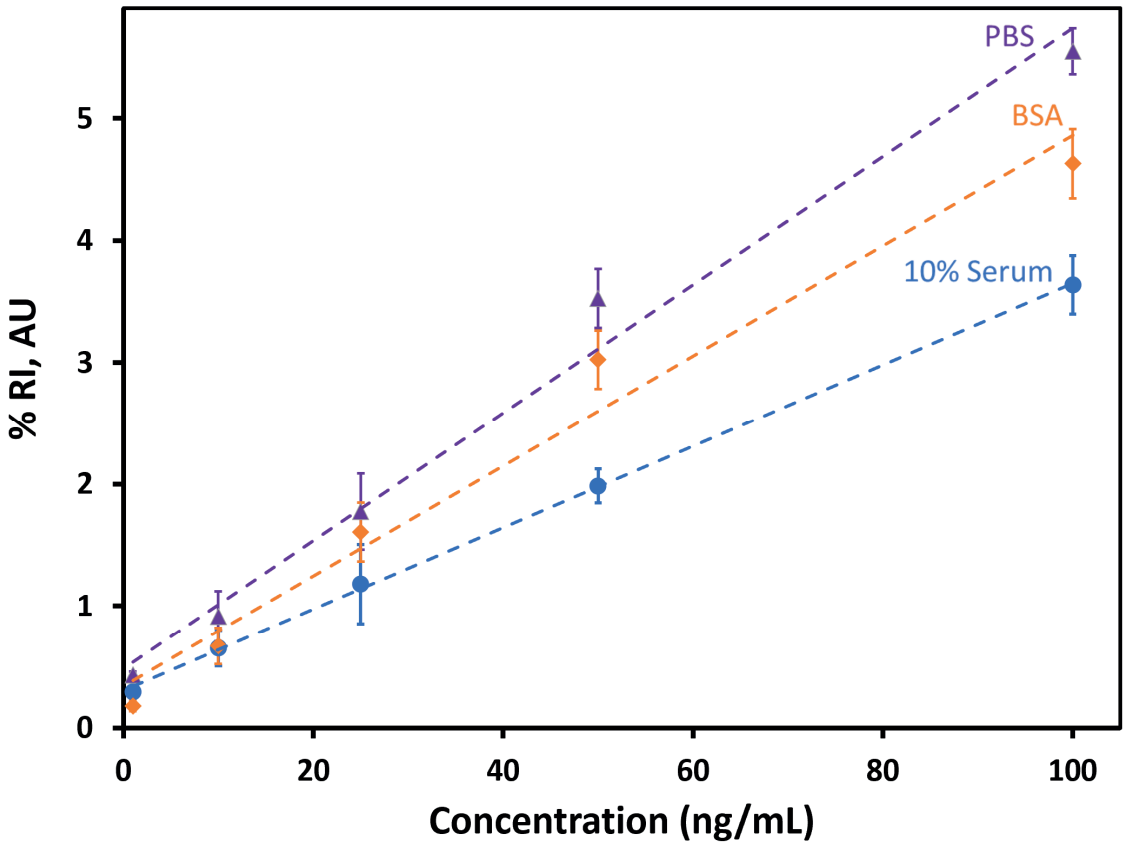


Figure 2.9. Calibration curves for the specific binding between GT_{1b} ganglioside functionalized substrate and anti-GT_{1b} spiked in PBS at pH of 7.4 (purple), 2.0 mg/mL BSA (orange), and 10 % diluted serum (blue) at below (1 ng/mL) and above (20 ng/mL) MS related antibodies concentrations.

A calibration curve for anti-GT_{1b} under different experimental conditions was generated using SPRi data, as shown in Figure 2.9. The sensing surface has a working range of 1 ng/mL to 100 ng/mL. The LOD for anti-Gt_{1b} in 10 serum was calculated to be 2.34 ng/mL using the 3σ principle. The results showed that this ganglioside microarray is capable of detecting MS antibodies in serum at clinically relevant levels, which are present in patient serum at concentrations between 3 ng/mL and 25 ng/mL.

Multiplexed Detection and Statistical Analysis of Ganglioside Antibodies in Serum: Precise detection of various antibodies for MS-related gangliosides in serum is a

complex process due to existence of cross reactions. We used hydrocarbon microarrays to investigate the cross reactivity with the multiplexed detection of 100 ng/mL for anti-GT_{1b}, anti-GM_1, and anti-GA1 spiked into 10 % human serum (Figure 2.10). The data presented was the average of at least 5 samples, and 0.1 mg/mL of GA_1, GM_1, and GT_{1b} ganglioside were used for microarray fabrication. The shifts caused by 100 ng/mL anti-GT_{1b} were 0.66 ± 0.23 % RIU on GA_1 surface, 1.21 ± 0.10 % RIU on GM_1 surface, and 4.62 ± 0.24 % RIU on GT_{1b} surface. The results indicate that anti-GT_{1b} shows the largest response for its ganglioside, indicating the selectivity of the antibody and also suggesting elevated binding strength due to a large headgroup.[43] The shifts caused by 100 ng/mL anti-GM_1 on the GA_1, GM_1, and $GT1_b$ microarray were 1.62 ± 0.21 % RIU, 3.35 ± 0.17 % RIU, and 0.98 ± 0.23 % RIU respectively, while 100 ng/mL of anti-GA_1 caused a shift that resulted in a 3.32 ± 0.17 % RIU for GA_1, 1.25 ± 0.23 % RIU for GM_1, and 0.58 ± 0.17 % RIU for GT_{1b}. Using the lowest binding as a correction for nonspecific binding, we calculated the specific binding for the three antibodies in spiked serum to be 3.64 ± 0.24 % RIU, 2.01 ± 0.17 % RIU, and 1.70 ± 0.17 % RIU for anti-GT_{1b}, anti-GM_1, and anti-GA_1, respectively. The cross reactivity between anti-GM_1 and anti-GA_1 can be attributed to the similar structure of the two gangliosides, where the only difference is that GM_1 has a SA group while GA_1 does not.[49]

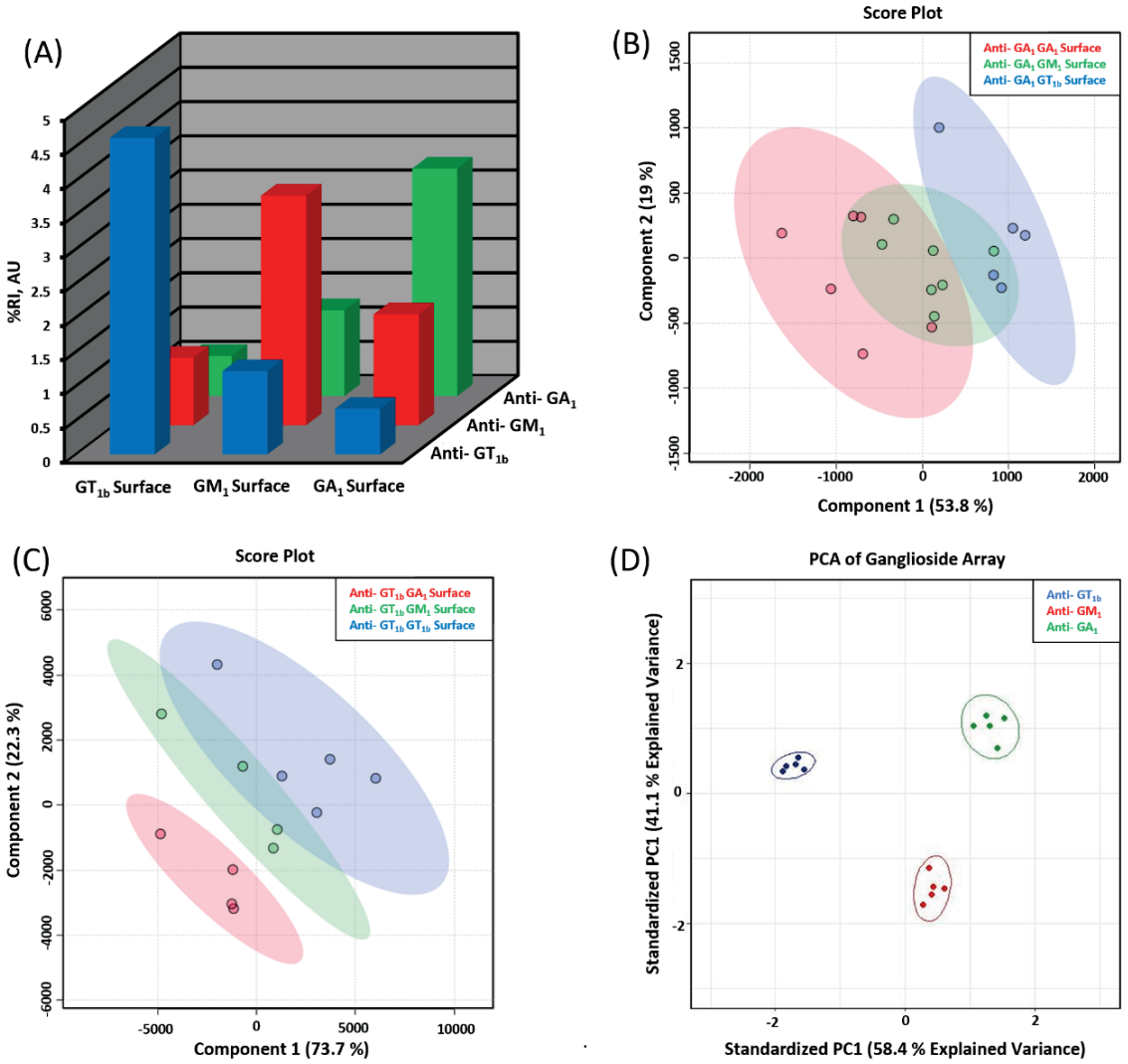


Figure 2.10. (A) Bulk changes caused by MS specific antibodies at 100 ng/mL spiked in 10 % serum on a high throughput multiplexed SPRi microarray. (Blue) % RI, AU caused by 100 ng/mL anti-GT_{1b} on 0.1 mg/mL GT_{1b}, GM_1, and GA_1 ganglioside surfaces, (Red) binding between 100 ng/mL anti-GM_1 and 0.1 mg/mL on a GT_{1b}, GM_1, and GA_1 ganglioside surfaces, and (Green) binding between 100 ng/mL anti-GA_1 and 0.1 mg/mL GT_{1b}, GM_1, and GA_1 gang100lioside surfaces. (B) Partial Least Squares Discriminant Analysis (PLS- DA) for 100 ng/mL anti-GA_1 on the three PFDTS ganglioside functionalized surfaces. (Red) anti-GA_1 on GA_1 ganglioside surface, (Green) anti-GA_1 on GM_1 ganglioside surface, and (Blue) anti-GA_1 on GT_{1b} ganglioside surface. (C) PLS- DA for anti-GT_{1b} on (Red) GA_1 ganglioside surface, (Green) GM_1 ganglioside surface, and (Blue) GT_{1b} ganglioside surface. (D) Principal components analysis (PCA) showing the ability to separate the three anti- ganglioside antibodies based on their induced response across the whole ganglioside microarray. (Blue) 100 ng/mL anti-GT_{1b}, (Red) 100 ng/mL anti-GM_1, and (Green) 100 ng/mL anti-GA_1.

To further evaluate the cross reactivity and differentiate antibody-ganglioside interactions, we conducted several statistical analyses of the microarray data including analysis of variance (ANOVA), partial least squares discriminant analysis (PLS-DA), and principal component analysis (PCA). ANOVA, Table 2.1, indicated antibody interactions with different ganglioside surfaces are statistically relevant with calculated p-values ≤ 0.05. The variance was found to vary between 0.06 and 0.03, with anti-GT_{1b}/GT_{1b} pair having the lowest variance while anti-GA_1/GA_1 having the highest variance.

Table 2.1. ANOVA of 100 ng/mL anti-GT_{1b}, 100 ng/mL anti-GM_1, and 100 ng/mL anti-GA_1 on the GT_{1b}, GM_1, and GA_1 ganglioside PFDTS functionalized SPRi microarray substrate.

Summary of GT1b Surface

Groups	*Count*	*Sum*	*Average*	*Variance*
100 ng/mL Anti-GT1b	5	23.097636	4.6195272	0.010248902
100 ng/mL Anti-GM1	5	4.90522	0.981044	0.027379828
100 ng/mL Anti-GA1	5	2.924113	0.5848226	0.052614597

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	49.45732101	2	24.7286605	535.4976162	1.85E-12	3.88529384
Within Groups	0.554146119	12	0.046178843			

Summary of Gm1 Surface

Groups	*Count*	*Sum*	*Average*	*Variance*
100 ng/mL anti-GT1b	5	6.066162	1.2132324	0.010248902
100 ng/mL Anti-GM1	5	16.748228	3.3496456	0.028826417
100 ng/mL Anti-GA1	5	6.2613254	1.25226508	0.058542104

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	14.94131659	2	7.470658297	246.1298534	1.82E-10	3.88529384
Within Groups	0.364230094	12	0.030352508			
Total	15.30554669	14				

Summary of GA1 Surface

Groups	*Count*	*Sum*	*Average*	*Variance*
100 ng/mL anti-GT1b	5	5.476465	1.095293	0.010209662
100 ng/mL Anti-GM1	5	8.085184	1.6170368	0.045632456
100 ng/mL Anti-GA1	5	16.59633	3.319266	0.027097434

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	13.52642791	2	6.763213956	244.6316791	1.88E-10	3.88529384
Within Groups	0.331758208	12	0.027646517			
Total	13.85818612	14				

PLS-DA was conducted using longer range of data from the SPRi sensorgrams, which allowed for evaluation of the binding interactions that include both kinetic and steady state components. PLS-DA results indicated marked differences for anti-GT_{1b} and

anti-GA_1 as the gangliosides differ the most. Figure 2.10B shows the score plot for anti-GA_1 binding associations with each ganglioside on the array. There was a substantial overlap between anti-GA_1 and GM_1 ganglioside, whereas no overlap was observed between anti-GA_1 and GT_{1b} ganglioside. Interactions of anti-GT_{1b} with various gangliosides on the array is shown in Figure 2.10C. As expected, anti-GT_{1b} and GM_1 ganglioside had some cross reactivity while no overlap was present for GA_1 ganglioside. The cross reactivity observed agrees well with data collected on conventional SPR (Figure 2.5 A-B) and on SPRi (Figure 2.7A, 2.8A, and 2.8B), and can be attributed to the presence of SA group(s). The larger structural difference between GT_{1b} and GA_1 decreases the cross reactivity, while this effect appears to be reversed for similar structures of GM_1 and GA_1. The analysis with PLS-DA using binding kinetic features allows determination of the extent of cross reactivity and selectivity of the antibodies. This approach may provide highly useful information on the nature of the immune interactions between structurally similar antigens.

In addition, PCA analysis was conducted based on end point values for each array component, as shown in Figure 2.10D. A clear separation between anti-GT_{1b}, anti-GM_1, and anti-GA_1 regions is obtained, which agrees well with the calibration curves in Figures 2.9 and 2.5C. This ability to identify specific antibody/ganglioside interaction is very significant, because it shows the microarray method is powerful for differentiating multiple antibodies in a single experiment, even with an extensive degree of cross reactivity.

These statistical analyses provide convincing evidence that the functionalized PFDTS ganglioside microarray is capable of probing the complex cross reactivity network between the MS specific antibodies, and is thus able to provide a wide range of information

compared to traditional end point assays. The effectiveness of the microarray can be attributed to the near superhydrophobic properties of the PFDTS surface, the carbohydrate head groups of the probes, and the added surface protection techniques. The array can also allow for high throughput antibody screening for disease diagnosis in patient samples, which may lead to faster diagnosis and more effective therapies.

2.4 Conclusion:

We have demonstrated a new and effective platform to screen blood samples for potential diagnosis of multiple sclerosis by assessing ganglioside antibody interactions and quantifications. The SPRi microarray biochip has shown to be able to detect and differentiate between MS specific antibodies at biological relevant concentrations. This method is an important step towards MS diagnosis as the complexity of MS is partially attributed to the clinical observation of high individual heterogeneity that certain antibodies present in one patient may not be expressed in another, and the concentrations may vary wildly. Therefore, detecting a broad range of antibodies with multiplexed capability and under identical assay conditions is critical. The ganglioside microarray can be easily extended to include other important glycolipids and sphingomyelins from the myelin sheath, which will allow for a broad-spectrum profiling of the patients' samples. Using anti-GT_{1b} as an example, we showed this microarray can detect biomarkers within the range of disease relevant concentrations of 3 ng/mL to 25 ng/mL. Statistical analysis presented here using ANOVA, PLS-DA, and PCA allowed for the analysis of features in the SPRi sensorgrams and the endpoint data, revealing unique characteristics that can be used for identification of specific analyte/antigen interactions. We believe the work

presented here has the potential to improve disease diagnosis, enhance the evaluation of disease progression, and may lead to improvements in drug development that aims at blocking antigen binding regions and inhibiting the immune system from targeting components of the CNS. Future work will expand to cover more immuno biomarkers in whole serum and inclusion of signal amplification for detecting those potential biomarkers with extremely low abundance.

2.5 References:

1. Goldenberg, M. M., Multiple sclerosis review. *Pharmacy and Therapeutics* **2012,** *37* (3), 175.

2. Organization, W. H., Atlas: multiple sclerosis resources in the world 2008. **2008**.

3. Compston, A.; Coles, A., Multiple sclerosis. *Lancet* **2002,** *359* (9313), 1221-31.

4. Polman, C. H.; Reingold, S. C.; Edan, G.; Filippi, M.; Hartung, H. P.; Kappos, L.; Lublin, F. D.; Metz, L. M.; McFarland, H. F.; O'Connor, P. W.; Sandberg-Wollheim, M.; Thompson, A. J.; Weinshenker, B. G.; Wolinsky, J. S., Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". *Ann Neurol* **2005,** *58* (6), 840-6.

5. Brady, S.; Siegel, G.; Albers, R. W.; Price, D., *Basic neurochemistry: molecular, cellular and medical aspects*. Elsevier: 2005.

6. Schnaar, R. L., Brain gangliosides in axon–myelin stability and axon regeneration. *FEBS letters* **2010,** *584* (9), 1741-1747.

7. Koutsouraki, E.; Hatzifilippou, E.; Cosetsidou, C.; Avdelidi, E.; Banaki, T.; Costa, V.; Baloyannis, S., Anti-GM1 antibodies in multiple sclerosis patients: P2674. *European Journal of Neurology* **2008,** *15*.

8. Reindl, M.; Khalil, M.; Berger, T., Antibodies as biological markers for pathophysiological processes in MS. *Journal of neuroimmunology* **2006,** *180* (1-2), 50-62.

9. Hogan, E. L.; Podbielska, M.; O'Keeffe, J., Implications of lymphocyte anergy to glycolipids in multiple sclerosis (MS): iNKT cells may mediate the MS infectious trigger. *Journal of clinical & cellular immunology* **2013,** *4* (3).

10. Marconi, S.; Acler, M.; Lovato, L.; De Toni, L.; Tedeschi, E.; Anghileri, E.; Romito, S.; Cordioli, C.; Bonetti, B., Anti-GD2-like IgM autoreactivity in multiple sclerosis patients. *Multiple Sclerosis Journal* **2006,** *12* (3), 302-308.

11. Pender, M. P.; Csurhes, P. A.; Wolfe, N. P.; Hooper, K. D.; Good, M. F.; McCombe, P. A.; Greer, J. M., Increased circulating T cell reactivity to GM3 and GQ1b gangliosides in primary progressive multiple sclerosis. *Journal of clinical neuroscience* **2003,** *10* (1), 63-66.

12. Mizutani, K.; Oka, N.; Kusunoki, S.; Kaji, R.; Mezaki, T.; Akiguchi, I.; Shibasaki, H., Sensorimotor demyelinating neuropathy with IgM antibody against

gangliosides GD1a, GT1b and GM3. *Journal of the neurological sciences* **2001,** *188* (1-2), 9-11.

13. Brownlee, W. J.; Hardy, T. A.; Fazekas, F.; Miller, D. H., Diagnosis of multiple sclerosis: progress and challenges. *The Lancet* **2017,** *389* (10076), 1336-1346.

14. Traboulsee, A.; Simon, J.; Stone, L.; Fisher, E.; Jones, D.; Malhotra, A.; Newsome, S.; Oh, J.; Reich, D.; Richert, N., Revised recommendations of the consortium of MS centers task force for a standardized MRI protocol and clinical guidelines for the diagnosis and follow-up of multiple sclerosis. *American Journal of Neuroradiology* **2016,** *37* (3), 394-401.

15. Laufs, H.; Daunizeau, J.; Carmichael, D. W.; Kleinschmidt, A., Recent advances in recording electrophysiological data simultaneously with magnetic resonance imaging. *Neuroimage* **2008,** *40* (2), 515-528.

16. McDonald, W. I.; Compston, A.; Edan, G.; Goodkin, D.; Hartung, H. P.; Lublin, F. D.; McFarland, H. F.; Paty, D. W.; Polman, C. H.; Reingold, S. C., Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society* **2001,** *50* (1), 121-127.

17. Solomon, A. J.; Bourdette, D. N.; Cross, A. H.; Applebee, A.; Skidd, P. M.; Howard, D. B.; Spain, R. I.; Cameron, M. H.; Kim, E.; Mass, M. K., The contemporary spectrum of multiple sclerosis misdiagnosis: a multicenter study. *Neurology* **2016,** *87* (13), 1393-1399.

18. Robinson, W. H.; DiGennaro, C.; Hueber, W.; Haab, B. B.; Kamachi, M.; Dean, E. J.; Fournel, S.; Fong, D.; Genovese, M. C.; De Vegvar, H. E. N., Autoantigen microarrays for multiplex characterization of autoantibody responses. *Nature medicine* **2002,** *8* (3), 295-301.

19. Derkus, B.; Bozkurt, P. A.; Tulu, M.; Emregul, K. C.; Yucesan, C.; Emregul, E., Simultaneous quantification of Myelin Basic Protein and Tau proteins in cerebrospinal fluid and serum of Multiple Sclerosis patients using nanoimmunosensor. *Biosensors and Bioelectronics* **2017,** *89*, 781-788.

20. Frohman, E.; Havrdova, E.; Lublin, F.; Barkhof, F.; Achiron, A.; Sharief, M.; Stuve, O.; Racke, M.; Steinman, L.; Weiner, H., Most patients with multiple sclerosis or a clinically isolated demyelinating syndrome should be treated at the time of diagnosis. *Archives of neurology* **2006,** *63* (4), 614-619.

21. Solomon, A. J.; Weinshenker, B. G., Misdiagnosis of multiple sclerosis: frequency, causes, effects, and prevention. *Current neurology and neuroscience reports* **2013,** *13* (12), 403.

22. Solomon, A. J.; Corboy, J. R., The tension between early diagnosis and misdiagnosis of multiple sclerosis. *Nature Reviews Neurology* **2017,** *13* (9), 567-572.

23. Solomon, A. J.; Klein, E. P.; Bourdette, D., "Undiagnosing" multiple sclerosis: the challenge of misdiagnosis in MS. *Neurology* **2012,** *78* (24), 1986-1991.

24. Boissy, A. R.; Ford, P. J., A touch of MS: therapeutic mislabeling. *Neurology* **2012,** *78* (24), 1981-1985.

25. Kodoyianni, V., Label-free analysis of biomolecular interactions using SPR imaging. *Biotechniques* **2011,** *50* (1), 32-40.

26. He, Y.; Yang, M.; Zhao, S.; Cong, C.; Li, X.; Cheng, X.; Yang, J.; Gao, D., Regulatory mechanism of localized surface plasmon resonance based on gold nanoparticles-coated paclitaxel nanoliposomes and their antitumor efficacy. *ACS Sustainable Chemistry & Engineering* **2018,** *6* (10), 13543-13550.

27. Kreysing, E.; Hassani, H.; Hampe, N.; Offenhäusser, A., Nanometer-Resolved Mapping of Cell–Substrate Distances of Contracting Cardiomyocytes Using Surface Plasmon Resonance Microscopy. *ACS nano* **2018,** *12* (9), 8934-8942.

28. Lambert, A.; Yang, Z.; Cheng, W.; Lu, Z.; Liu, Y.; Cheng, Q., Ultrasensitive detection of bacterial protein toxins on patterned microarray via surface plasmon resonance imaging with signal amplification by conjugate nanoparticle clusters. *ACS sensors* **2018,** *3* (9), 1639-1646.

29. Nguyen, H. H.; Park, J.; Kang, S.; Kim, M., Surface plasmon resonance: a versatile technique for biosensor applications. *Sensors* **2015,** *15* (5), 10481-10510.

30. Banerjee, I.; Pangule, R. C.; Kane, R. S., Antifouling coatings: recent developments in the design of surfaces that prevent fouling by proteins, bacteria, and marine organisms. *Advanced materials* **2011,** *23* (6), 690-718.

31. Abbas, A.; Linman, M. J.; Cheng, Q., Patterned resonance plasmonic microarrays for high-performance SPR imaging. *Analytical chemistry* **2011,** *83* (8), 3147-3152.

32. Duan, J.; Wang, H.; Cheng, Q. J. A. c., On-plate desalting and SALDI-MS analysis of peptides with hydrophobic silicate nanofilms on a gold substrate. **2010,** *82* (22), 9211-9220.

33. Xue, C.-H.; Ma, J.-Z., Long-lived superhydrophobic surfaces. *Journal of materials chemistry A* **2013,** *1* (13), 4146-4161.

34. Srinivasan, U.; Houston, M. R.; Howe, R. T.; Maboudian, R., Alkyltrichlorosilane-based self-assembled monolayer films for stiction reduction in silicon micromachines. *Journal of Microelectromechanical Systems* **1998,** *7* (2), 252-260.

35. Abbas, A.; Linman, M. J.; Cheng, Q., Sensitivity comparison of surface plasmon resonance and plasmon-waveguide resonance biosensors. *Sensors and Actuators B: Chemical* **2011,** *156* (1), 169-175.

36. Damodaran, V. B.; Murthy, N. S., Bio-inspired strategies for designing antifouling biomaterials. *Biomaterials research* **2016,** *20* (1), 18.

37. Zarei, M.; Bindila, L.; Souady, J.; Dreisewerd, K.; Berkenkamp, S.; Müthing, J.; Peter-Katalinić, J., A sialylation study of mouse brain gangliosides by MALDI a-TOF and o-TOF mass spectrometry. *Journal of mass spectrometry* **2008,** *43* (6), 716-725.

38. Ivleva, V. B.; Sapp, L. M.; O'Connor, P. B.; Costello, C. E. J. J. o. t. A. S. f. M. S., Ganglioside analysis by thin-layer chromatography matrix-assisted laser desorption/ionization orthogonal time-of-flight mass spectrometry. **2005,** *16* (9), 1552-1560.

39. Chan, K.; Lanthier, P.; Liu, X.; Sandhu, J. K.; Stanimirovic, D.; Li, J., MALDI mass spectrometry imaging of gangliosides in mouse brain using ionic liquid matrix. *Analytica chimica acta* **2009,** *639* (1-2), 57-61.

40. Weishaupt, N.; Caughlin, S.; Yeung, K. K.-C.; Whitehead, S. N., Differential anatomical expression of ganglioside GM1 species containing d18: 1 or d20: 1 sphingosine detected by MALDI imaging mass spectrometry in mature rat brain. *Frontiers in neuroanatomy* **2015,** *9*, 155.

41. Valdes-Gonzalez, T.; Goto-Inoue, N.; Hirano, W.; Ishiyama, H.; Hayasaka, T.; Setou, M.; Taki, T., New approach for glyco-and lipidomics–Molecular scanning of human brain gangliosides by TLC-Blot and MALDI-QIT-TOF MS. *Journal of neurochemistry* **2011,** *116* (5), 678-683.

42. Wood, A.; Chen, B.; Pathan, S.; Bok, S.; Mathai, C.; Gangopadhyay, K.; Grant, S.; Gangopadhyay, S. J. R. a., Influence of silver grain size, roughness, and profile on the extraordinary fluorescence enhancement capabilities of grating coupled surface plasmon resonance. **2015,** *5* (96), 78534-78544.

43. Imberty, A.; Varrot, A., Microbial recognition of human cell surface glycoconjugates. *Current opinion in structural biology* **2008,** *18* (5), 567-576.

44. Shanta, P. V.; Li, B.; Stuart, D. D.; Cheng, Q. J. A. C., Plasmonic Gold Templates Enhancing Single Cell Lipidomic Analysis of Microorganisms. **2020,** *92* (9), 6213-6217.

45. Wilkop, T.; Wang, Z.; Cheng, Q., Analysis of μ-contact printed protein patterns by SPR imaging with a LED light source. *Langmuir* **2004,** *20* (25), 11141-11148.

46. Phillips, K. S.; Han, J.-H.; Martinez, M.; Wang, Z.; Carter, D.; Cheng, Q., Nanoscale glassification of gold substrates for surface plasmon resonance analysis of protein toxins with supported lipid membranes. *Analytical chemistry* **2006,** *78* (2), 596-603.

47. Phillips, K. S.; Wilkop, T.; Wu, J.-J.; Al-Kaysi, R. O.; Cheng, Q., Surface plasmon resonance imaging analysis of protein-receptor binding in supported membrane arrays on gold substrates with calcinated silicate films. *Journal of the American Chemical Society* **2006,** *128* (30), 9590-9591.

48. Hu, W.; He, G.; Zhang, H.; Wu, X.; Li, J.; Zhao, Z.; Qiao, Y.; Lu, Z.; Liu, Y.; Li, C. M. J. A. c., Polydopamine-functionalization of graphene oxide to enable dual signal amplification for sensitive surface plasmon resonance imaging detection of biomarker. **2014,** *86* (9), 4488-4493.

49. Koga, M.; Tatsumoto, M.; Yuki, N.; Hirata, K. J. J. o. N., Neurosurgery; Psychiatry, Range of cross reactivity of anti-GM1 IgG antibody in Guillain-Barré syndrome. **2001,** *71* (1), 123-124.

Chapter 3: Surface Plasmon Resonance Imaging (SPR*i*) in Combination with Machine Learning for Microarray Analysis of Multiple Sclerosis Biomarkers in Whole Serum

3.1 Introduction:

Rapid detection and monitoring of disease biomarker levels is vital to medical diagnosis and therapeutic intervention, and thus constitutes an important part of research endeavors for the advancement of medical sciences.[1] These biomarkers allow for early diagnosis and thus enable disease differentiation, leading to faster implementation of treatments and targeted therapies.[2] For multiple sclerosis (MS), several antibody and protein biomarkers were found to target cell membrane components of the myelin sheath, such as gangliosides and sulfatides. The myelin sheath is a lipid rich substance that surrounds and insulates the neurons of the central nervous system (CNS), allowing for transmission of electrical pulses that control various functions throughout the body.[3-6] Among various membrane components of the myelin sheath, gangliosides have received extensive research attention as they are significant for maintaining structural stability, assist in cell to cell interactions, and aid in the regeneration and growth of axons.[7-9] It has been observed that with the progression of MS, the myelin sheath's integrity diminishes, severely impacting the CNS' ability to communicate with the rest of the body.[10] The damaged neural areas will cease to function normally, resulting in the symptoms associated with MS.[11]

Current diagnosis of MS relies heavily on characterizing damage to the CNS by scanning for plaques or scar tissue, which indicate that trauma or an autoimmune attack has occurred.[12, 13] Evaluation of the severity of the plaques is normally achieved by using a combination of several techniques, including magnetic resonance imaging (MRI), evoked potential, spinal taps, and blood tests,[12] while only spinal taps are routinely used for direct detection of MS protein biomarkers.[13] Collection of cerebrospinal fluid (CSF), however, requires the use of lumbar punctures, which are incredibly painful, difficult to perform, and only allow for a small amount of CSF to be collected,[14] leading to some diminished interest in the monitoring of MS progression.[15] Concurrently, there has been a great deal of interest in the development of new approaches for quantifying MS biomarkers in blood.[16] Presently, blood tests are routinely performed to screen for established markers of other diseases that have similar symptoms as MS for the purpose to rule them out in diagnosis.[17] They are not normally used for direct MS diagnosis due to limited biomarker presence in blood resulting in much lower concentrations when compared to CSF samples.[18] Therefore, moving to a blood-test based detection method for MS markers would require technical development for sensitivity improvement and robustness that could 1) quantify concentrations and analyte/antigen interactions concurrently, 2) identify and differentiate cross reactivity between biomarkers, and 3) eliminate background signals from various other components in the patients' blood sample.

While fluorescence and chemiluminescence detection schemes have been used,[19] surface plasmon resonance imaging (SPRi), a label-free, real time, and direct detection method, has increasingly been used broadly to detect disease biomarkers in various

biological samples.[20] The benefits offered by SPRi includes compatibility with microarrays that allow for screening of multiple biomarkers simultaneously in a high throughput and multiplexed manner, which can drastically improve the monitoring of disease progression.[21, 22] A major drawback, however, is the nonspecific binding when dealing with complex samples, such as blood.[20] Another challenge is the cross reactivity among various biomarkers in the sample, which is a particularly troublesome issue for anti-ganglioside antibody detection where the difference among carbohydrate headgroups is small. This convolutes signals and makes detection unreliable due to false positive and/or negative results, causing major concerns for use in clinical studies.[23] Antifouling surfaces and new SPRi methodologies have therefore been the focus of many works to reduce these undesirable interferences.[19, 22, 24-29]

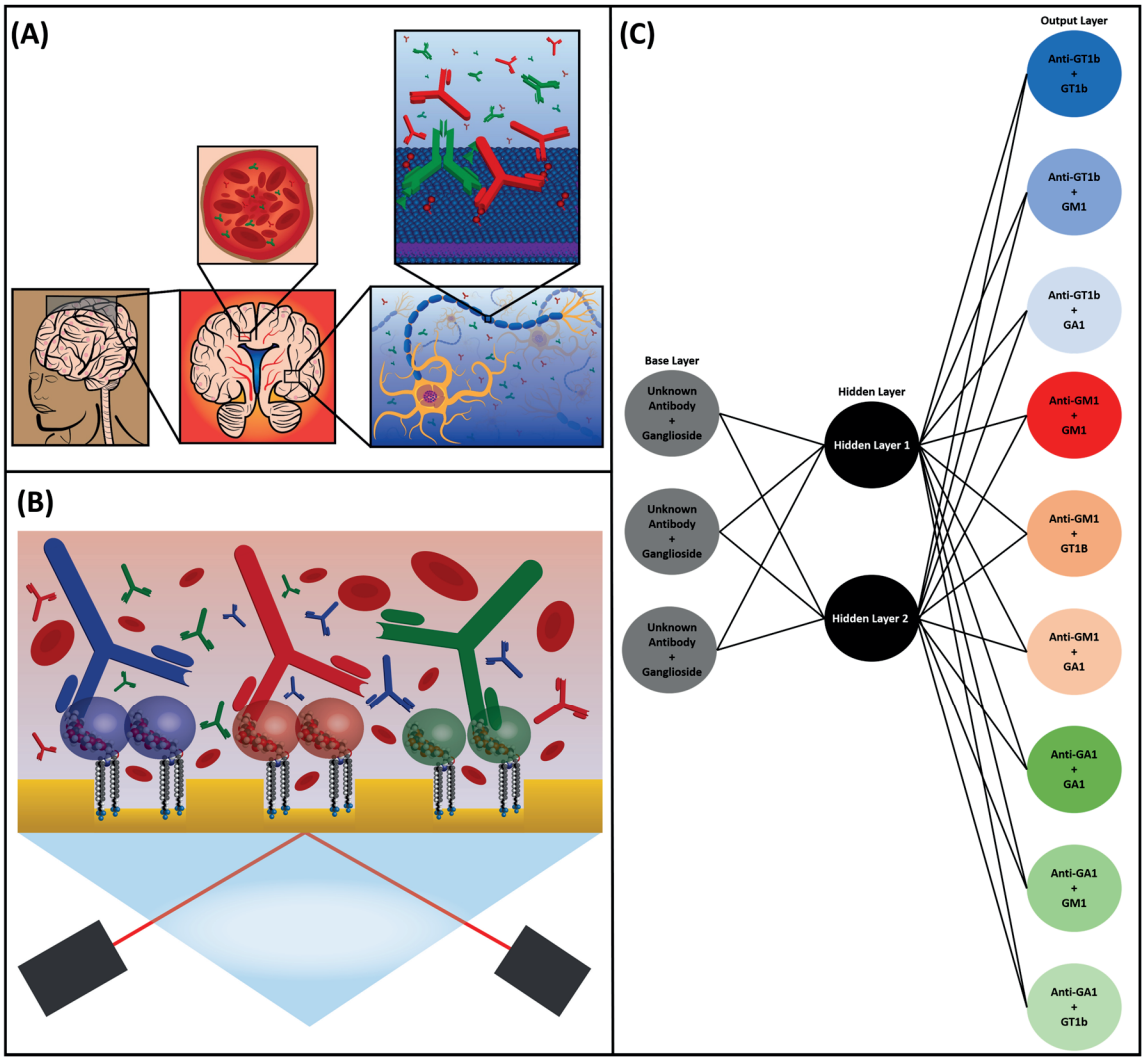


Figure 3.1. (A) Graphical representation of the biological process of antibodies attacking the myelin sheath in multiple sclerosis during an autoimmune attack. (B) Capture and detection scheme of anti- ganglioside antibodies associated with multiple sclerosis via the ganglioside microarray and SPRi. (C) Visualization of the machine learning algorithm for a neural network process including base layer, hidden layers, and output layers for all potential analyte antigen interaction in this study.

Recently we reported a near super hydrophobic, perfluorodecyl-trichlorosilane (PFDTS) surface for the detection of MS biomarkers. The antifouling properties of the ganglioside PFDTS substrate were evaluated and characterized, and detection of 3 anti-ganglioside antibodies in 10 % serum was successfully demonstrated.[22] In this work we

expand the study to investigate the sensing performance in clinical conditions by coupling machine learning to the differentiation of MS specific antibodies in undiluted serum (Figure 3.1). Extensive controls were used to evaluate cross reactivity between the investigated antibodies and gangliosides, which is critical for obtaining an accurate diagnosis in a clinical setting. To assess cross reactivity among structurally similar carbohydrate antigens, we have performed modeling and statistical analysis using various machine learning (ML) algorithms for post-acquisition data analysis. Datasets of endpoint results, association, steady state, and dissociation energies were utilized, which provided a more comprehensive understanding into the observed interactions than what traditional methods can achieve.[30] Categorization by ML relies on complex algorithms to detect patterns in the raw data where similar observations can be grouped or clustered together.[31] This in-depth analysis allows for the discovery of previously overlooked patterns that can be used to train the ML models to aid in the identification and differentiation of analytes present in a complex biological sample.[32] The findings of the presented study establishes a new methodology to address the technical difficulties of identifying analyte/antigen interactions in complex media and that by correctly training ML models, they can be implemented to improve biomarker detection for disease diagnosis.

3.2 Materials and Methods:

Materials and Reagents: Monosialoganglioside GM_1 was purchased from Matreya (Pleasant Gap, PA). Trisialoganglioside GT_{1b} was obtained from Biosynth (Itsaca, Il). Asialoganglioside GA_1 was acquired from Sigma-Aldrich (St. Louis, MO).

1H,1H,2H,2H-Perfluorodecyltrichlorosilane (PFDTS) was purchased from Fisher Scientific (Pittsburgh, PA). Anti asialoganglioside GA_1 human anti mouse monoclonal antibody and anti monoganglioside GM_1 rabbit polyclonal antibody, were both obtained from Abcam (Cambridge, UK). Anti Trisialoganglioside GT_{1b} ganglioside mouse monoclonal antibody was obtained from Millipore Sigma (Billerica, MA). Human serum was purchased from Innovative Research (Upper Marlboro, MD).

Fabrication of SPRi Substrates: The SPRi arrays biochips were fabricated via the protocol reported in a previous paper.[33] In short, glass slides were spin-coated with hexamethyldisilazane (HMDS) to promote adhesion. After baking at 110 °C for approximately 1 minute, UV exposure via a Karl-Suss MA-6 system allowed for the creation of an array pattern on the photoresist, after which standard photoresist development protocols were implemented. A 2 nm layer of titanium was first deposited to act as an adhesion layer, after which a layer of 200 nm of gold was deposited to form wells. The remaining photoresist was removed from the surface with acetone. A second 2nm layer of titanium was deposited followed by a 48 nm layer of gold to form the sensing surface inside of the wells. PECVD was then used to deposit 1-3 nm of SiO_2 on the microarray chips, which would later be treated with PFDTS. The final product was a microarray consisting of 10x10 well arrays that were 200 nm deep and 600 μm in diameter.

Surface Functionalization and Preparation: Functionalization of the chips was carried out with a similar protocol used previously.[22] The surface of the chips was submerged in 1 mM PFDTS in toluene. After 30 minutes, the chip was removed from the

solution and rinsed with toluene, ethanol, and deionized water and dried under nitrogen gas. Once the chips were completely dry, 1.5 μL of the 100 μg/mL stock solution for each ganglioside was incubated and allowed to dry in air to create four different working channels. An S-shaped PDMS flow cell was used to create eight working wells per channel where functionalization and interactions occurred under the same configuration for reproducibility. The first channel was left un-functionalized as an internal reference to measure chip to chip variation, whereas 1.5 μL of 100 μg/mL stock solutions was employed for generating ganglioside channels with GA_1, GM_1, and GT_{1b}.

SPRi analysis: SPRi measurements were conducted on a home-built setup; a detailed description of which can be found in previous work.[34] The functionalized chips were mounted onto an optical stage that houses a PDMS flow cell. The array was placed in contact with an equilateral SF2 prism (n = 1.65) with a layer of refractive index matching fluid (Cargill Laboratories, Cedar Grove, NJ). A 648 nm light emitting diode (LED) was used as the light source for SPR excitation. Reflected images of the microarray were captured by a cooled 12-bit CCD camera (QImaging Retiga 1300) and data acquisition was controlled via a home built LabView program. Intensity data was normalized by using the intensity from the *p*-polarized light over the *s*-polarized beam and described as a percentage.

Statistical analysis and Machine Learning: Analysis of variance (ANOVA) was conducted in Excel with the Analysis ToolPak add-in and used the end point data obtained with SPRi. Partial least squares discrimination analysis (PLS-DA) plots were produced

with MetaboAnalyst. Principal component analysis (PCA) was completed with the prcomp function in R and graphed with the ggbiplot package with an ellipse probability set to 95 % using the endpoint data. Neural Network (nnet) was used to analyze both the endpoint data and SPRi sensorgram data. Nnet was conducted using the nnet package for R and was plotted using an expanded grid in R. K nearest neighbor (Knn) from the caret package was used to evaluate regions of the sensorgram related to the antibody spiked in whole serum interaction study. The Knn model was visualized with ggbiplot packages. For ML 70 % of the data was used to train the models and 30 % to test. Each model had random iterations incorporated into them so that all of the data could be used to test the success of the model.

3.3 Results and Discussion:

SPR imaging analysis of anti-gangliosides in Serum: Microarrays offer the benefit of detecting many biomarkers simultaneously. However, reliably and effectively using microarrays in a clinical setting requires an in-depth understanding of background signals from the biological sample and the cross reactivity between the biomarkers of interest that could be present in the sample.[25, 35] The microarray used in this study has a 4 x 8 arrangement with the 3D printed looped flow cell, capable of monitoring the presence of many biomarkers concurrently. Previously, we showed that our myelin sheath mimic is capable of detecting and differentiating MS specific antibodies in diluted serum at 100 ng/mL.[22] In this study, we expand the investigation and focus on detecting and differentiating three MS specific anti-ganglioside antibodies in whole serum at disease-relevant concentrations ranging from 3 to 25 ng/mL.[10] Antibodies for GT_{1b}, GM_1, and GA_1 gangliosides were used as they have been associated with symptoms commonly observed

in MS.[4, 10, 36] Anti-GT_{1b} has been linked to the loss of muscle control in the limbs, whereas anti-GM_1 and anti-GA_1 are believed to play significant roles in the damage of myelin associated with the optic nerves, as they both have been correlated to changes in and loss of vision.[10, 36-38] Detection of these antibodies, as well as other anti-ganglioside antibodies, should drastically increase the confidence and speed at which a diagnosis is determined when coupled with currently established MS detection methods.

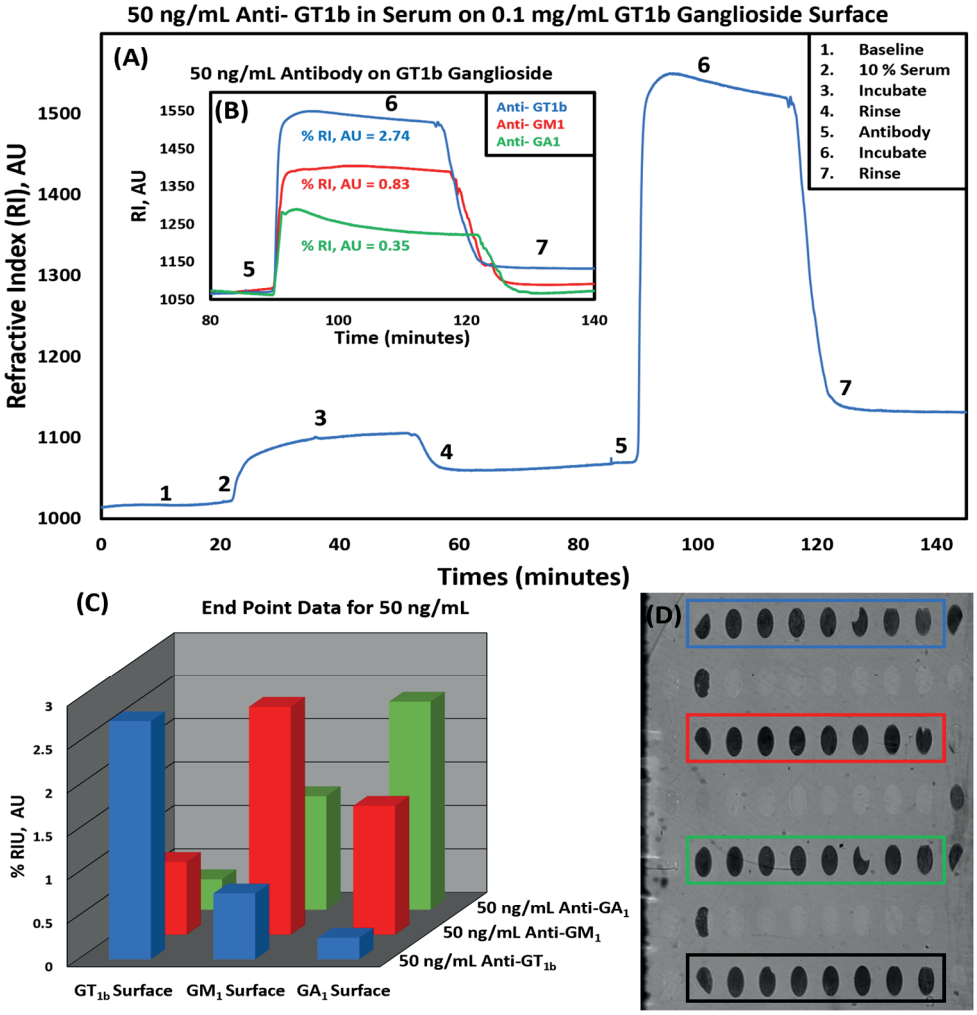


Figure 3.2. (A) Entire sensorgram for 50 ng/mL of anti-$GT1_b$ in serum on a $GT1_b$ ganglioside surface. (B) Average of 5 sensorgrams to depict association, steady state, and dissociation binding interactions for 50 ng/mL anti-$GT1_b$ (Blue), anti-GM_1 (Red), anti-GA_1 (Green) interacting with a $GT1_b$ ganglioside surface. (C) Average of all observed bulk changes caused by MS specific antibodies at 50 ng/mL in serum on the PFDTS functionalized ganglioside microarray. (Blue) % RI, AU caused by 50 ng/mL anti-GT_{1b} on 0.1 mg/mL GT_{1b}, GM_1, and GA_1 ganglioside surfaces, (Red) binding between 50 ng/mL anti-GM_1 and 0.1 mg/mL on a GT_{1b}, GM_1, and GA_1 ganglioside surfaces, and (Green) binding between 100 ng/mL anti-GA_1 and 0.1 mg/mL GT_{1b}, GM_1, and GA_1 ganglioside surfaces. (D) Image of ganglioside microarray by the CCD camera used in the SPR imaging experiments. Each color indicates the functionalization of the surface PFDTS only (Black), GA_1 ganglioside surface (Green), GM_1 ganglioside surfaces (Red), and GT_{1b} ganglioside surfaces (Blue).

Figure 3.2 shows the SPRi results for specific anti-gangliosides antibodies under various conditions and an image of the ganglioside microarray. The first step in the experiment was to inject 10 % serum diluted with PBS to block the surface to account for cross reactivity and nonspecific binding. Different concentrations of serum dilutions were tested, but higher concentrations did not offer any benefits over the 10 % dilution. As shown in Figure 3.1 A and B, there is a large angle shift due to the change in refractive index units (RIU) once the spiked whole serum is introduced. This shift can be attributed to the high concentration of proteins and other biological components present in the sample.[35] Once the rinse cycle is initiated, the vast majority of the material that caused the large shift is rinsed off, leaving behind only the specific antibody/ganglioside interaction of interest and other materials of extremely high affinity. To confirm the specific interaction, we conducted cross reactivity evaluations with several antibodies for each concentration to determine how much of the observed shifts were due to specific analyte/antigen interactions, as shown in Figure 3.2B. The small amount of nonspecific binding can be attributed to the unique properties of the near super hydrophobic surface and the selectivity of the sialic acids (SA) present on the antigenic sensing sites of the gangliosides.[22] Figure 3.2C shows the bulk changes (in % RIU) of the investigated antibodies at 50 ng/mL with the ganglioside microarray. High cross reactivity can be observed between anti-GM_1 and anti-GA_1 and their respective gangliosides, whereas there is negligible cross reactivity between anti-GT_{1b} and anti-GA_1. This observation can be attributed to the number of SA for each ganglioside.[22] This trend was present in all of the investigated concentrations and can be seen in the 3D bar graphs of Figures 3.5. The

PFDTS carbohydrate chips used here appear effective in addressing large levels of nonspecific binding from complex media and cross reactivity between analytes, one of the major challenges of label-free detection methods.[35]

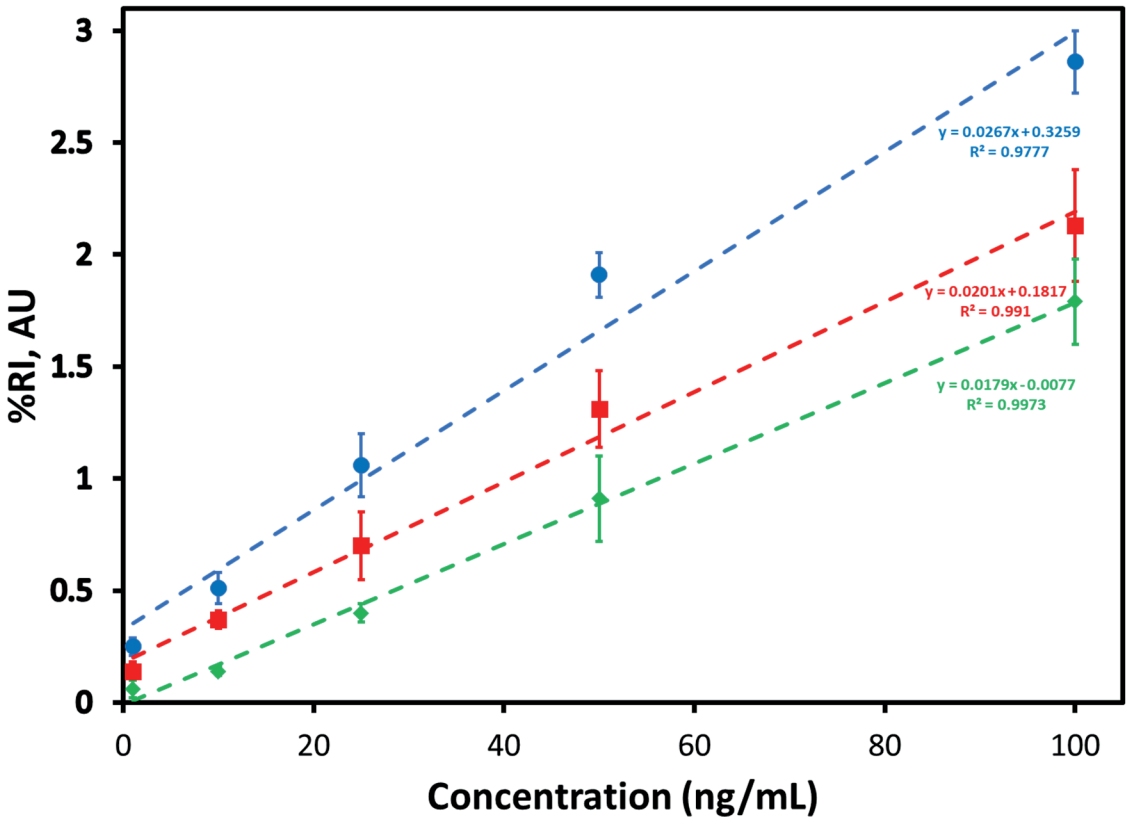


Figure 3.3. Calibration curve for the specific binding of all of the investigated antibodies in serum. Anti-$GT1_b$ (Blue), anti-GM_1 (Red), and anti-GA_1 (Green). Each data point is the average of at least 5 wells.

A calibration curve for anti-GT_{1b}, anti-GM_1, and anti-GA_1 after background subtraction is shown in Figure 3.3. Each surface has a working range of 1 ng/mL to 100 ng/mL in undiluted serum. The limits of detection (LOD) using the 3σ method were calculated to be 4.5 ng/mL, 5.6 ng/mL, and 6.6 ng/mL for anti-GT_{1b}, anti-GM_1, and anti-GA_1 respectively. Based on these calculations, the detection limit of the carbohydrate microarrays appears to fall within the concentration range of antibodies typically seen in

patient serum samples,[18, 39] demonstrating that our microarray is capable of detecting, quantifying, and differentiating MS biomarkers in a clinical setting. This trends validates that the presented methodology offers unique and clear benefits when compared to other recently developed detection methods for MS specific biomarkers, as presented in Table 3.1. While each of the listed techniques have aided in the progression towards more reliable detection of MS, they all lack in one or several capabilities that the presented method provides. Most notably though is in regards to simultaneously screening, quantifying, and differentiating multiple biomarkers in biological samples.

While the developed method has met the technical requirements for the routine use of blood tests for the evaluation of MS biomarkers in patient samples and even may streamline the diagnostic procedure, the observed cross reactivity suggested simple quantification by binding signals may be an oversimplified approach and could ignore potential interferences to the observed results, as is the case for the studies shown in Table 3.1. Therefore, an accurate detection and effective differentiation between analyte/antigen interactions would require in-depth statistical analysis.

Table 3.1. Comparison of recent biosensors developed for the detection and monitoring of MS biomarkers.

	Bioreceptor	Method	Amplification	LOD	Detected in Biological Sample	Reference
anti-GT_{1b} anti-GM_1 anti-GA_1	GT_{1b} ganglioside GM_1 ganglioside GA_1 ganglioside	SPRi	No	4.5 ng/mL 5.6 ng/mL 6.6 ng/mL	Yes whole serum	This work
anti-GT_{1b} anti-GM_1 anti-GA_1	GT_{1b} ganglioside GM_1 ganglioside GA_1 ganglioside	SPRi	No	2.34 ng/mL Only tested 100 ng/mL Only tested 100 ng/mL	Yes 10 % serum	40
miR-17	DNA probe	Localized SPR	Yes	1 pM	No PBS	41
miR-422 miR-223 miR-216 miR-23A	DNA probe	SPRi	Yes	0.55 pM 0.88 pM 1.19 pM 1.79 pM	No PBS	42
miR-145	DNA probe and fluorescent silver nanocluster	Fluorescence spectrophotometer	Yes	0.1 nM	Yes 50 % serum	43
Unidentified antibodies present in purified MS patient serum samples	Synthetic glycopeptide CSF114(Glc) with ferrocenyl moiety	Cyclic and square wave voltammetry	Yes Indirectly by excess free glycopeptide Fc-CSF114(Glc) present in solution	Not determined only reported a difference between healthy and sick samples	Yes Purified patient serum samples of 0.04 to 4 ug/mL	44
Unidentified antibodies present in patient serum samples	Synthetic glycopeptide CSF114(Glc)	SPR	No	Not reported only compared response between healthy and sick samples	Yes 1 % and 2 % serum	45

Cross reactivity and statistical analysis: Robust statistical analysis was performed to analyze and characterize data from both SPRi sensorgrams and endpoint results. The endpoint data is the observed change in RIU of the analyte-antigen interactions after accounting for observed cross reactivity between the other biomarkers and nonspecific binding from serum. The calibration curves seen in Figure 3.3 are generated using the collected end point data after confirming all data was statistically relevant using analysis of variance (ANOVA) as shown in Table 3.2. Further analysis was conducted

using principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA), in order to characterize the data to reveal impact by specific, nonspecific, and cross reactivity interactions, yielding more thorough evaluation on antibody/ganglioside interactions than what traditional endpoint assay studies are capable of. The use of these statistical analyses gives a good assessment of the effectiveness of the ganglioside microarray's ability to screen MS biomarkers.

Table 3.2. Table depicting analysis of variance for all utilized end point data values before accounting for nonspecific and cross reactivity at each concentration utilized in machine learning algorithms and as a whole dataset. The P-values and variance values calculated with ANOVA in Excel via data analysis tool indicate that all observed antibody ganglioside interactions are statistically relevant by being ≤ 0.05.

Anti-GT1b ANOVA

SUMMARY	GT1B	GM1	GA1			
100 ng/mL						
Count	5	5	5			
Variance	0.015368	0.018441	0.002701			
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	31.34399	2	15.672	1287.733	9.95E-15	3.885294
Within Groups	0.146043	12	0.01217			
50 ng/mL						
Count	5	5	5			
Variance	0.01025	0.02632	0.00062			
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	17.33081	2	8.665407	699.011	3.80E-13	3.885294
Within Groups	0.14876	12	0.012397			
25 ng/mL						
Count	5	5	5			
Variance	0.020566	0.01993	0.00753			
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	4.788357	2	2.394179	149.5556	3.29E-09	3.885294
Within Groups	0.192103	12	0.016009			
10ng/mL						
Count	5	5	5			
Variance	0.00433	0.01937	0.01827			
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	5.844093	2	2.922047	208.8668	4.74E-10	3.885294
Within Groups	0.16788	12	0.01399			
1 ng/mL						
Count	5	5	5			
Variance	0.00037	0.00588	0.00967			
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.745293	2	0.372647	70.22236	2.38E-07	3.885294
Within Groups	0.06368	12	0.005307			
Total						
Count	25	25	25			
Variance	0.894912	0.029253	0.008163			
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Sample	7.394362	4	1.84859	154.3781	7.40E-31	2.525215
Columns	45.78952	2	22.89476	1911.97	4.64E-55	3.150411
Interaction	14.26303	8	1.782879	148.8904	1.26E-36	2.096968
Within	0.718466	60	0.011974			

Anti-GM1 ANOVA

SUMMARY	GT1B	GM1	GA1
100 ng/mL			
Count	5	5	5
Variance	0.004381	0.049561	0.004989

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	15.38125	2	7.690623	327.6031	3.38E-11	3.885294
Within Groups	0.281705	12	0.023475			

50 ng/mL			
Count	5	5	5
Variance	0.04938	0.02825	0.049982

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	8.176867	2	4.088433	81.05977	1.07E-07	3.885294
Within Groups	0.605247	12	0.050437			

25 ng/mL			
Count	5	5	5
Variance	0.008651	0.02258	0.04023

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	1.947242	2	0.973621	40.87339	4.40E-06	3.885294
Within Groups	0.285845	12	0.02382			

10 ng/mL			
Count	5	5	5
Variance	0.009075	0.00142	0.00437

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.190687	2	0.095343	19.24209	0.00018	3.885294
Within Groups	0.059459	12	0.004955			

1 ng/mL			
Count	5	5	5
Variance	0.049916	0.015519	0.00065

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.168813	2	0.084406	2.899313	0.049392	3.885294
Within Groups	0.34935	12	0.029113			

Total			
Count	25	25	25
Variance	0.110885	0.605513	0.143275

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Sample	8.521274	4	2.130319	80.81601	1.82E-23	2.525215
Columns	15.33557	2	7.667785	290.886	1.33E-31	3.150411
Interaction	10.52928	8	1.31616	49.93001	1.09E-23	2.096968
Within	1.581606	60	0.02636			

Anti-GA1 ANOVA

SUMMARY	GT1B	GM1	GA1			
100 ng/mL						
Count	5	5	5			
Variance	0.003642	0.00703	0.034402			
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	13.65719	2	6.828593	454.4854	4.89E-12	3.885294
Within Groups	0.180299	12	0.015025			
50 ng/mL						
Count	5	5	5			
Variance	0.00253	0.04838	0.03493			
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	10.45736	2	5.22868	45.35632	2.54E-06	3.885294
Within Groups	1.383361	12	0.11528			
25 ng/mL						
Count	5	5	5			
Variance	0.00152	0.01923	0.02968			
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	3.355053	2	1.677527	99.79338	3.33E-08	3.885294
Within Groups	0.20172	12	0.01681			
10 ng/mL						
Count	5	5	5			
Variance	0.002	0.04287	0.00032			
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	3.112013	2	1.556007	103.2976	2.74E-08	3.885294
Within Groups	0.18076	12	0.015063			
1 ng/mL						
Count	5	5	5	15		
Variance	0.000317	0.006419	0.000717	0.030701		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.400002	2	0.200001	80.51486	1.11E-07	3.885294
Within Groups	0.029808	12	0.002484			
Total						
Count	25	25	25			
Variance	0.008757	0.171289	0.432556			
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Sample	5.610728	4	1.402682	42.59269	6.87E-17	2.525215
Columns	23.86584	2	11.93292	362.3452	3.19E-34	3.150411
Interaction	7.115773	8	0.889472	27.00896	3.49E-17	2.096968
Within	1.975948	60	0.032932			

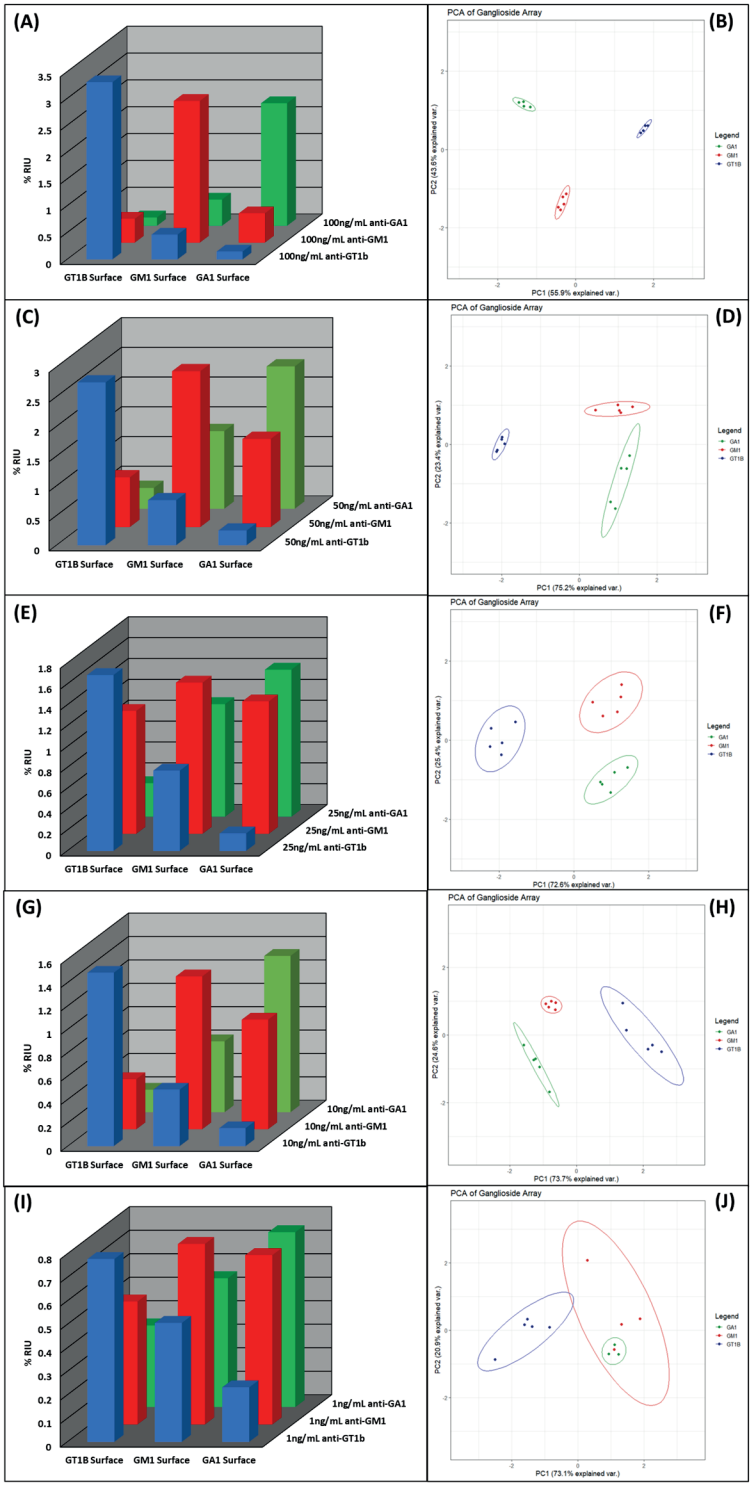


Figure 3.4. Raw end point data values used in presented study displayed as 3D bar graphs and PCA of the data. There is clear confusion occurring for 1 ng/mL as seen in the PCA analysis in (2J).

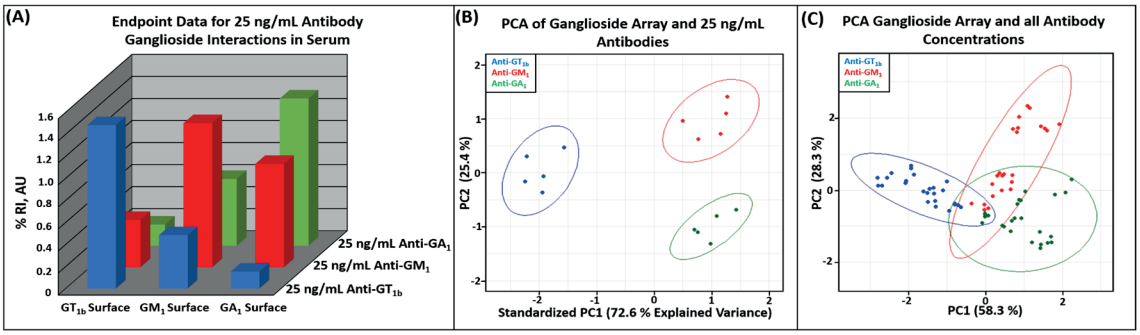


Figure 3.5. (A) 3D bar graph showing % RI, AU of all three antibodies interacting with the ganglioside microarray at 25 ng/mL in serum. Anti-GT_{1b} interactions (Blue), anti-GM_1 interactions in (Red), and anti-GA_1 interactions in (Green). (B) Principal component analysis (PCA) showing the ability to separate the three anti-ganglioside antibodies based on their induced response across the whole microarray at 25 ng/mL. (C) PCA of all antibody/ganglioside interactions at each concentration showing overlap of anti-GT_{1b} beginning at 10 ng/mL and at 25 ng/mL for both anti-GM_1 and anti-GA_1.

Principal component analysis was first carried out with the endpoint data for each antibody interacting with three different ganglioside surfaces individually and as an entire dataset. PCA, an unsupervised statistical model with a clustering statistical algorithm that looks for linear patterns in complex datasets,[46,47] showed that at higher concentrations it could easily determine the specific antibody/ganglioside interactions, but was less effective in differentiating interactions at concentrations below 10 ng/mL for anti-GT_{1b} and below 25 ng/mL for anti-GM_1 and anti-GA_1, as shown in Figure 3.5. The different number of SAs on the ganglioside antigens may explain why anti-GT_{1b} antibodies could still be differentiated at 10 ng/mL, as anti-GT_{1b} specifically targets three SAs whereas anti-GM_1 and anti-GA_1 target antigenic site with one and zero SA, respectively.[48] Figure 3.4 shows that when all concentrations are plotted, there is a significant overlap between the antibodies at concentrations below 10 ng/mL, but at 25 ng/mL only overlap is between anti-GM_1 and anti-GA_1. This agrees well with our previous observation in 10 % serum where the majority of cross reactivity occurred between anti-GT_{1b} and anti-GM_1 or anti-

GM_1 and anti-GA_1 but little between anti-GT_{1b} and anti-GA_1.[22] It is apparent that PCA had difficulty differentiating the antibodies at concentrations below 10 ng/mL as the measurements themselves started to show uncertainty (LODs determined around 7 ng/mL).

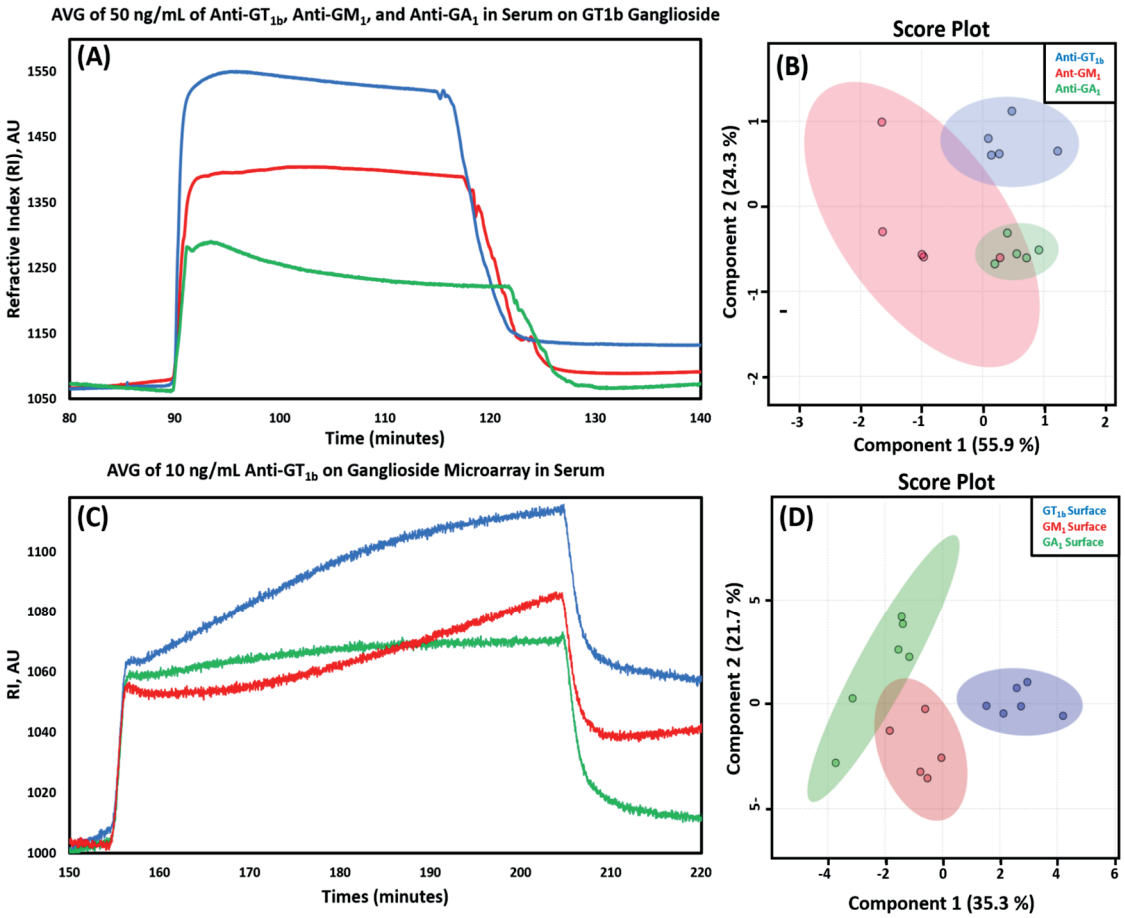


Figure 3.6. (A) Average of sensorgrams for binding associated regions between 50 ng/mL of the three investigated antibodies with a GT_{1b} ganglioside functionalized PFDTS surface. Anti-GT_{1b} interactions (Blue), anti-GM_1 interactions in (Red), and anti-GA_1 interactions in (Green). (B) Sample of Partial Least Squares Discriminant Analysis (PLS-DA) for all three antibodies at 50 ng/mL on a GT_{1b} ganglioside surface. (C) Average association, steady state, and dissociation regions for anti-GT_{1b} interacting with the 3 different functionalized ganglioside PFDTS surfaces. (Blue) GT_{1b} ganglioside surface, (Red) GM_1 ganglioside surface, and (Green) GA_1 ganglioside surface. (D) Sample of PLS-DA analysis of anti-GT_{1b} at 10 ng/mL classification based upon interactions with (Blue) GT_{1b} ganglioside surface, (Red) GM_1 ganglioside surface, and (Green) GA_1 ganglioside surface.

The PCA analysis with endpoint data provides a good glimpse of the interaction properties between antibodies and gangliosides, but the results are limited. Further characterization of this complex property was carried out with kinetic interaction data from the SPRi sensorgrams. To achieve this, we utilized PLS-DA to analyze the regions associated with binding kinetics in the sensorgrams related to the antibody ganglioside interactions. PLS-DA is a supervised or classification based statistical method that looks for trends in the whole dataset and makes a prediction based upon these trends to determine the relationship that one dataset has to another.[49] We performed PLS-DA on the sensorgrams for 50 ng/mL of the three MS specific antibodies in serum interacting with a GT_{1b} ganglioside surface (Figure 3.6A) and for 10 ng/mL of anti-GT_{1b} on the 3 different ganglioside surfaces (Figure 3.6C). Figure 3.6B shows minimal overlap between anti-GT_{1b} and anti-GM_1 and between anti-GM_1 and anti-GA_1, but none between anti-GT_{1b} and anti-GA_1. The interactions observed between the three antibodies and the ganglioside surface can be attributed to anti-GA_1 not targeting SAs whereas both anti-GT_{1b} and anti-GM_1 do. In Figure 3.6D it is clear that 10 ng/mL anti-GT_{1b} specifically binds to the GT_{1b} ganglioside surface and there is only minor overlap between GM_1 and GA_1, while the majority of binding occurs to GT_{1b} and GM_1 ganglioside surfaces. There is no observable overlap between anti-GT_{1b} interactions with a GA_1 ganglioside surface, indicating that anti-GT_{1b} specifically targets the SA on the carbohydrates head group. PLS-DA using the kinetic interaction data appears to differentiate antibodies and ganglioside interactions more effectively than PCA did.

The promising results with PLS-DA and PCA prompted us to explore the datasets further to understand the cross reactivity between the biomarkers. Although these statistical methods are effective to reveal cross reactivity, they are limited in that only one dataset was compared to another, rather than finding patterns hidden in the endpoint data, sensorgrams, and binding kinetics to predict the type of interactions that are occurring. Therefore, further statistical analysis with machine learning (ML) was performed to evaluate the SPRi raw data.

Neural Networks and k Nearest Neighbor Algorithms: ML has become increasingly relevant and present in various aspects of scientific investigation and society as a whole.[50] It appears to be only a matter of time before ML algorithms are used routinely in disease diagnostics due to its ability to handle large and complex datasets.[30] In this study we trained and tested a neural network (nnet) and a *k*-nearest neighbor (knn) model using endpoint data for both specific and cross reactive interactions. In addition, we performed these analyses using data from the sensorgrams that are linked to association, steady state, and dissociation energies of the analyte/antigen interactions based upon the changes in time and RIU values, which gave us access to over 65,000 data points per sensorgram. By training and testing these models with the collected data, we can evaluate the effectiveness of the algorithms and the functionalized substrates for detecting and differentiating between antibody ganglioside interactions in a pseudo clinical setting. It will allow us to determine if the combination of ML to label-free sensing methodology could facilitate the general high throughput screening of antibodies/markers, which may drastically improve disease diagnosis and the monitoring of their progression.

Nnets are essentially a virtual nervous system,[30] making it an ideal ML model to use in the detection and evaluation of MS. Nnets are composed of three layers: an input, hidden, and output layer.[32] A visualization of the nnet used in our study is shown in Scheme 1C. The input layer consists of data that the user feeds into the model to train and test the success rate of the algorithm;[51] for our study, our input layers included the time, RIU, endpoint data, and concentration depending on how we were analyzing the data. The hidden layer is the intermittent computations that occur to define patterns and investigate the data so that the algorithm can make predictions.[30] The output layer allows the user to interpret the results generated by the model, which can either be used for classification, as for the endpoint data, or a prediction, which was done for the sensorgram data.[51] Nnets have the ability to learn by themselves and produce outputs that are not limited to the inputs originally provided to them, making them ideal for aiding in the detection and monitoring of diseases where biomarkers in one patient can vary drastically to those in another patient. Nevertheless, these new outputs would need to be verified before being included in the diagnosis criteria.

The second model used in our work was kNN. kNN is a non-parametric classification model that works by analyzing and comparing a single data point to the entirety of the data series before moving to the next data point in the series; this allows for the detection and monitoring of unique trends that other ML algorithms might overlook.[52] In brief, the algorithm attempts to predict the correct class of the test data by calculating and identifying trends among a few data points (referred to as neighbors) and defines trends amongst the entire data series or the entire neighborhood.[30] The model then compares the

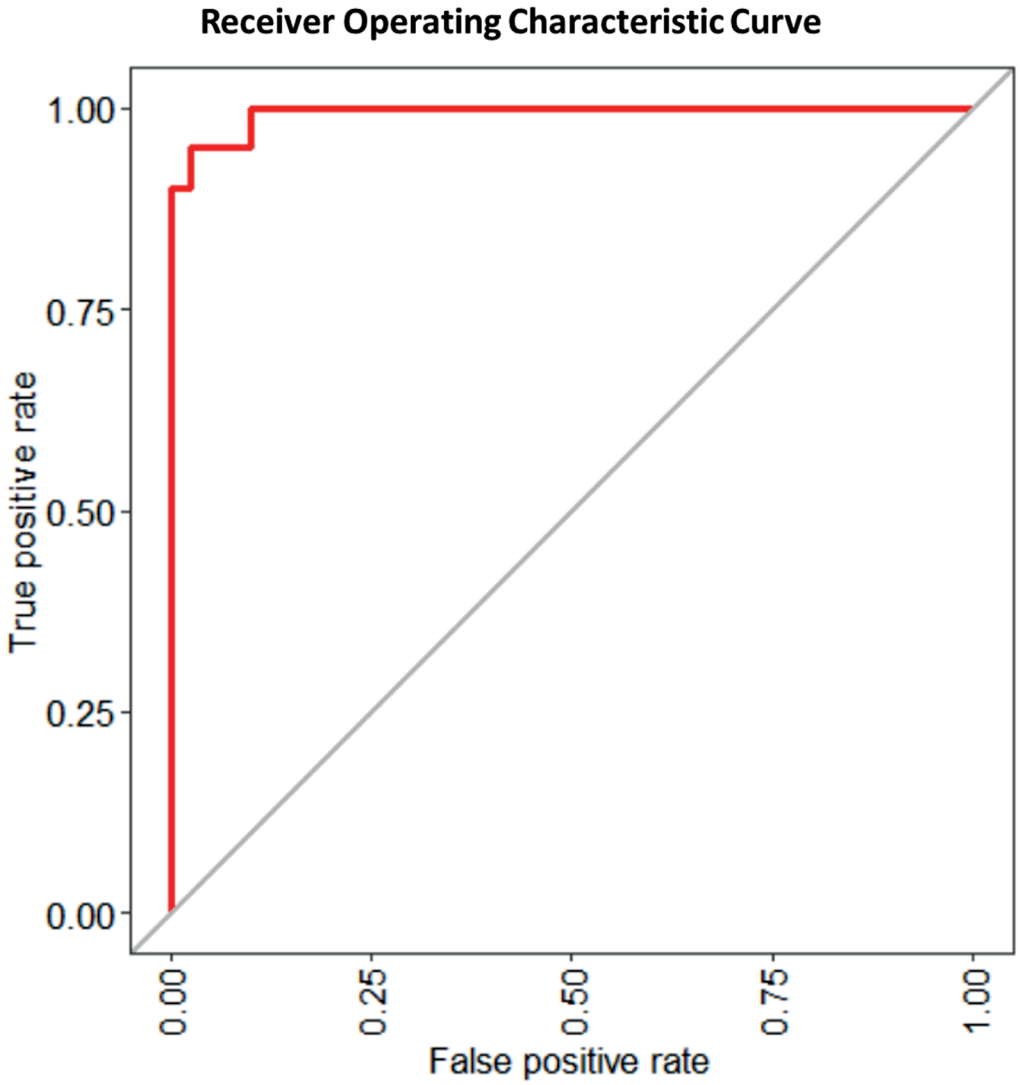
observed trends in the test data between individual data points, groups of data points, and the entire data series to trends observed in the training set.[51] Unlike the nnet, all potential analyte/antigen interactions would need to be predefined for the model to accurately identify, differentiate, and predict between healthy and sick patient samples as well as which analyte/antigen interactions were most likely occurring. Given the way that the kNN algorithm operates, we were only able to reliably utilize it on the SPRi sensorgram data.

To train and determine the accuracy of the algorithms for various datasets, we implemented multiple random iterations by randomly selecting 70 % of the collected data to be used as a training set and the other 30 % to test the iteration of the model.[32, 51, 53] This allowed us to use all of the collected data to train and test the model, resulting in a more reliable algorithm than what only a portion of the data would achieve.[51, 53] While increasing the percentage and number of training sets would undoubtedly improve the chance of the models to achieve the goal, it also carries the risk of over fitting,[51] which is a major concern and also the reason why random iterations were utilized in this work.[54] The accuracies for the ML algorithms discussed here are the average of all potential random iterations conducted in the current study.

There are a few limitations to consider when using ML algorithms for this type of data analysis. One must account for all potential cross reactive interactions and nonspecific binding for the model to be considered truly reliable.[32] When potential interactions are not taken into account or considered during the training period, it is very likely that false positive and negative results will occur.[30] Nevertheless, with a model properly set-up and

running, training can be completed relatively quickly as long as the data is reliable and reproducible.[55]

Neural Network Data Analysis: Figure 3.7 shows a receiver operating curve (ROC) for the endpoint data for all antibody/ganglioside interactions at each studied concentration, which constitutes a total of 225 observations each of which contain the baseline, incubation, and post rinse cycles present in each sensorgram. An ROC depicts the reliability of the algorithm to correctly classify models at all classification thresholds based upon whether the model correctly or incorrectly identified which antibody/ganglioside interactions were observed.[56] Our ROC curve indicates that the model has an overall accuracy of 94 %. The table shown below the ROC curve is the average of all possible random iterations that could occur, showing how accurately the model can identify for the antibody/ganglioside interactions occurring in serum. If the model was trained with all possible biomarkers associated with a disease, it could reliably differentiate and identify them based upon their specific, nonspecific, and cross reactive interactions, which would drastically improve the reliability of end point assays. We have applied the nnet to the endpoint data, which focused on using concentration and specific RIU results. In addition, we applied further analysis with the sensorgram data using both the nnet and kNN models.



Nnet Confusion Matrix on Endpoint Data

94 % Accuracy	GA_1 Gang.	GM_1 Gang.	GT_{1b} Gang.
Anti- GA_1	21	1	0
Anti- GM_1	1	20	1
Anti- GT_{1b}	0	1	21

Figure 3.7. Receiver operating characteristic curve (ROC) for the analysis of the endpoint data containing all antibody/ganglioside interaction with a neural network (nnet). Below the ROC curve is a representative confusion matrix of the nnet testing datasets using random iterations, to evaluate the accuracy of the model to identify the specific analyte antigen interactions of interest.

The nnet modeling based on the sensorgram data was focused on three antibodies at 50 ng/mL in whole serum on a GT_{1b} ganglioside surface and anti-GT_{1b} at 10 ng/mL in serum. This dataset was selected with the consideration that GT_{1b} has the most SA groups of the three investigated gangliosides and has known cross reactivity with anti-GM_1 and little with anti-GA_1. Anti-GT_{1b} at 10 ng/mL was selected as anti-GT_{1b} had the lowest LOD (4.5 ng/mL) and LOQ (15 ng/mL) of the three investigated antibodies in whole serum and has notable cross reactivity with the GM_1 ganglioside surface and none with a GA_1 ganglioside surface as shown previously with our statistical analysis.

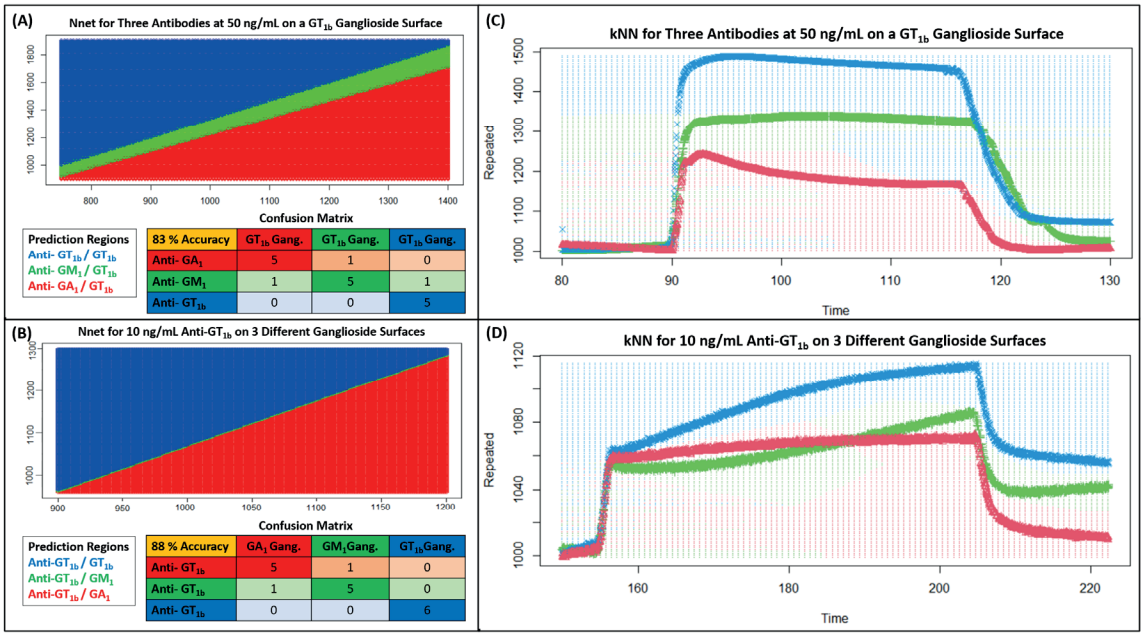


Figure 3.8. (A) nnet of the sensorgram for (Blue) 50 ng/mL anti-GT_{1b} on a GT_{1b} ganglioside surface, (Green) 50 ng/mL anti-GM_1 on a GT_{1b} ganglioside surface, and (Red) 50 ng/mL anti-GA_1 on a GT_{1b} ganglioside surface. (B) nnet for 10 ng/mL anti-GT_{1b} with (Blue) a GT_{1b} ganglioside surface, (Green) GM_1 ganglioside surface, and (Red) a GA_1 ganglioside surface. (C) K nearest neighbor model (kNN) depicting the binding interactions between (Red) 50 ng/mL anti-GA_1 on a GT_{1b} ganglioside surface, (Green) 50 ng/mL anti-GM_1 on a GT_{1b} ganglioside surface, and (Blue) 50 ng/mL anti-GT_{1b} on a GT_{1b} ganglioside surface. (D) KNN of 10 ng/mL anti-GT_{1b} interacting with a (Blue) GT_{1b} ganglioside surface, a (Green) GM_1 ganglioside surface, and a (Red) GA_1 ganglioside surface.

Figures 3.9A and 3.9B show the nnet plots based on the location of individual time points and their relation to specific RIU values. This is visualized as an expanded grid, depicting specific prediction regions based upon observed trends in the sensorgrams shape, as well as steady state and kinetic information. A confusion matrix is shown below each expanded grid depicting the model's ability to classify which antibody/ganglioside interactions are occurring. As shown in Figures 3.9A and 3.9B, there are clearly three distinct predictive regions based upon the location of these time points and their relation to RIU values. In Figure 3.8A, there is a noticeable overlap between anti-GT_{1b} and anti-GM_1, as well as between anti-GM_1 and anti-GA_1, but no overlap between anti-GT_{1b} and anti-GA_1. In Figure 3.8B, a similar trend is observed, but the region of identification for 10 ng/mL anti-GT_{1b} interacting with a GM_1 ganglioside surface is very small compared to the dominate regions, which are the anti-GT_{1b} on a GT_{1b} ganglioside surface and anti-GT_{1b} on a GA_1 ganglioside surface. This indicates that the algorithm is able to differentiate between one antibody and three ganglioside surfaces more effectively than three antibodies and one surface. Using the sensorgram data for the three antibodies at 50 ng/mL on a GT_{1b} surface, the model has an accuracy of 83%; and for 10 ng/mL anti-GT_{1b} interacting with three different ganglioside surfaces has an accuracy of 88%. Both accuracies are lower than those of the nnet endpoint data analysis, which can be attributed to the fact that this dataset has a much more complex features due to inclusion of association, steady state, and dissociation patterns.

K-Nearest Neighbor Data Analysis: The kNN algorithm was applied to the same sensorgram dataset with the same training to test ratio of 70 to 30 %. As seen in Figures

3.9C and 3.9D, the most likely paths to occur based upon the specific analyte/antigen interactions from the training datasets are the solid lines. The colored individual data points that surround the predicted paths are the cluster regions where specific analyte/antigen interactions are most likely to occur. Notable confusion can be observed in both Figures 3.9C and 3.9D in relation to the model's clustering analysis capabilities in regards to regions related to association and dissociation interactions. However, in regions related to the steady state and post rinse cycles, the algorithm is highly successful. Even with these regions of high confusion, the model produced high accuracy rates above 90 %. Specifically, we obtained 94 % accuracy for 50 ng/mL of three antibodies on a GT_{1b} ganglioside surface and 96 % for 10 ng/mL anti-GT_{1b} interacting with three different ganglioside surfaces. The higher accuracy for differentiating between ganglioside surfaces instead of antibodies agrees well with the findings of the nnet and PLS-DA, as discussed previously.

The high accuracy observed for the kNN model in comparison to the nnet algorithm can be attributed to the different approach by which the algorithms analyze the data series.[51, 52] Both algorithms identify that there is a higher probability of experiencing confusion between anti-GM_1 with the other two antibodies due to higher cross reactivity. In addition, they agree on no confusion between anti-GT_{1b} and anti-GA_1. Both models also show that differentiating between one antibody and three ganglioside surfaces is much more reliable than differentiating between three antibodies and one ganglioside surface. These findings agree well with the statistical analysis previously discussed to characterize the microarrays ability to differentiate between analyte/antigen interactions. It also shows

that these models are capable of identifying an antibody that is interacting specifically with a ganglioside surface based upon endpoint data, association, steady state, and dissociation kinetics.

Using the sensorgram dataset in combination with the endpoint data appears to be more effective than relying on either one separately for disease biomarker characterization. The combination allows for a more comprehensive review of the observed interactions, and also for faster identification of abnormalities that individual ML algorithms may miss. This is critical to the analysis of patient samples, where there is a high likelihood that unforeseen interactions could have occurred that were not accounted for in the training series, such as those from patient's therapeutics/drugs and other disease biomarkers that are not affiliated with MS.[55] Relying on endpoint data alone may cause higher than expected false positive or negative results due to these unaccounted interactions. We demonstrated that utilizing binding kinetic interactions with the endpoint data, the unaccounted interactions could be more easily identified and then corrected by the models. The findings of this study have the potential to drastically improve label free detection methods as well as the reliability of screening many biomarkers simultaneously.

3.4 Conclusion:

In this work, we have shown that SPRi microarray biochips in combination with robust statistical algorithms are capable of detecting, identifying, and differentiating antibody/ganglioside interactions in whole serum samples. The work addressed a major concern of using antibodies for the diagnosis of MS in a clinical setting, which is the high individual heterogeneity in patients and the widely varied concentration range. We

demonstrated that developing a detection scheme allowing a range of antibodies measured with multiplexed capability, under identical assay conditions, and being able to account for cross reactivity and nonspecific interactions is critical. Using the PFDTS surface, the microarray can be easily extended to include more antigens desired, and the hydrophobic regions will minimize the interferes from the background proteins. This is ideal for glycolipids and sphingomyelins, which are the major components of the myelin sheath. All three targeted antibodies were detectable and quantifiable within biologically relevant concentrations between 3 ng/mL to 25 ng/mL. The statistical analysis and machine learning algorithms implemented in this study allowed for the observation and evaluation of unique trends and features between the antibodies and antigens, which allowed us to conduct a more intense evaluation and gain a broader understanding than what traditional assays can achieve. The method demonstrated here may improve patient-specific evaluation of MS biomarkers, and find use in helping understand the disease progression. As can be seen from these results, robust data analysis protocols are integral for future disease detection studies based on the complexity of biological interactions.

3.5 References:

1. Broza, Y. Y.; Zhou, X.; Yuan, M.; Qu, D.; Zheng, Y.; Vishinkin, R.; Khatib, M.; Wu, W.; Haick, H., Disease Detection with Molecular Biomarkers: From Chemistry of Body Fluids to Nature-Inspired Chemical Sensors. *Chem Rev* **2019,** *119* (22), 11761-11817.

2. Clark, L. F.; Kodadek, T. J. A. c. n., The immune system and neuroinflammation as potential sources of blood-based biomarkers for Alzheimer's disease, Parkinson's disease, and Huntington's disease. *ACS chemical neuroscience* **2016,** *7* (5), 520-527.

3. Graner, M.; Pointon, T.; Manton, S.; Green, M.; Dennison, K.; Davis, M.; Braiotta, G.; Craft, J.; Edwards, T.; Polonsky, B. J. P. o., Oligoclonal IgG antibodies in multiple sclerosis target patient-specific peptides. *PloS one* **2020,** *15* (2), e0228883.

4. Kuerten, S.; Lanz, T. V.; Lingampalli, N.; Lahey, L. J.; Kleinschnitz, C.; Mäurer, M.; Schroeter, M.; Braune, S.; Ziemssen, T.; Ho, P. P. J. P. o. t. N. A. o. S., Autoantibodies against central nervous system antigens in a subset of B cell–dominant multiple sclerosis patients. *Proceedings of the National Academy of Sciences* **2020,** *117* (35), 21512-21518.

5. Mizutani, K.; Oka, N.; Kusunoki, S.; Kaji, R.; Mezaki, T.; Akiguchi, I.; Shibasaki, H. J. J. o. t. n. s., Sensorimotor demyelinating neuropathy with IgM antibody against gangliosides GD1a, GT1b and GM3. *Journal of the neurological sciences* **2001,** *188* (1-2), 9-11.

6. Pender, M. P.; Csurhes, P. A.; Wolfe, N. P.; Hooper, K. D.; Good, M. F.; McCombe, P. A.; Greer, J. M., Increased circulating T cell reactivity to GM3 and GQ1b gangliosides in primary progressive multiple sclerosis. *Journal of clinical neuroscience* **2003,** *10* (1), 63-66.

7. Schnaar, R. L., Brain gangliosides in axon–myelin stability and axon regeneration. *FEBS letters* **2010,** *584* (9), 1741-1747.

8. Cawley, J. L.; Jordan, L. R.; Wittenberg, N. J., Detection and Characterization of Vesicular Gangliosides Binding to Myelin-Associated Glycoprotein on Supported Lipid Bilayers. *Anal Chem* **2021,** *93* (2), 1185-1192.

9. Nowack, L.; Teschers, C. S.; Albrecht, S.; Gilmour, R., Oligodendroglial glycolipids in (Re)myelination: implications for multiple sclerosis research. *Nat Prod Rep* **2021,** *38* (5), 890-904.

10. Ivanova, M.; Zakharova, M. J. H. P., Antibodies against myelin lipids in multiple sclerosis. *Human Physiology* **2017,** *43* (8), 875-880.

11. Höftberger, R.; Guo, Y.; Flanagan, E. P.; Lopez-Chiriboga, A. S.; Endmayr, V.; Hochmeister, S.; Joldic, D.; Pittock, S. J.; Tillema, J. M.; Gorman, M. J. A. n., The pathology of central nervous system inflammatory demyelinating disease accompanying myelin oligodendrocyte glycoprotein autoantibody. *Acta neuropathologica* **2020**, 1-18.

12. Ghasemi, N.; Razavi, S.; Nikzad, E., Multiple Sclerosis: Pathogenesis, Symptoms, Diagnoses and Cell-Based Therapy. *Cell J* **2017,** *19* (1), 1-10.

13. Shedko, E.; Tyumentseva, M. J. N.; Physiology, B., Molecular Biomarkers in the Cerebrospinal Fluid in Multiple Sclerosis. *Neuroscience and Behavioral Physiology* **2020,** *50* (5), 527-533.

14. Costerus, J. M.; Brouwer, M. C.; van de Beek, D., Technological advances and changing indications for lumbar puncture in neurological disorders. *Lancet Neurol* **2018,** *17* (3), 268-278.

15. Derkus, B.; Bozkurt, P. A.; Tulu, M.; Emregul, K. C.; Yucesan, C.; Emregul, E., Simultaneous quantification of Myelin Basic Protein and Tau proteins in cerebrospinal fluid and serum of Multiple Sclerosis patients using nanoimmunosensor. *Biosensors and Bioelectronics* **2017,** *89*, 781-788.

16. Lycke, J.; Zetterberg, H., The role of blood and CSF biomarkers in the evaluation of new treatments against multiple sclerosis. *Expert Rev Clin Immunol* **2017,** *13* (12), 1143-1153.

17. Brownlee, W. J.; Hardy, T. A.; Fazekas, F.; Miller, D. H., Diagnosis of multiple sclerosis: progress and challenges. *Lancet* **2017,** *389* (10076), 1336-1346.

18. Ziemssen, T.; Akgun, K.; Bruck, W., Molecular biomarkers in multiple sclerosis. *J Neuroinflammation* **2019,** *16* (1), 272.

19. Yang, M.; Huang, J.; Fan, J.; Du, J.; Pu, K.; Peng, X., Chemiluminescence for bioimaging and therapeutics: recent advances and challenges. *Chem Soc Rev* **2020,** *49* (19), 6800-6815.

20. Sharafeldin, M.; Davis, J. J., Point of Care Sensors for Infectious Pathogens. *Anal Chem* **2021,** *93* (1), 184-197.

21. Lambert, A.; Yang, Z.; Cheng, W.; Lu, Z.; Liu, Y.; Cheng, Q., Ultrasensitive Detection of Bacterial Protein Toxins on Patterned Microarray via Surface Plasmon Resonance Imaging with Signal Amplification by Conjugate Nanoparticle Clusters. *ACS Sens* **2018,** *3* (9), 1639-1646.

22. Malinick, A. S.; Lambert, A. S.; Stuart, D. D.; Li, B.; Puente, E.; Cheng, Q., Detection of Multiple Sclerosis Biomarkers in Serum by Ganglioside Microarrays and Surface Plasmon Resonance Imaging. *ACS Sens* **2020,** *5* (11), 3617-3626.

23. Mescheriakova, J. Y.; Wong, Y. Y. M.; Runia, T. F.; Jafari, N.; Samijn, J. P.; de Beukelaar, J. W.; Wokke, B. H.; Siepman, T. A.; Hintzen, R. Q. J. J. n., Application of the 2017 revised McDonald criteria for multiple sclerosis to patients with a typical clinically isolated syndrome. *JAMA neurology* **2018,** *75* (11), 1392-1398.

24. Damodaran, V. B.; Murthy, N. S., Bio-inspired strategies for designing antifouling biomaterials. *Biomater Res* **2016,** *20* (1), 18.

25. D'Agata, R.; Bellassai, N.; Jungbluth, V.; Spoto, G. J. P., Recent Advances in Antifouling Materials for Surface Plasmon Resonance Biosensing in Clinical Diagnostics and Food Safety. *Polymers* **2021,** *13* (12), 1929.

26. Lambert, A. S.; Valiulis, S. N.; Malinick, A. S.; Tanabe, I.; Cheng, Q., Plasmonic Biosensing with Aluminum Thin Films under the Kretschmann Configuration. *Anal Chem* **2020,** *92* (13), 8654-8659.

27. Rikkert, L. G.; de Rond, L.; van Dam, A.; van Leeuwen, T. G.; Coumans, F. A. W.; de Reijke, T. M.; Terstappen, L.; Nieuwland, R., Detection of extracellular vesicles in plasma and urine of prostate cancer patients by flow cytometry and surface plasmon resonance imaging. *PLoS One* **2020,** *15* (6), e0233443.

28. Wu, W.; Yu, X.; Wu, J.; Wu, T.; Fan, Y.; Chen, W.; Zhao, M.; Wu, H.; Li, X.; Ding, S. J. B.; Bioelectronics, Surface plasmon resonance imaging-based biosensor for multiplex and ultrasensitive detection of NSCLC-associated exosomal miRNAs using DNA programmed heterostructure of Au-on-Ag. *Biosensors and Bioelectronics* **2021,** *175*, 112835.

29. Nair, S.; Gomez-Cruz, J.; Manjarrez-Hernandez, A.; Ascanio, G.; Sabat, R. G.; Escobedo, C., Rapid label-free detection of intact pathogenic bacteria in situ via surface plasmon resonance imaging enabled by crossed surface relief gratings. *Analyst* **2020,** *145* (6), 2133-2142.

30. Cui, F.; Yue, Y.; Zhang, Y.; Zhang, Z.; Zhou, H. S., Advancing Biosensors with Machine Learning. *ACS Sens* **2020,** *5* (11), 3346-3364.

31. Das, S.; Dey, A.; Pal, A.; Roy, N. J. I. J. o. C. A., Applications of artificial intelligence in machine learning: review and prospect. *International Journal of Computer Applications* **2015,** *115* (9).

32. Volk, M. J.; Lourentzou, I.; Mishra, S.; Vo, L. T.; Zhai, C.; Zhao, H., Biosystems Design by Machine Learning. *ACS Synth Biol* **2020,** *9* (7), 1514-1533.

33. Abbas, A.; Linman, M. J.; Cheng, Q., Patterned resonance plasmonic microarrays for high-performance SPR imaging. *Anal Chem* **2011,** *83* (8), 3147-52.

34. Wilkop, T.; Wang, Z.; Cheng, Q., Analysis of micro-contact printed protein patterns by SPR imaging with a LED light source. *Langmuir* **2004,** *20* (25), 11141-8.

35. Masson, J. F., Surface Plasmon Resonance Clinical Biosensors for Medical Diagnostics. *ACS Sens* **2017,** *2* (1), 16-30.

36. Wanleenuwat, P.; Iwanowski, P., Role of B cells and antibodies in multiple sclerosis. *Mult Scler Relat Disord* **2019,** *36*, 101416.

37. Kappler, K.; Hennet, T., Emergence and significance of carbohydrate-specific antibodies. *Genes Immun* **2020,** *21* (4), 224-239.

38. Hogan, E. L.; Podbielska, M.; O'Keeffe, J., Implications of lymphocyte anergy to glycolipids in multiple sclerosis (MS): iNKT cells may mediate the MS infectious trigger. *Journal of clinical & cellular immunology* **2013,** *4* (3).

39. Häusser-Kinzel, S.; Weber, M. S. J. F. i. i., The role of B cells and antibodies in multiple sclerosis, neuromyelitis optica, and related disorders. *Frontiers in immunology* **2019,** *10*, 201.

40. Malinick, A. S.; Lambert, A. S.; Stuart, D. D.; Li, B.; Puente, E.; Cheng, Q., Detection of Multiple Sclerosis Biomarkers in Serum by Ganglioside Microarrays and Surface Plasmon Resonance Imaging. *ACS Sensors* **2020,** *5* (11), 3617-3626.

41. Miti, A.; Thamm, S.; Müller, P.; Csáki, A.; Fritzsche, W.; Zuccheri, G., A miRNA biosensor based on localized surface plasmon resonance enhanced by surface-bound hybridization chain reaction. *Biosensors and Bioelectronics* **2020,** *167*, 112465.

42. Sguassero, A.; Artiga, Á.; Morasso, C.; Jimenez, R. R.; Rapún, R. M.; Mancuso, R.; Agostini, S.; Hernis, A.; Abols, A.; Linē, A., A simple and universal enzyme-free approach for the detection of multiple microRNAs using a single nanostructured enhancer of surface plasmon resonance imaging. *Analytical bioanalytical chemistry* **2019,** *411* (9), 1873-1885.

43. Mansourian, N.; Rahaie, M.; Hosseini, M., A Nanobiosensor based on fluorescent DNA-hosted silver nanocluster and HCR amplification for detection of MicroRNA involved in progression of multiple sclerosis. *Journal of fluorescence* **2017,** *27* (5), 1679-1685.

44. Bellagha-Chenchah, W.; Sella, C.; Fernandez, F. R.; Peroni, E.; Lolli, F.; Amatore, C.; Thouin, L.; Papini, A., Interactions between human antibodies and synthetic conformational peptide epitopes: innovative approach for electrochemical

detection of biomarkers of multiple sclerosis at platinum electrodes. *Electrochimica Acta* **2015,** *176*, 1239-1247.

45. Real-Fernández, F.; Passalacqua, I.; Peroni, E.; Chelli, M.; Lolli, F.; Papini, A. M.; Rovero, P., Glycopeptide-based antibody detection in multiple sclerosis by surface plasmon resonance. *Sensors* **2012,** *12* (5), 5596-5607.

46. Giuliani, A., The application of principal component analysis to drug discovery and biomedical data. *Drug Discov Today* **2017,** *22* (7), 1069-1076.

47. Wetzel, S. J., Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders. *Phys Rev E* **2017,** *96* (2-1), 022140.

48. Koga, M.; Tatsumoto, M.; Yuki, N.; Hirata, K., Range of cross reactivity of anti-GM1 IgG antibody in Guillain-Barré syndrome. *Journal of Neurology, Neurosurgery Psychiatry* **2001,** *71* (1), 123-124.

49. Lee, L. C.; Liong, C. Y.; Jemain, A. A., Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. *Analyst* **2018,** *143* (15), 3526-3539.

50. Das, S.; Dey, A.; Pal, A.; Roy, N., Applications of artificial intelligence in machine learning: review and prospect. *International Journal of Computer Applications* **2015,** *115* (9).

51. Shahid, N.; Rappon, T.; Berta, W., Applications of artificial neural networks in health care organizational decision-making: A scoping review. *PLoS One* **2019,** *14* (2), e0212356.

52. Peterson, L. E., K-nearest neighbor. *Scholarpedia* **2009,** *4* (2), 1883.

53. Uçar, M. K.; Nour, M.; Sindi, H.; Polat, K., The effect of training and testing process on machine learning in biomedical datasets. *Mathematical Problems in Engineering* **2020,** *2020*.

54. Ying, X. In *An overview of overfitting and its solutions*, Journal of Physics: Conference Series, IOP Publishing: **2019**; p 022022.

55. Peiffer-Smadja, N.; Rawson, T. M.; Ahmad, R.; Buchard, A.; Georgiou, P.; Lescure, F.-X.; Birgand, G.; Holmes, A. H., Machine learning for clinical decision support in infectious diseases: a narrative review of current applications. *Clinical Microbiology Infection* **2020,** *26* (5), 584-595.

56. Tong, X.; Feng, Y.; Li, J. J., Neyman-Pearson classification algorithms and NP receiver operating characteristics. *Sci Adv* **2018,** *4* (2), eaao1659.

Chapter 4: Curved Membrane Mimics for Quantitative Probing of Protein-Membrane Interactions by Surface Plasmon Resonance

4.1 Introduction:

Cellular membranes encircle and reside within all cells, playing crucial aspects in almost all cellular functions. Lipids, being the backbone of cellular membranes, divide the line between extracellular and intracellular interactions and control the packaging and transportation of various biological components across and within cells.[1] Various cellular functions such as cell division, endo- and exocytosis, organelle trafficking, and cell motility rely on the composition, structure, and curvature of cell membranes to function properly.[2] Because of the dynamic and integral roles that lipid membranes play in all organisms, a great deal of research has focused on the development of membrane mimics to characterize and investigate their various fundamental biophysical interactions.[3-6]

To date, the vast majority of membrane mimics have relied on the use of self-assembled monolayers (SAM) and supported lipid bilayers (SLB) to investigate these complex interactions.[4, 6-12] While SAM and SLB are powerful investigation tools, especially as biomimetics, they do come with several drawbacks when investigating biophysical interactions. Most notably is that both SAM and SLB are planar surfaces lacking any curvature, resulting in a misrepresentation when compared to the natural interactions they attempt to mimic.[6] To date, the vast majority of research regarding protein lipid interactions has focused on how the composition of planar lipid membrane mimics affect biophysical interactions.[3, 6, 12] However, it is becoming more evident that curvature plays a vital role in these interactions and needs to be taken into account during these

investigations.[13-15] While SAMs and SLBs have aided in the examination of various biological interactions,[7, 8, 16] there is a great deal of information left to be investigated and interpreted, especially in regards to proteins interacting with curved membranes.[13-15, 17, 18] To date, several methods have been developed to investigate these interactions through the use of unique structures present on the biosensor's surface and/or via specific surface functionalization procedures.[13, 19-28] While these methods have aided in the investigation of curved membranes, they are either limited by the composition of the membrane, the amount of curvature, reproducibility of the platform, and/or the fluidity and malleability of the membrane itself. This has led to a clear need for the development of new methods that can reliably generate reproducible curved membrane surfaces that can be modified with ease.

Developing a reliable curved membrane mimicking platforms, that can easily be incorporated into various analytical methods, is key to probing curvature dependent biophysical interactions. This is important as many biological entities, such as proteins, viruses, and cells, rely on membrane curvature for communication, transportation, and regulation.[15, 17, 18, 29] The ability to reliably generate different angles of curvature that are stable and can be composed of various biological components is necessary to deciphering these interactions.[15, 18, 30] By controlling cellular membrane mimics in terms of composition and curvature, a deeper understanding of disease specific biological interactions can be pursued.[15, 31, 32] The severity of several complex diseases such as cancers, Alzheimer's, Parkinson's, Huntington's disease, multiple sclerosis, muscular dystrophy, and COVID-19 can be linked to various biological interactions that rely on curved membranes.[15, 31, 33-35]

Specifically, curvature sensing proteins, such as alpha synuclein and bridging integrator 1, have been linked to several of the mentioned diseases.[15, 32, 36] Investigating these proteins and others biophysical interactions with curved membranes may aid in the development of new drugs, drug delivery methods, and novel disease detection methods.

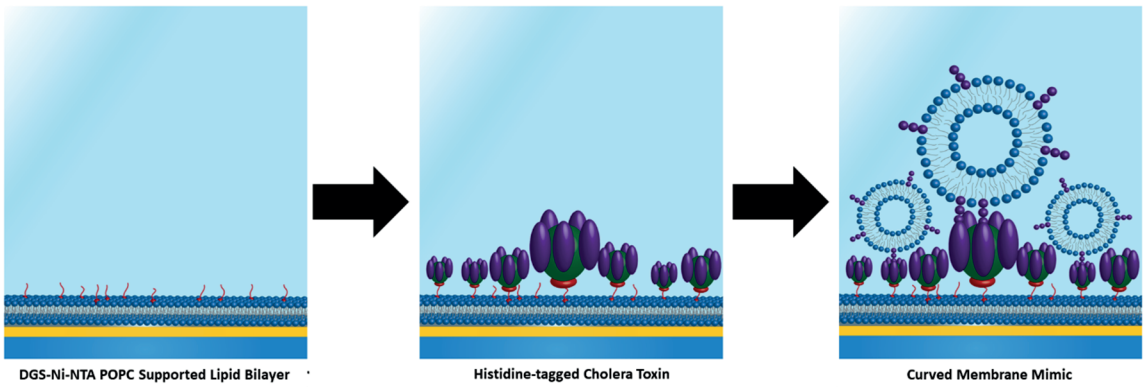


Figure 4.1. Depiction of the formation of a supported lipid bilayer on a gold silicated SPR biochip composed of 5 % DGS-Ni-NTA and 95 % POPC. The surface is then functionalized further with a histidine-tagged cholera toxin (his-CT) to allow for the capture of varying curved membranes as shown in the final portion of the scheme where a vesicle composed of gangliosides is successfully captured onto the surface allowing for the formation of a curved membrane mimic.

Here we report a novel method to create curved membrane surfaces with lipid vesicles ranging in size between 30 nm to 200 nm. The vesicles can easily be modified to incorporate various types of glycans and other membrane components, allowing one to control the composition and size of the curved membrane mimics. Vesicle sizes between 30 nm to 200 nm were selected as this is the common range that endosomes and exosomes exist in, and represent the curvatures that many proteins interact with.[17, 33, 37] Gangliosides were selected as the primary membrane component for investigation as they are present in various highly curved membranes, such as the myelin sheath[7, 8, 21, 31] A range of 0.05 % to 5 % for the gangliosides was selected as these are the biologically relevant percentages.[38,

[39] By changing the type and amount of ganglioside incorporated into different sized curved membranes, an investigation into how these variables affect protein binding can be pursued. To evaluate our platforms ability to identify and differentiate between multivalent or monovalent biophysical interactions cholera toxin (CT) was used. To date CT has been reliably documented in its binding kinetics and biophysical interactions with various gangliosides on planar lipid membrane mimics, but never before with curved membrane mimics. Here, we focus on the characterization of this newly developed platform and present future paths to utilize it in future studies.

4.2 Materials and Methods:

Materials and Reagents: Asialoganglioside GA_1 was acquired from Sigma-Aldrich (St. Louis, MO). Monosialoganglioside GM_1 was purchased from Matreya (Pleasant Gap, PA). Trisialoganglioside GT_{1b} was obtained from Biosynth (Itsaca, Il). Tetrasialoganglioside GQ_{1b} was acquired from Sigma-Aldrich (St. Louis, MO). 5B-sub-unit histidine tagged cholera toxins was purchased from Sigma-Aldrich (St. Louis, MO). 1,2-dioleoyl-sn-glycero-3-[(N-(5-amino-1-carboxypentyl)iminodiacetic acid)succinyl] (nickel salt) DGS-Ni-NTA, 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine POPC, mini stainless steel extruder, 1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine-N-(7-nitro-2-1,3-benzoxadiazol-4-yl) (ammonium salt) NBD, and 30 nm, 100 nm, and 200 nm polycarbonate thin film membranes were all purchased from Avanti Polar Lipids (Alabaster, Al). Premium Plain BK-7 glass microscope slides and phosphate buffered saline (PBS) concentrate were purchased from Fisher Scientific (Pittsburgh, PA).

Lipid Vesicle Preparation: Stocks of 5 mg/mL POPC and 5 mg/mL GA_1, GM_1, GT_{1b}, and GQ_{1b} were diluted in 1:9 methanol chloroform solution to the designated concentration and stored in a −80 °C freezer. Lipid vesicle formation was achieved by aliquoting the lipid stock solution into glass vials and drying under nitrogen to form a thin lipid film which was left to dry for 24 hours in a vacuum desiccator. The dried lipids were resuspended in 1× PBS (10 mM Na2HPO4, 1.8 mM KH2PO4, 137 mM NaCl, 2.7 mM KCL, pH 7.4) to a final concentration of 1 mg/mL. After resuspension the solutions were vigorously vortexed until cloudy after which they would undergo bath sonication for 30 minutes at a constant temperature of 60 °C. The lipid vesicle solutions were then extruded through the desired polycarbonate thin film filter (30 nm, 100 nm, and 200 nm) to produce small unilamellar vesicles of uniform size. All lipid vesicles suspensions were stored at 4 °C and used within 2 days of preparation to ensure consistent vesicle structure.

Nano Tracking Analysis: Lipid vesicle suspensions of 100 nm and 200 nm were diluted from the stock solution of 1 mg/mL down to 10 ug/mL. 30 nm lipid vesicle suspensions were diluted down to a concentration of 500 ug/mL due to presence of high background noise at 10 ug/mL during the NTA experiments. This difference in concentration is why there is a higher intensity for the 30 nm vesicles compared to the 100 nm and 200 nm vesicle samples. Analysis of the vesicles was conducted on a NanoSight NS300 with a flow rate of 5 mL/hr. After completion of the experiments the results of the analysis were graphed in excel.

Fluorescence Recovery After Photobleaching: Fluorescence microscopy and bleaching images were acquired on an inverted Leica TCS SP5 II. For both the SLB and curved membrane mimic investigations a 2 % molar ratio of NBD-PE lipids were incorporated into lipid vesicle preparation methods so that fluorescent visualization of lipid membranes could be achieved. Fluorescently labeled lipids were incorporated into either the underlying lipid membrane or attached vesicles to ensure that fluorescent signal was only attributable to one lipid layer at a time. Excitation of the NBD labeled lipids was achieved with an argon laser (488 nm) at 10-20 % laser power. While, photobleaching was achieved by irradiation of a small area for 1 second at 100 % power using 476 and 488 nm argon laser lines. Recovery images were taken every second with 2-line averaging using the LAS AF software package. After data acquisition the images were processed in ImageJ with the Fiji package and intensity values for bleach and control areas were obtained using a macro. These intensity values were used to calculate fractional recovery following methods demonstrated by Axelrod et al.[40] Diffusion coefficients are not presented due to the large immobile fraction of this system which significantly limited ability to accurately fit recovery curves.

Fabrication of Surface Plasmon Resonance Sensor Chips: Fabrication of the SPR sensor chips was performed following a previous procedure published by our group under clean room protocols in a nanofabrication facility.[8] After the glass microchips were cleaned in piranha solution 2 nm of chromium (0.5 Å/s) followed by 48 nm of gold (2.0 Å/s), were deposited on to the cleaned glass slides via electron beam physical vapor deposition (EBPVD) (Temescal, Berkeley, CA). After deposition the slides were removed

from the EBPVD so that 1-3 nm of SiO_2 could be added onto the gold layer via plasma enhanced chemical vapor deposition (PECVD) using a Unaxis Plasmatherm 790 system (Santa Clara, CA).

SPR Analysis: A NanoSPR5–321 (NanoSPR, Chicago, IL), a dual-channel SPR spectrometer with a GaAs semiconductor laser light source set at a wavelength of 670 nm, was used for all spectroscopic measurements for conventional SPR biosensing. The device utilizes a prism with refractive index of n=1.61 and a 30 μL flow cell. PBS (phosphate buffered saline) running buffer at a pH of 7.4 was used in all experiments with a flow speed of 5 mL/hr.

Statistical Analysis: Principal component analysis (PCA) was accomplished with the prcomp function in R and visualized through the ggbiplot package with an ellipse probability set to 95 % using the collected endpoint data with a total set of over 240 data points. Analysis of variance (ANOVA) was conducted in Excel with the Analysis ToolPak add-in and used the same end point data utilized in PCA. All of the utilized data was found to be statistically relevant.

Monte Carlo Lipid Models: Monte Carlo methods were utilized to iteratively build lipid vesicles with varying ganglioside compositions and to calculate the distance between these randomly distributed gangliosides. This was achieved through a home-built R script that calculates number of lipids within a vesicle of given size and then randomly generates a size appropriate sphere with that number of lipids. The minimum distance between individual gangliosides within this model is then calculated and iterated to achieve

a dataset of over 30,000 lipid distances. Thus, providing an expected ganglioside distribution for each vesicle parameter. The equations used for these calculations were achieved through the R package "rgl"[41] and was employed for 3D visualization of the resulting ganglioside distributed vesicle models.

4.3 Results and Discussion:

Formulation and Characterizations of Curved Membrane Mimics: Curved membrane mimics offer unique research opportunities for the detection of various biomarkers and the investigation of complex biophysical interactions.[42] However, their utilization with SPR based technologies has been severely limited to date.[16, 22-25] To address this, a highly reliable, reproducible, and effective methodology to generate a biomimetic curved membrane platforms was developed and characterized. The process to generate the platform is shown in Figure 4.1 as a cartoon. Figure 4.2 shows the generation of the platform via a SPR spectroscopy sensorgram on a gold silicated biochips with accompanying cartoon illustrations for each step.

The first step in the process to generate the curved biomimetic membrane platform is to create a SLB composed of 5 % DGS-Ni-NTA and 95 % POPC on a silicated gold surface. 5 % DGS-Ni-NTA was selected as it would allow for an abundant amount to be evenly distributed throughout the SLB in an easy to access and readily available manner for the capture of histidine tagged proteins.[6, 43] Once the SLB had fully formed and was stable, the surface was saturated with 10 ug/mL of histidine tagged cholera toxin (his-CT) to ensure all DGS-Ni-NTA sites were occupied. His-CT was selected to act as the support

between our SLB and the ganglioside containing vesicles as CT is known to have very strong binding affinity with the ganglioside GM_1,[44] and high cross reactivity with other gangliosides.[45] The main difference amongst the gangliosides in their respective antigenic binding sites is the number of sialic acids present: GA_1 has no sialic acids, GM_1 has one, GT_{1b} has three, and GQ_{1b} has four.[7, 45-47] The final step in the presented protocol was to inject and incubate 1 mg/mL of 100 nm 1 % GM_1 and 99 % POPC lipid vesicles, thus creating the curved membrane mimicking substrate.

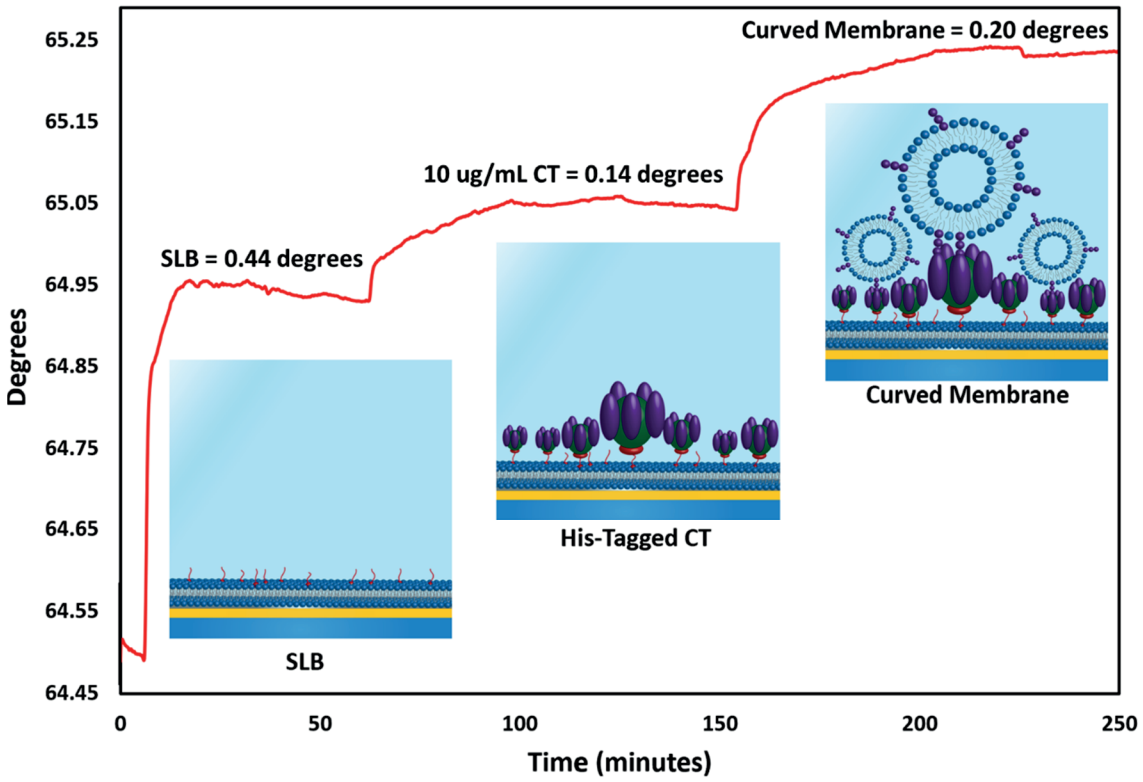


Figure 4.2. Sample SPR sensorgram showing the formation of a 100 nm 1 % GM1 curved membrane mimic with a visual representation for each of the specific steps shown in the sensorgram. First is the formation of the 5 % DGS-Ni-NTA and 95 % POPC SLB, after which 10 ug/mL of his-CT is saturated on the surface so that the curved membrane mimic can form through the interaction with the B-subunit of the his-CT and the GM_1 gangliosides present inside the 1 % GM_1 99 % POPC 100 nm lipid vesicles.

As can be seen in Figure 4.2, the curved membrane mimicking step created a large shift roughly half of that compared to the SLB shift. This led to the question of whether the tethered vesicles were remaining intact or becoming a second SLB on top of the his-CT. To investigate the stability of the SLB with and without his-CT and the tethered vesicles, Fluorescence Recovery After Photobleaching (FRAP) was implemented, as shown in Figure 4.3.

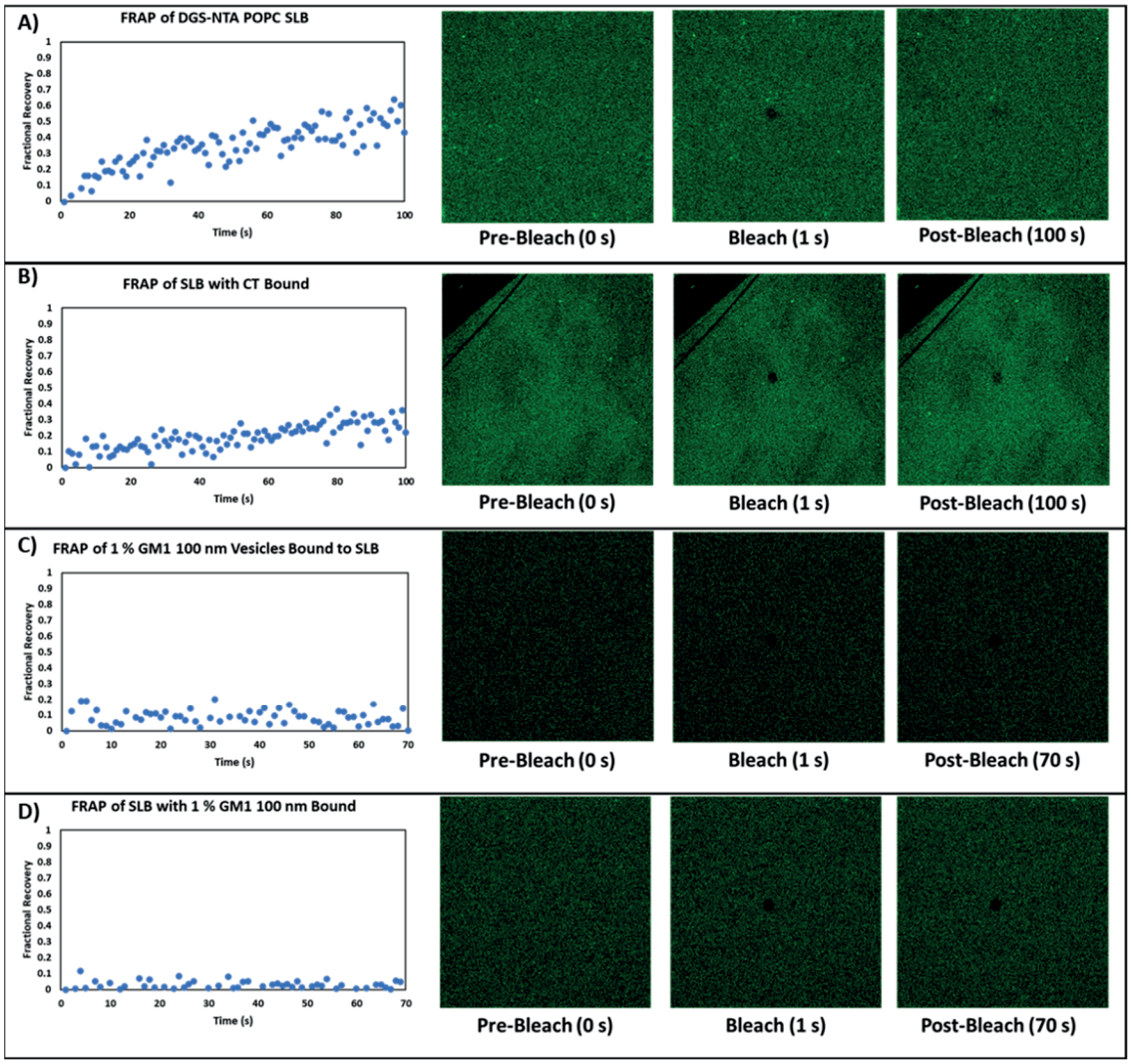


Figure 4.3. Fluorescence Recovery After Photobleaching (FRAP) of A) SLB composed of 5 % DGS-Ni-NTA, 2 % NBD-PE, and 93 % POPC, B) SLB composed of 5 % DGS-Ni-NTA, 2 % NBD-PE, and 93 % POPC with 10 ug/mL of his-CT bound to the DGS-Ni-NTA, C) 1 % GM_1, 2 % NBD-PE, and 97 % POPC 100 nm lipid vesicle onto of the SLB with his-CT present, and D) SLB composed of 5 % DGS-Ni-NTA, 2 % NBD-PE, and 93 % POPC with both 10 ug/mL of his-CT and 1 % GM_1 ganglioside 99 % 100 nm lipid vesicle bound.

Figure 4.3A represents the expected FRAP of a SLB, with lipids that are fluid and obtain substantial recovery over a period of 100s. However, the mobile fraction is substantially lower than that for a purely POPC membrane, which can be attributed to the presence of 5 % DGS-Ni-NTA. Prior literature has shown that the inclusion of large

moieties like DGS-Ni-NTA,[43] GM_1,[48] and proteins[49] can reduce the mobility of SLB. Once the his-CT is bound (Figure 4.3B), a substantial amount of the SLB's fluidity is lost, likely due to the presence of his-CT's extra mass.[49] Figure 4.3C showcases the FRAP of the 1 % GM_1, 2 % NBD-PE, and 97 % POPC 100 nm tethered vesicles attached to the his-CT functionalized SLB, indicating that there is no fluidity. This lack of fluidity indicates that the tethered vesicles retain their curvature to some degree, and do not form a second SLB. While some deformation may be present, it is highly unlikely that enough has occurred to induce vesicle merging. This also indicates that the packing density of the vesicles has not reached a point where vesicle merging is occurring. The final FRAP investigation looked at the underlying SLB with the his-CT and tethered vesicle present, Figure 4.3D. Clearly, the SLB has lost all fluidity and is unable to recover, which agrees with what has been reported in literature and what was expected.[43, 48, 49] Looking at the data as a whole it can be assumed that the his-CT protein layer effectively blocks vesicle rupture and merging between the underlying SLB and the tethered vesicles, thus a curved membrane mimicking platform was successfully generated.

The data gathered from both FRAP, Figure 4.3, and SPR sensorgram data, Figure 4.2, indicated that the creation of a reliable and reproducible curved membrane biomimetic platform can be achieved. However, how changing the composition and size of the vesicles would impact the observed biophysical interactions required further investigation. Of most interest, in regards to his-CT, was whether the degree of monovalent or multivalent interactions could be monitored and/or controlled between the protein and ganglioside containing vesicles. To investigate this, vesicles of specific sizes, 30 nm, 100 nm, and 200

nm, were extruded, and the incorporation of different gangliosides at varying percentages ranging between 0.05 % to 5 % were used, while keeping the SLB composition and amount of his-CT constant.

Curved Membrane Platform Expansion and Capability Evaluation: Before investigating the different sized vesicles and compositions on SPR, it was confirmed that the polycarbonate thin films could reliably generate vesicles of the same size batch to batch. Figure 4.4 highlights the reproducibility of the extruded vesicles for 30 nm, 100 nm, and 200 nm vesicles via nano tracking analysis (NTA). An average of three different batches, each of which underwent three NTA examinations per experiment, were used to determine the average size of the vesicles. The collected results agree well with what has been reported in literature.[16, 22] With the vesicle sizes confirmed, a correlation between vesicle composition and curvature through SPR shifts could be pursued.

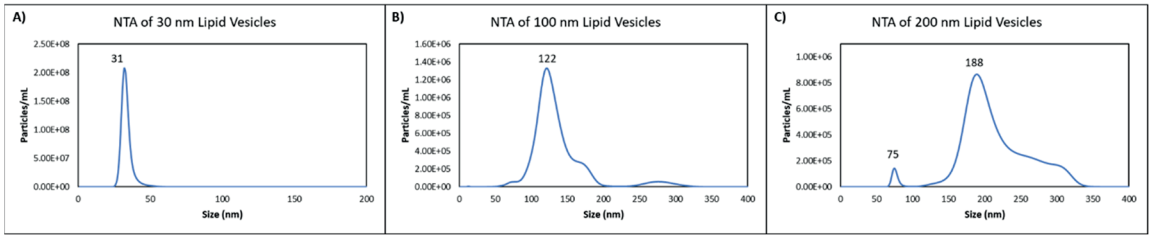


Figure 4.4. Nano Tracking Analysis (NTA) of A) 30 nm, B) 100 nm, and C) 200 nm 1 % GM_1 ganglioside and 99 % POPC lipid vesicles after extrusion. The numbers shown above each peak are the most common values identified in the solution, averaged amongst three different samples. Both the 100 nm and 200 nm vesicles are at a concentration of 10 ug/mL, but the 30 nm were at 500 ug/mL due to the instrument experiencing difficulty at being able to track the small vesicles at 10 ug/mL.

To limit the number of investigated variables the percentage of DGS-Ni-NTA in the SLB and the concentration of his-CT were held constant. Based upon a total of 255 experiments it was found that, the 5 % DGS-Ni-NTA containing SLB had an average SPR

shift of 0.45 degrees with a standard deviation (STD) of 0.01, and for the 10 ug/mL of his-CT an average SPR shift of 0.16 degrees with a STD of 0.02 degrees. To investigate how changing the composition of the vesicles would affect the biophysical interactions with the captured his-CT, GA_1, GM_1, GT_{1b}, and GQ_{1b} gangliosides were incorporate into the 30 nm, 100 nm, and 200 nm vesicles at percentages of 0.05 %, 0.1 %, 1 %, and 5 %. The use of gangliosides offers several unique avenues to explore in the context of diseases, especially to those associated with the central nervous system.[38] However, before this platform can be implemented for the investigation of disease related biophysical interactions, an understanding of how varying the size and composition of the tethered vesicles affects the formation of the curved membrane mimics must be pursued.

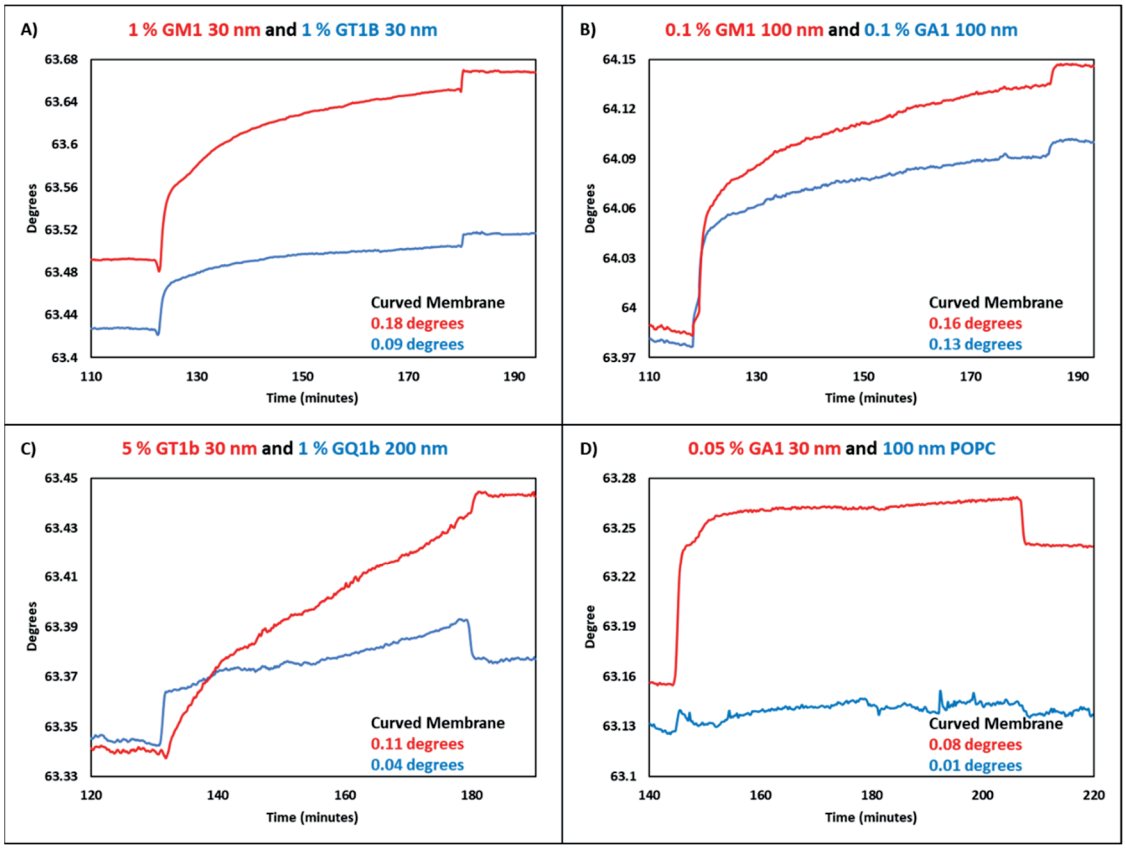


Figure 4.5. Example SPR sensorgrams of creating varying curved membrane surfaces after the 5 % DGS-Ni-NTA and 95 % POPC SLB and 10 ug/mL his-CT were incubated by introducing A) 1 % GM_1 30 nm in red and 1 % GT_{1b} 30 nm in blue, B) 0.1 % GM_1 100 nm in red and 0.1 % GA_1 100 nm in blue, C) 5 % GT_{1b} 30 nm in red and 1 % GQ_{1b} 200 nm in blue, and D) 0.05 % GA_1 30 nm in red and 100 nm POPC vesicles in blue.

Figure 4.5 showcases several examples of how changing the vesicles composition and size affect the observed binding kinetics with his-CT. Comparing 5 % GT_{1b} 30 nm vesicles (Figure 4.5C, red) and 1 % GM_1 30 nm vesicle (Figure 4.5A, red), the ganglioside present in the vesicles can be more significant than having a higher percentage of a ganglioside with known weaker binding affinity. In contrast, when gangliosides that have similar binding affinities with CT are incorporated at similar percentages and in the same vesicles sizes, as shown in Figure 4.5B for GM_1 (red) and GA_1 (blue), the observed binding kinetics are nearly identical. These trends agree well with what has been reported in

literature on both SLB and SAM,[44, 45, 50, 51] but is the first with curved membranes. While the majority of the observed binding interactions can be attributed to which ganglioside is present,[44, 45, 50, 51] it is possible that other factors maybe influencing these observations, such as vesicle size and whether monovalent or multivalent interactions are occurring.

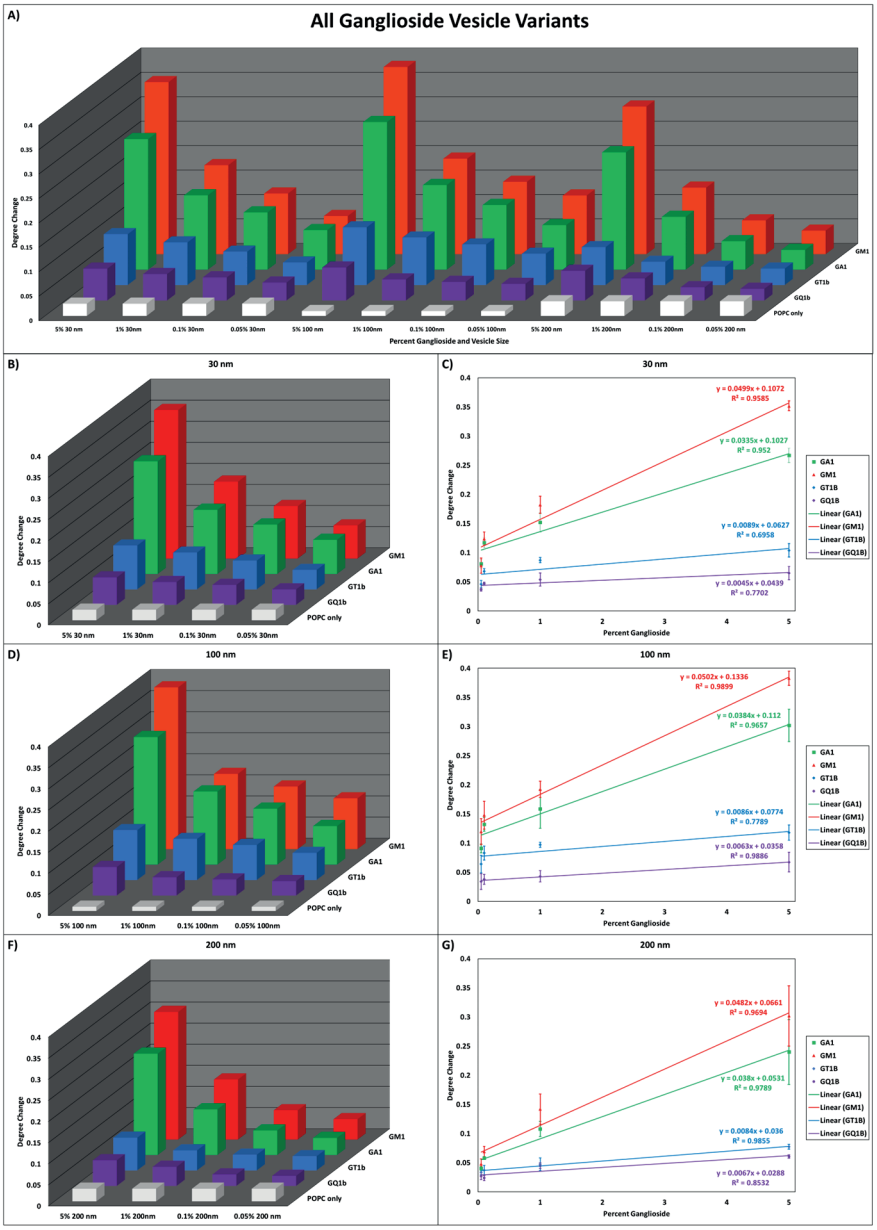


Figure 4.6. A) 3D bar graph showcasing the overall end point data of all investigated interactions white bars are for lipid vesicles containing only POPC, purple is for all GQ_{1b} ganglioside containing lipid vesicles, blue is for all GT_{1b} ganglioside containing lipid vesicle, green is all GA_1 ganglioside containing vesicles, and red is for all GM_1 containing ganglioside vesicles. The 3D bar graphs of B), D), and E) show the individual end point data values for 30 nm, 100 nm, and 200 nm respectively. Scotter plots for C) 30 nm, E) 100 nm, and G) 200 nm represent the linear trends observed for each of the investigated gangliosides at that specific vesicles size at increasing percentages based upon the end point data.

The first step in determining what factors besides binding affinities were influencing the observed biophysical interactions was to plot the collected end point data as a 3D bar graph (Figure 4.6A). Doing so allowed for the observation that the 100 nm vesicles had the largest observable SPR shifts, 30 nm vesicles had the second largest shifts, and 200 nm vesicles had the smallest shifts for all ganglioside types and at each percentage. Due to SPR's sensitivity being reliant on changes in mass and refractive index, it can be assumed that the observed changes from the vesicle sizes can mostly be attributed to varying masses.[4] Because the amount of DGS-Ni-NTA and his-CT were held constant, the number of vesicles that could be tethered should remain constant between vesicle sizes. Thus, the difference between 100 nm and 30 nm vesicles can be assumed to be mostly dependent on the difference in mass between the two. For the 200 nm vesicles the substantial loss in sensitivity, as well as an increase in variation, can be attributed to the 200 nm vesicles likely being slightly outside of the 300 nm detection range of the SPR evanescent wave.[7-9] Based upon these observations the 200 nm tethered vesicles cannot reliable be used for future SPR biosensing applications.

The second step was to separate Figure 4.6A into its respective components based upon vesicle size, as can be seen for 30 nm (Figure 4.6 B and C), 100 nm (Figure 4.6 D and E), and 200 nm (Figure 4.6 F and G) vesicles as 3D bar graphs and scatter plots. Breaking up the data in this manner facilitated the investigation of the importance of which ganglioside was present and the percentage of that ganglioside for each vesicle size. The first observation is that as the antigenic binding sites of the ganglioside differ from GM_1's,

a steady loss in binding affinity can be observed. This trend agrees well with what was previously discussed regarding changes in affinity.[45, 51]

When focusing in on the scatter plots, it can be observed there is a positive linear trend between percent ganglioside and SPR shifts for all of the investigated vesicle sizes. It is important to note that GT_{1b} and GQ_{1b} containing vesicles had the least amount of linearity, which can be attributed to their weaker binding affinities with CT.[45, 51] Nevertheless, the positive linear trends indicate that the percent of gangliosides present in each vesicle, regardless of size and ganglioside incorporated, plays a significant role in the observed biophysical interactions. The most likely explanation for this trend is that the percent ganglioside can be correlated to whether monovalent or multivalent interactions are occurring. However, it is also possible that this trend may indicate that as the percentage of ganglioside increases the chances for vesicle deformation to occur also increases. Investigating deformation of the tethered vesicle is outside the scope of SPR's capabilities and would require the use of other methods, such as quartz crystal microbalance with dissipation (QCM-D), to be implemented. QCM-D has previously been used to monitor deformations in planar membrane mimics,[52] and will be applied in a similar fashion for the tethered vesicles in future experiments.

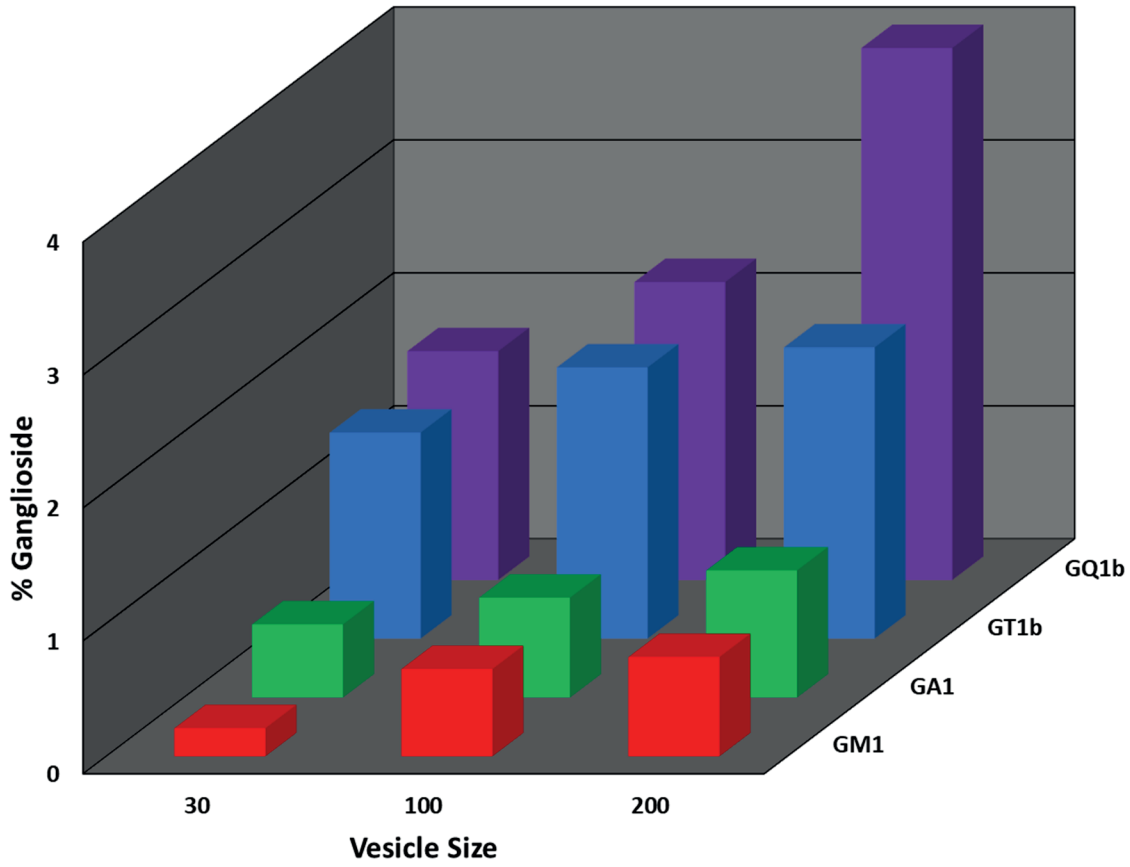


Figure 4.7. Visual representation through the use of a 3D bar graphs showcasing the limits of detection (LOD) for the varying vesicle sizes and ganglioside compositions calculated based upon the calibration curves shown in Figure 4.6. Red is for GM_1, green is for GA_1, blue is for GT_{1b}, and purple for GQ_{1b} for left to right 30 nm, 100 nm, and 200 nm lipid vesicles.

To determine if the linear trends could be used as a means to identify between monovalent or multivalent interactions, the lowest percentage for each ganglioside needed to tether the vesicles were calculated. The 3D bar graph in Figure 4.7 visualizes the calculated limits of detection (LOD), via the 3σ method, for each of the investigated gangliosides and vesicle sizes. Based upon the calculated LOD's it was found that, for GM_1 containing vesicles the minimum percent of gangliosides needed for 30 nm vesicles to be tethered was 0.40 %, for 100 nm vesicle was 0.66 %, and for 200 nm vesicles was 0.75 %.

For GA_1 the LODs were calculated to be 0.55 %, 0.75 %, and 0.96 % for 30 nm, 100 nm, and 200 nm vesicles respectively. GT_{1b} containing vesicles had LODs of 1.29 % for 30 nm vesicles, 1.57 % for 100 nm vesicles, and 1.58 % for 200 nm vesicles. For GQ_{1b} containing vesicles the calculated LODs were 1.55 %, 2.61 %, and 3.95 % for 30 nm, 100 nm, and 200 nm vesicles respectively. These values can be used to determine the likelihood that either monovalent or multivalent interactions are occurring. However, to confidently make this claim the application of robust post data acquisition analysis tools are needed, as well as taking into account the pentameric structure of CT's five B subunits.[53, 54]

Based upon the collected observations and taking into account vesicle size and affinity, it was hypothesized values below the LOD can be associated with monovalent interactions, and those above the LOD are multivalent interactions. With this hypothesis, it needed to be determined whether the type of ganglioside, percent of ganglioside, or vesicle size were the leading causes of the observed biophysical interactions. To achieve this several statistical analysis tools were implemented including analysis of variance (ANOVA), Principal Component Analysis (PCA), and Monte Carlo simulations.

Mathematical Modeling and Statistical Analysis of Vesicle Structure and Biophysical Interactions: PCA was implemented to gain a deeper understanding of the significance that the size of the vesicles, ganglioside present, and percent composition of that ganglioside played in the observed biophysical interactions. Before PCA could be used however, the data needed to be shown to have no internal bias and be statistically relevant. To achieve this, analysis of variance (ANOVA) was applied to the end point data, as shown in Table 4.1.

Table 4.1. ANOVA data for all investigated gangliosides and vesicles sizes.

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
5 % GA1 30 nm	5	1.335	0.267	0.000145
5 % GM1 30 nm	5	1.76	0.352	7.00E-05
5 % GT1b 30 nm	5	0.52	0.104	0.00013
5 % GQ1b 30 nm	5	0.325	0.065	0.000125

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.27499	3	0.091663333	780.1134752	1.50E-17	3.238871517
Within Groups	0.00188	16	0.0001175			
Total	0.27687	19				

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
1 % GA1 30 nm	5	0.76	0.152	0.00027
1 % GM1 30 nm	5	0.91	0.182	0.00022
1 % GT1b 30 nm	5	0.436	0.0872	0.0000192
1 % GQ1b 30 nm	5	0.27	0.054	0.00013

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.0514704	3	0.0171568	107.3642053	8.22E-11	3.238871517
Within Groups	0.0025568	16	0.0001598			
Total	0.0540272	19				

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
0.1 % GA1 30nm	5	0.585	0.117	0.00007
0.1 % GM1 30nm	5	0.62	0.124	0.00013
0.1 % GT1b 30nm	5	0.341	0.0682	0.0000217
0.1 % GQ1b 30nm	5	0.236	0.0472	3.70E-06

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.0209442	3	0.0069814	123.8935226	2.76E-11	3.238871517
Within Groups	0.0009016	16	0.00005635			
Total	0.0218458	19				

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
0.05 % GA1 30nm	5	0.404	0.0808	0.0000557
0.05 % GM1 30nm	5	0.39	0.078	0.00017
0.05 % GT1b 30nm	5	0.23	0.046	0.0000425
0.05 % GQ1b 30nm	5	0.183	0.0366	0.0000058

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.00749855	3	0.002499517	36.4892944	2.20E-07	3.238871517
Within Groups	0.001096	16	0.0000685			
Total	0.00859455	19				

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
5 % GA1 100 nm	5	1.51	0.302	0.00077
5 % GM1 100 nm	5	1.915	0.383	0.000145
5 % GT1b 100 nm	5	0.59	0.118	0.00017
5 % GQ1b 100 nm	5	0.337	0.0674	0.0002838

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.3348036	3	0.1116012	326.1285798	1.49E-14	3.238871517
Within Groups	0.0054752	16	0.0003422			
Total	0.3402788	19				

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
1 % GA1 100nm	5	0.793	0.1586	0.0010898
1 % GM1 100nm	5	0.962	0.1924	0.0001888
1 % GT1b 100nm	5	0.486	0.0972	0.0000192
1 % GQ1b 100nm	5	0.215	0.043	9.50E-05

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.065746	3	0.021915333	62.93892399	4.47E-09	3.238871517
Within Groups	0.0055712	16	0.0003482			
Total	0.0713172	19				

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
0.1 % GA1 100nm	5	0.661	0.1322	0.0001402
0.1 % GM1 100nm	5	0.74	0.148	0.00057
0.1 % GT1b 100nm	5	0.415	0.083	0.000145
0.1 % GQ1b 100nm	5	0.19	0.038	0.00007

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.0373674	3	0.0124558	53.8512754	1.39E-08	3.238871517
Within Groups	0.0037008	16	0.0002313			
Total	0.0410682	19				

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
0.05 % GA1 100nm	5	0.455	0.091	0.000055
0.05 % GM1 100nm	5	0.6	0.12	0.0005
0.05 % GT1b 100nm	5	0.32	0.064	0.00023
0.05 % GQ1b 100nm	5	0.17	0.034	0.00018

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.02031375	3	0.00677125	28.06735751	1.31E-06	3.238871517
Within Groups	0.00386	16	0.00024125			
Total	0.02417375	19				

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
5 % GA1 200 nm	5	1.2	0.24	0.0031
5 % GM1 200 nm	5	1.51	0.302	0.00267
5 % GT1b 200 nm	5	0.388	0.0775	0.00001825
5 % GQ1b 200 nm	5	0.304	0.0607	8.23E-06

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.21413365	3	0.071377883	49.25601975	2.65E-08	3.238871517
Within Groups	0.02318592	16	0.00144912			
Total	0.23731957	19				

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
1 % GA1 200nm	5	0.54	0.108	0.00017
1 % GM1 200nm	5	0.71	0.142	0.00067
1 % GT1b 200nm	5	0.237	0.0474	0.0001188
1 % GQ1b 200 nm	5	0.225	0.045	0.000025

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.0339516	3	0.0113172	46.01423053	4.31E-08	3.238871517
Within Groups	0.0039352	16	0.00024595			
Total	0.0378868	19				

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
0.1 % GA1 200nm	5	0.29	0.058	6.50E-06
0.1 % GM1 200nm	5	0.345	0.069	8.00E-05
0.1 % GT1b 200nm	5	0.186	0.0372	6.92E-05
0.1 % GQ1b 200nm	5	0.135	0.027	0.00002

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.0054924	3	0.0018308	41.6801366	8.68E-08	3.238871517
Within Groups	0.0007028	16	0.000043925			
Total	0.0061952	19				

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
0.05 % GA1 200 nm	5	0.201	0.0402	0.0000127
0.05 % GM1 200 nm	5	0.24	0.048	7.00E-05
0.05 % GQ1b 200 nm	5	0.168	0.0336	0.0000173
0.05 % GT1b 200 nm	5	0.119	0.0238	3.22E-05

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.001578	3	0.000526	15.91527988	4.63E-05	3.238871517
Within Groups	0.0005288	16	0.00003305			
Total	0.0021068	19				

By holding the vesicle sizes constant as shown in Figures 4.8A, 4.8B, and 4.8C for 30 nm, 100 nm, and 200 nm lipid vesicles respectively, the significance of ganglioside species and percentage was achieved. PCA determined that the dominant component for differentiation was which ganglioside was incorporated into the vesicle, which was expected due to the importance of binding affinities. However, the percentage of the ganglioside was found to be the most significant factor for the observed misidentification. Minor confusion occurs between GA_1 and GM_1 containing vesicles at percentages of 0.05 % for 30 nm and 100 nm lipid vesicle and below 0.1 % for 200 nm vesicles. GM_1 experienced no confusion with either GT_{1b} nor GQ_{1b} for any sized vesicles. GA_1 and GT_{1b} containing vesicles experienced considerable confusion for 100 nm vesicle below 5 %, though no confusion occurred in 30 nm and 200 nm lipid vesicles. GT_{1b} and GQ_{1b} experienced significant confusion at concentrations below 1 % for 100 nm lipid vesicles, confusion at 0.05 % for 200 nm vesicles, and no confusion at 30 nm vesicles. The observed confusion between the gangliosides and varying percentages can largely be attributed to the high cross reactivity they each have with CT.[45, 46, 51]

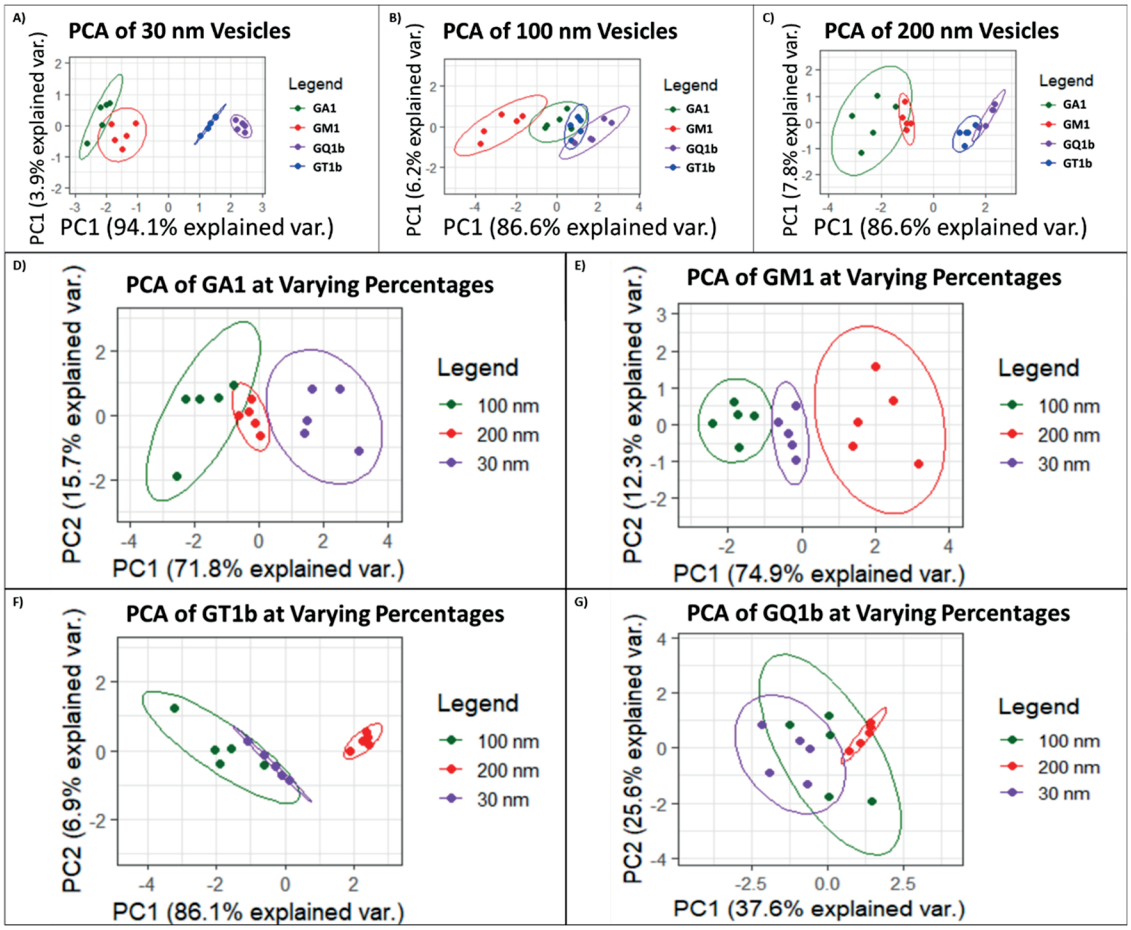


Figure 4.8. PCA of all collected data based upon the control factor for the vesicles where A) is comparing variance based upon 30 nm vesicles and ganglioside present, B) 100 nm vesicles and ganglioside present, and C) 200 nm vesicles and ganglioside present, D) percentage GA_1 ganglioside and size of vesicles, E) percentage of GM_1 ganglioside and size of vesicles, F) percentage of GT_{1b} ganglioside and size of vesicles, and G) percentage of GQ_{1b} ganglioside and size of vesicles.

Figures 4.8D – 4.8G for GA_1, GM_1, GT_{1b}, and GQ_{1b} respectively showcase the significance of vesicle size and ganglioside percentage. In this dataset, it was found that the dominant factor was the percentage of that specific ganglioside over the vesicle's size. GA_1 and GM_1 containing vesicles can reliably be differentiated by vesicle size up to 95 % confidence. This is of great interest as it shows that SPR has the potential to be utilized as a new characterization tool for nanostructures functionalized onto a surface, which to date

has not been extensively investigated. For GT_{1b} and GQ_{1b} differentiation of the vesicle sizes proved to be considerably difficult for PCA and is most likely due to their substantially weaker binding affinity with CT.[45, 46, 51] However, as the percentages of GT_{1b} and GQ_{1b} increased differentiation began to be observed, which indicates that their interactions with CT rely heavily on multivalent interactions.

The collected SPR data and presented PCAs clearly indicated that the most important features in order of significance were which ganglioside was present, the percentage of that ganglioside, and the size of the vesicles. However, the importance of ganglioside type and percent ganglioside present, led to the question of whether there was a missing component in the characterization of these biophysical interactions. Looking back at the results for PCA and LOD, it became apparent that an evaluation of the statistical likelihood for multivalent or monovalent interactions to occur was needed. To determine this required calculating the arc distances for the investigated ganglioside percentage at each vesicles size in a three dimensional (3D) space, while also taking into account the structure of the antigenic binding domain of CT.

To visualize and investigate the distribution of the gangliosides in a 3D space, Monte Carlo simulations were implemented. Monte Carlo simulations are essentially a set of algorithms that rely on repeated random sampling techniques to calculate a numerical result of a specific parameter. The utilization of Monte Carlo methods are most often applied for probability distributions, optimization, or numerical integration.[55, 56] These simulations are useful for building iterative random models that at large scales provide

important information about the system in question, especially in the context of spatial distributions.[55, 56]

The Monte Carlo simulations used in this study were designed to calculate the arc distance between the investigated ganglioside percentages and vesicles sizes. This approach allowed for the investigation of whether monovalent or multivalent interactions were occurring between a single vesicle and a captured his-CT. To achieve this thousands of vesicles were modeled for each vesicle size and ganglioside percentage, after which each simulation was compiled together to obtain a robust arc length distribution. Using the arc distance distributions and the structure of CT's pentamer GM_1 binding domain an investigation of whether monovalent or multivalent interactions are occurring could be pursued.

The structure of CT is composed of one A subunit and five B subunits. The A subunit of CT is the enzymatic portion of the protein, which has a mass of 28 kDa, and is not present in the his-CT used in the presented study, as it only contains the five subunit B portion of CT. The five B subunits, each of which have an individual mass of 11 kDa and a combined mass of 55 kDa, are the receptor binding portions of CT, which has evolved to interact with the sialic acid containing lipid GM_1.[53] To obtain the spatial distribution of CT's five B subunit required looking into literature. Recently, it was reported both through simulations and X-ray crystallography, that CT's pentavalent B subunits are arranged in such a way that they have a 3-nm spacing between one another.[54] This information was used in combination with the Monte Carlo simulations to determine whether monovalent or multivalent interactions were occurring.

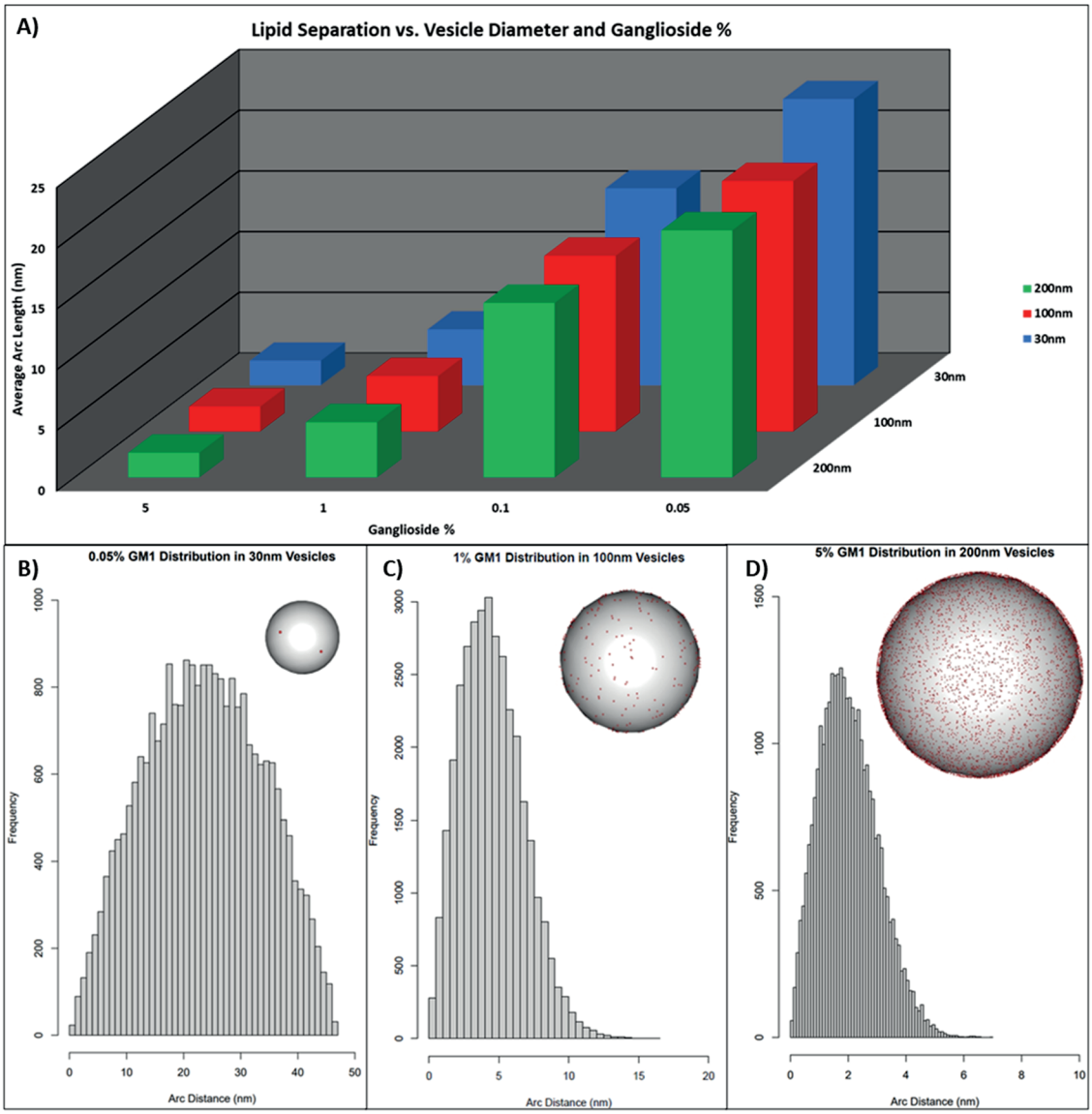


Figure 4.9. A) overall representation of average arc length relation to ganglioside % in each of the three investigated vesicle sizes 30 nm blue, 100 nm red, and 200 nm green, B) example of arc distance results of a 0.05 % ganglioside distributed in a 30 nm vesicle, C) arc distance distribution of a 1 % ganglioside distribution in a 100 nm vesicle, and D) example of 5 % ganglioside in a 200 nm vesicle. The red dots indicate the potential location of GM_1 gangliosides in one iteration taken at a specific time point.

With the calculated average arc distances for every vesicle size and ganglioside percentage investigated in this study, shown as a 3D bar graph in Figure 4.9 and the Monte Carlo Simulations in Figure 4.10, we can determine how the gangliosides present on the

surface of the vesicles would interact with the captured his-CT. The average calculated arc distances for 30 nm lipid vesicles according to the Monte Carlo simulations were 2.03 nm, 4.59 nm, 16.20 nm, and 23.56 nm for 5 %, 1 %, 0.1 %, and 0.05 % respectively. For 100 nm vesicles it was found that the average arc distances were 2.02 nm, 4.53 nm, 14.43 nm, and 20.59 nm at 5 %, 1 %, 0.1 %, and 0.05 % respectively. When 200 nm vesicles were investigated it was found the average arc distances were 2.02 nm, 4.53 nm, 14.16 nm, and 20.33 nm for 5 %, 1 %, 0.1 %, and 0.05 % respectively.

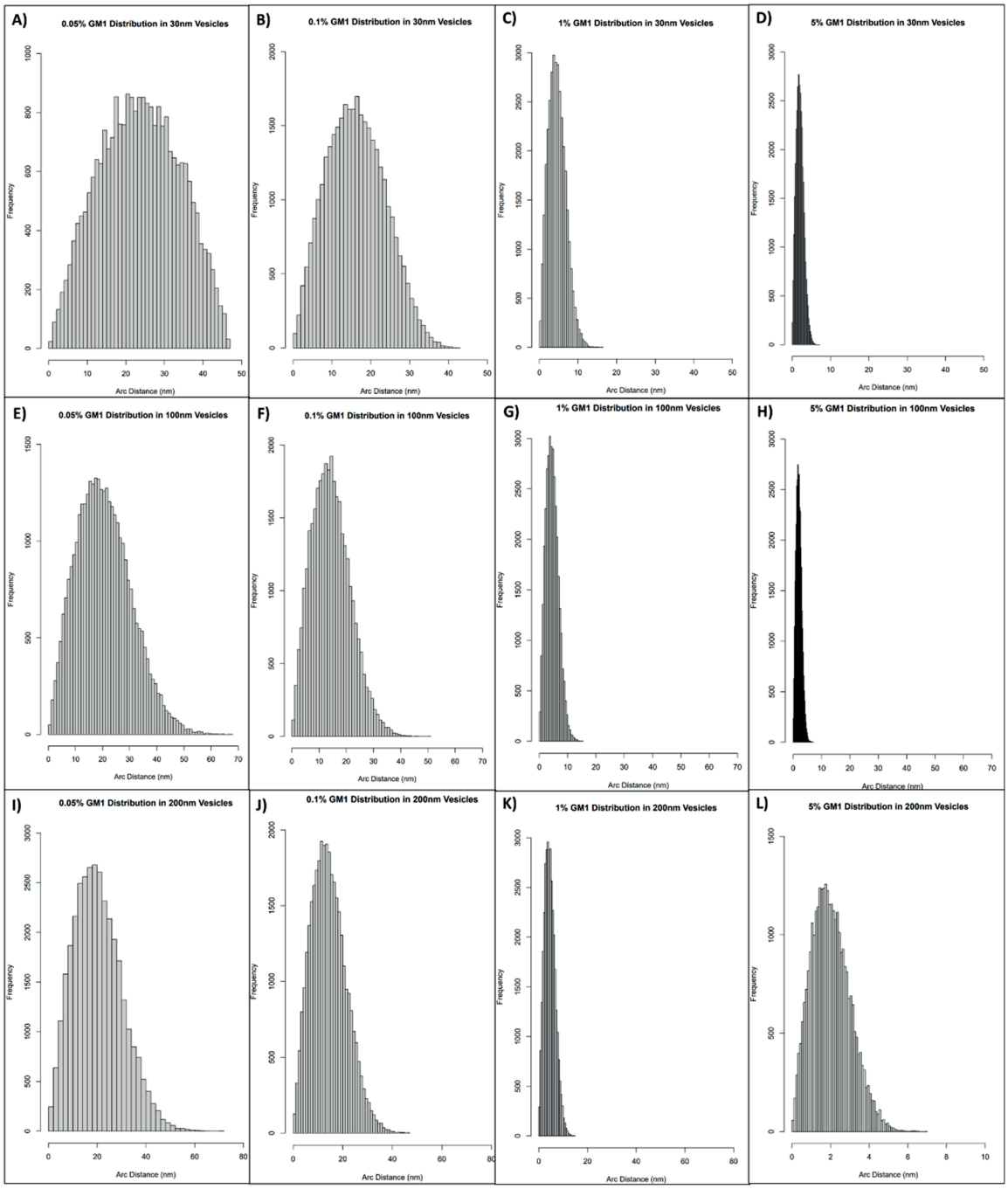


Figure 4.10. All average arc distances calculated using Monte Carlo methods. Varying percentages for 30 nm are shown from A to D, E to H are for 100 nm vesicles at different percentages, and I to L are for 200 nm vesicles.

Based upon the collected Monte Carlo data and the assumption that the five B subunits of CT are 3 nm apart it can be inferred that all five of B subunits would be occupied

at 5 % for each of the vesicle sizes. For 1 % ganglioside containing vesicles statistically three to four of the five B subunits would be occupied at every vesicle size. At 0.1 % statistically one to two B subunits could be occupied at any given point, and at 0.05 % it is guaranteed that only monovalent interactions would be occurring. These results agree well with the proposed LOD calculations for GA_1 and GM_1 to predict whether multivalent or monovalent interactions were occurring, but did not agree well with the calculated LOD's for GQ_{1b} and GT_{1b} at 1 % for all vesicles sizes. The observed discrepancy between the Monte Carlo simulations and the calculated LOD's for GQ_{1b} and GT_{1b} at 1 % occupancy for all vesicle sizes can be attributed to the weaker binding affinities with CT.[47] This further indicated the importance of the antigenic head group of the molecule being incorporated into the vesicles. Therefore it can be concluded, that the lower binding affinity of GT_{1b} and GQ_{1b} are the dominating factors at play even if multivalent interactions are occurring. This observation agrees well with the observed SPR sensorgrams, shown in Figure 4.5, and the end point data shown in Figure 4.6A. In addition to agreeing with the LODs, the Monte Carlo results agrees well with the PCA's conclusion that the primary cause for the observed biophysical interactions is the percentage of ganglioside over the vesicle size.

With the information gained from PCA, SPR, and Monte Carlo simulations it can be concluded that the leading biophysical interactions are the binding affinities of the gangliosides and whether or not multivalent interactions are occurring. However, these observations are only true for CT and may be drastically different for other proteins. For this reason, we plan to focus future efforts on investigating known curvature dependent

biophysical interactions, such as those affiliated with bridging integrator 1 and alpha-synuclein.[15, 18]

4.4 Conclusion:

The presented work demonstrated that the combination of SPR with statistical analysis tools can be a powerful investigation strategy to decipher complex biophysical interactions. In addition, the creation and characterization of a novel biomimetic curved membrane platform was presented. The developed platform was shown to be able to reliably generate various curved membrane mimics in terms of compositions and curvature sizes, while maintaining structural stability. The combination of this platform and statistics facilitated the identification of whether monovalent or multivalent interactions were occurring with his-CT. This platform should facilitate the investigation of various analyte antigen interactions that rely on curvature but to date have only been investigated with planar substrates. The information that could be collected from this platform may drastically aid in the development of new drugs, drug delivery methods, and the development of new disease detection strategies. Future work will focus on using this platform to characterize curvature sensing proteins in biological matrices in a manner that has not been pursued before.

4.5 References:

1. Giacomello, M.; Pyakurel, A.; Glytsou, C.; Scorrano, L., The cell biology of mitochondrial membrane dynamics. *Nature reviews Molecular cell biology* **2020,** *21* (4), 204-224.

2. De Carvalho, C. C.; Caramujo, M. J., The various roles of fatty acids. *Molecules* **2018,** *23* (10), 2583.

3. Luchini, A.; Vitiello, G., Mimicking the mammalian plasma membrane: An overview of lipid membrane models for biophysical studies. *J Biomimetics* **2020,** *6* (1), 3.

4. Hinman, S. S.; McKeating, K. S.; Cheng, Q., Surface plasmon resonance: material and interface design for universal accessibility. *Analytical chemistry* **2018,** *90* (1), 19.

5. Baek, J. M.; Ryu, Y.-S., Surface sensitive analysis device using model membrane and challenges for biosensor-chip. *BioChip Journal* **2020,** *14* (1), 110-123.

6. Jackman, J. A.; Cho, N.-J., Supported lipid bilayer formation: beyond vesicle fusion. *Langmuir* **2020,** *36* (6), 1387-1400.

7. Malinick, A. S.; Lambert, A. S.; Stuart, D. D.; Li, B.; Puente, E.; Cheng, Q., Detection of multiple sclerosis biomarkers in serum by ganglioside microarrays and surface plasmon resonance imaging. *ACS sensors* **2020,** *5* (11), 3617-3626.

8. Malinick, A. S.; Stuart, D. D.; Lambert, A. S.; Cheng, Q., Surface plasmon resonance imaging (SPRi) in combination with machine learning for microarray analysis of multiple sclerosis biomarkers in whole serum. *Biosensors Bioelectronics: X* **2022,** *10*, 100127.

9. Hinman, S. S.; Ruiz, C. J.; Drakakaki, G.; Wilkop, T. E.; Cheng, Q., On-demand formation of supported lipid membrane arrays by trehalose-assisted vesicle delivery for SPR imaging. *ACS applied materials Interfaces* **2015,** *7* (31), 17122-17130.

10. Hinman, S. S.; Ruiz, C. J.; Drakakaki, G.; Wilkop, T. E.; Cheng, Q., On-demand formation of supported lipid membrane arrays by trehalose-assisted vesicle delivery for SPR imaging. *ACS applied materials interfaces* **2015,** *7* (31), 17122-17130.

11. Hinman, S. S.; Ruiz, C. J.; Cao, Y.; Ma, M. C.; Tang, J.; Laurini, E.; Posocco, P.; Giorgio, S.; Pricl, S.; Peng, L., Mix and match: coassembly of amphiphilic dendrimers and phospholipids creates robust, modular, and controllable interfaces. *ACS applied materials interfaces* **2017,** *9* (1), 1029-1035.

12. Mazur, F.; Bally, M.; Städler, B.; Chandrawati, R., Liposomes and lipid bilayers in biosensors. *Advances in colloid interface science* **2017,** *249*, 88-99.

13. Callan-Jones, A.; Bassereau, P., Curvature-driven membrane lipid and protein distribution. *Current Opinion in Solid State Materials Science* **2013,** *17* (4), 143-150.

14. Harayama, T.; Riezman, H., Understanding the diversity of membrane lipid composition. *Nature reviews Molecular cell biology* **2018,** *19* (5), 281-296.

15. Oliveira, L.; Gasser, T.; Edwards, R.; Zweckstetter, M.; Melki, R.; Stefanis, L.; Lashuel, H. A.; Sulzer, D.; Vekrellis, K.; Halliday, G. M., Alpha-synuclein research: Defining strategic moves in the battle against Parkinson's disease. *npj Parkinson's Disease* **2021,** *7* (1), 1-23.

16. Yang, Z.; Malinick, A. S.; Yang, T.; Cheng, W.; Cheng, Q., Gold nanoparticle-coupled liposomes for enhanced plasmonic biosensing. *Sensors Actuators Reports* **2020,** *2* (1), 100023.

17. Simunovic, M.; Evergren, E.; Callan-Jones, A.; Bassereau, P., Curving cells inside and out: roles of BAR domain proteins in membrane shaping and its cellular implications. *Annual review of cell developmental biology* **2019,** *35*, 111-129.

18. Hsieh, W.-T.; Hsu, C.-J.; Capraro, B. R.; Wu, T.; Chen, C.-M.; Yang, S.; Baumgart, T., Curvature sorting of peripheral proteins on solid-supported wavy membranes. *Langmuir* **2012,** *28* (35), 12838-12843.

19. Sundh, M.; Svedhem, S.; Sutherland, D. S., Formation of supported lipid bilayers at surfaces with controlled curvatures: Influence of lipid charge. *The Journal of Physical Chemistry B* **2011,** *115* (24), 7838-7848.

20. Belkilani, M.; Farre, C.; Chevalier, Y.; Minot, S.; Bessueille, F.; Abdelghani, A.; Jaffrezic-Renault, N.; Chaix, C., Mechanisms of Influenza Virus HA2 Peptide Interaction with Liposomes Studied by Dual-Wavelength MP-SPR. *ACS Applied Materials Interfaces* **2022**.

21. Cawley, J. L.; Jordan, L. R.; Wittenberg, N. J., Detection and characterization of vesicular gangliosides binding to myelin-associated glycoprotein on supported lipid bilayers. *Analytical Chemistry* **2020,** *93* (2), 1185-1192.

22. Park, H.; Sut, T. N.; Yoon, B. K.; Zhdanov, V. P.; Cho, N.-J.; Jackman, J. A., Unraveling how multivalency triggers shape deformation of sub-100 nm lipid vesicles. *The Journal of Physical Chemistry Letters* **2021,** *12* (28), 6722-6729.

23. Belkilani, M.; Shokouhi, M.; Farre, C.; Chevalier, Y.; Minot, S.; Bessueille, F.; Abdelghani, A.; Jaffrezic-Renault, N.; Chaix, C., Surface plasmon resonance monitoring

of mono-rhamnolipid interaction with phospholipid-based liposomes. *Langmuir* **2021,** *37* (26), 7975-7985.

24. Norling, K.; Sjöberg, M.; Bally, M.; Zhdanov, V. P.; Parveen, N.; Höök, F., Dissimilar Deformation of Fluid-and Gel-Phase Liposomes upon Multivalent Interaction with Cell Membrane Mimics Revealed Using Dual-Wavelength Surface Plasmon Resonance. *Langmuir* **2022,** *38* (8), 2550-2560.

25. Bompard, J.; Maniti, O.; Aboukhachfe, R.; Ausserre, D.; Girard-Egrot, A., BALM: Watching the Formation of Tethered Bilayer Lipid Membranes with submicron lateral resolution. *Langmuir* **2021,** *37* (31), 9457-9471.

26. Black, J. C.; Cheney, P. P.; Campbell, T.; Knowles, M. K., Membrane curvature based lipid sorting using a nanoparticle patterned substrate. *Soft Matter* **2014,** *10* (12), 2016-2023.

27. Cheney, P. P.; Weisgerber, A. W.; Feuerbach, A. M.; Knowles, M. K., Single lipid molecule dynamics on supported lipid bilayers with membrane curvature. *Membranes* **2017,** *7* (1), 15.

28. Parthasarathy, R.; Yu, C.-h.; Groves, J. T., Curvature-modulated phase separation in lipid bilayer membranes. *Langmuir* **2006,** *22* (11), 5095-5099.

29. Bozelli Jr, J. C.; Epand, R. M., Membrane shape and the regulation of biological processes. *Journal of molecular biology* **2020,** *432* (18), 5124-5136.

30. Casares, D.; Escribá, P. V.; Rosselló, C. A., Membrane lipid composition: effect on membrane and organelle structure, function and compartmentalization and therapeutic avenues. *International journal of molecular sciences* **2019,** *20* (9), 2167.

31. Lauwers, E.; Goodchild, R.; Verstreken, P., Membrane lipids in presynaptic function and disease. *Neuron* **2016,** *90* (1), 11-25.

32. Liu, S.; Xiong, X.; Zhao, X.; Yang, X.; Wang, H., F-BAR family proteins, emerging regulators for cell membrane dynamic changes—from structure to human diseases. *Journal of hematology oncology* **2015,** *8* (1), 1-14.

33. Terakawa, M. S.; Lin, Y.; Kinoshita, M.; Kanemura, S.; Itoh, D.; Sugiki, T.; Okumura, M.; Ramamoorthy, A.; Lee, Y.-H., Impact of membrane curvature on amyloid aggregation. *Biochimica et Biophysica Acta -Biomembranes* **2018,** *1860* (9), 1741-1764.

34. Chng, C.-P.; Cho, N.-J.; Hsia, K. J.; Huang, C., Role of Membrane Stretch in Adsorption of Antiviral Peptides onto Lipid Membranes and Membrane Pore Formation. *Langmuir* **2021,** *37* (45), 13390-13398.

35. Zhao, M.; Maani, N.; Dowling, J. J., Dynamin 2 (DNM2) as cause of, and modifier for, human neuromuscular disease. *Neurotherapeutics* **2018,** *15*, 966-975.

36. Colina-Tenorio, L.; Horten, P.; Pfanner, N.; Rampelt, H., Shaping the mitochondrial inner membrane in health and disease. *Journal of Internal Medicine* **2020,** *287* (6), 645-664.

37. Pegtel, D. M.; Gould, S., Exosomes. *Annual review of biochemistry* **2019,** *88*, 487-514.

38. Ledeen, R.; Wu, G., Gangliosides of the nervous system. *Gangliosides* **2018**, 19-55.

39. Cochran Jr, F. B.; Yu, R. K.; Ledeen, R. W., Myelin gangliosides in vertebrates. *Journal of neurochemistry* **1982,** *39* (3), 773-779.

40. D. AXELROD, D. E. K., J. SCHLESSINGER, E. ELSON, and W. W. WEBB, MOBILITY MEASUREMENT BY ANALYSIS OF FLUORESCENCE PHOTOBLEACHING RECOVERY KINETICS. *BIOPHYSICAL JOURNAL* **1976,** *16*.

41. Adler, D. M. a. D. *rgl: 3D Visualization Using OpenGL*, 2022.

42. Olsén, E.; Jõemetsa, S.; González, A.; Joyce, P.; Zhdanov, V. P.; Midtvedt, D.; Hook, F., Diffusion of lipid nanovesicles bound to a lipid membrane is associated with the partial-slip boundary condition. *Nano letters* **2021,** *21* (19), 8503-8509.

43. Zou, H.; Wu, J.; Fang, Y., [The preparation of supported lipid bilayers and its functional study]. *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi* **2019,** *36* (1), 85-93.

44. Iwamori, M.; Takamizawa, K.; Momoeda, M.; Iwamori, Y.; Taketani, Y., Gangliosides in human, cow and goat milk, and their abilities as to neutralization of cholera toxin and botulinum type A neurotoxin. *Glycoconjugate journal* **2008,** *25* (7), 675-683.

45. Kuziemko, G. M.; Stroh, M.; Stevens, R. C., Cholera toxin binding affinity and specificity for gangliosides determined by surface plasmon resonance. *Biochemistry* **1996,** *35* (20), 6375-6384.

46. Iwamori, M.; Takamizawa, K.; Momoeda, M.; Iwamori, Y.; Taketani, Y., Gangliosides in human, cow and goat milk, and their abilities as to neutralization of cholera toxin and botulinum type A neurotoxin. *Glycoconjugate journal* **2008,** *25* (7), 675-683.

47. Angström, J.; Teneberg, S.; Karlsson, K.-A., Delineation and comparison of ganglioside-binding epitopes for the toxins of Vibrio cholerae, Escherichia coli, and

Clostridium tetani: evidence for overlapping epitopes. *Proceedings of the National Academy of Sciences* **1994,** *91* (25), 11859-11863.

48. B. GOINS, M. M., B. G. BARISAS, AND E. FREIRE, LATERAL DIFFUSION OF GANGLIOSIDE GMl IN PHOSPHOLIPID BILAYER MEMBRANES. *Biophysical Journal* **1986,** *49* (4), 849-856.

49. Yamazaki, V.; Sirenko, O.; Schafer, R. J.; Groves, J. T., Lipid mobility and molecular binding in fluid lipid membranes. *J Am Chem Soc* **2005,** *127* (9), 2826-2827.

50. Fukuta, S.; Magnani, J. L.; Twiddy, E. M.; Holmes, R. K.; Ginsburg, V., Comparison of the carbohydrate-binding specificities of cholera toxin and Escherichia coli heat-labile enterotoxins LTh-I, LT-IIa, and LT-IIb. *Infection immunity* **1988,** *56* (7), 1748-1753.

51. Terrettaz, S.; Stora, T.; Duschl, C.; Vogel, H., Protein binding to supported lipid membranes: investigation of the cholera toxin-ganglioside interaction by simultaneous impedance spectroscopy and surface plasmon resonance. *Langmuir* **1993,** *9* (5), 1361-1369.

52. Seantier, B.; Breffa, C.; Felix, O.; Decher, G., Dissipation-enhanced quartz crystal microbalance studies on the experimental parameters controlling the formation of supported lipid bilayers. *The Journal of Physical Chemistry B* **2005,** *109* (46), 21755-21765.

53. Zhang, R.-G.; Scott, D. L.; Westbrook, M. L.; Nance, S.; Spangler, B. D.; Shipley, G. G.; Westbrook, E. M., The three-dimensional crystal structure of cholera toxin. *Journal of molecular biology* **1995,** *251* (4), 563-573.

54. Groza, R.; Ewers, H., Membrane deformation by the cholera toxin beta subunit requires more than one binding site. *Proceedings of the National Academy of Sciences* **2020,** *117* (30), 17467-17469.

55. Chan, W. K. V., *Theory and applications of Monte Carlo simulations*. BoD–Books on Demand: **2013**.

56. Harrison, R. L. In *Introduction to monte carlo simulation*, AIP conference proceedings, American Institute of Physics: **2010**; pp 17-21.

Chapter 5: Curvature-Tuning Membranes to Probe Affinity Fluctuation of Bridging Integrator 1 Protein

5.1 Introduction:

Over the last few decades many diseases at the forefront of modern medical investigations, such as various types of cancers,[1] neurological disorders,[2] cardiovascular,[3] skin,[4] and musculoskeletal diseases,[5] have been linked to curvature sensing proteins.[6] Due to the integral roles that curvature sensing proteins play in various biological interactions,[7-9] disruptions to their binding are believed to be a major component in these diseases.[10-12] While our understanding of the interrelations and biophysical behaviors between membranes and curvature sensing proteins is still growing,[13-15] several of these proteins have been affiliated with specific diseases.[10, 11, 16] This has led to a great deal of interest in using curvature sensing proteins as disease biomarkers and as drug targets.[11, 13] While the list of known curvature sensing proteins is still growing,[10, 17-19] there is one group of proteins that are repeatedly found to be integral in many biological systems, specifically Bin-amphiphysin-Rvs (BAR) domain containing proteins (BAR-proteins).[11, 16, 20-22]

BAR-proteins are known to bind directly to cellular membrane surfaces,[22] and aid in various cellular functions.[21] While our knowledge of BAR-proteins is arguably still in its infancy, several key features have been identified that are central to the activity of all BAR-proteins.[23] First, they sense membrane curvature, and are preferential to tubular or spherical membranes, like the myelin sheath, and have a clear preference for highly curved membranes.[24] Second, BAR-proteins wrap around membranes in the direction of the membrane's curvature.[16] Third, BAR-proteins prefer negatively charged membranes over

positively charged or neutral membranes,[25] and finally BAR-proteins act as recruitment sites for other types of proteins.[22]

To date, many diseases have been affiliated with specific BAR-protein-membrane and BAR-protein-protein interactions.[11, 26, 27] However, the investigation of BAR-proteins' interactions with membranes and other proteins have largely been qualitative and/or inconclusive.[23] This can largely be attributed to the fact that the majority of these studies have relied on either cell based assays or planar membrane mimics.[28-33] While cell based assays allow for curvature to be accounted for, the amount of curvature and the composition of the membranes cannot be reliable controlled. In regards to planar membrane mimics, such as supported lipid bilayers (SLB), membrane composition can be effectively controlled, but due to the lack of curvature, conclusions of these studies are highly questionable in terms of curvature sensing proteins.[34] While both methods have provided vital information regarding our understanding of BAR-proteins, their inability to reliably control membrane composition and curvature has led to a clear need for the use of new ways to investigate curvature sensing proteins.[28-33, 35-38] To address this, we implemented our previously developed and characterized curved membrane mimicking platform, and used the BAR-protein, amphiphysin II or bridging integrator 1 (BIN1), as a model system, as shown in Figure 5.1.

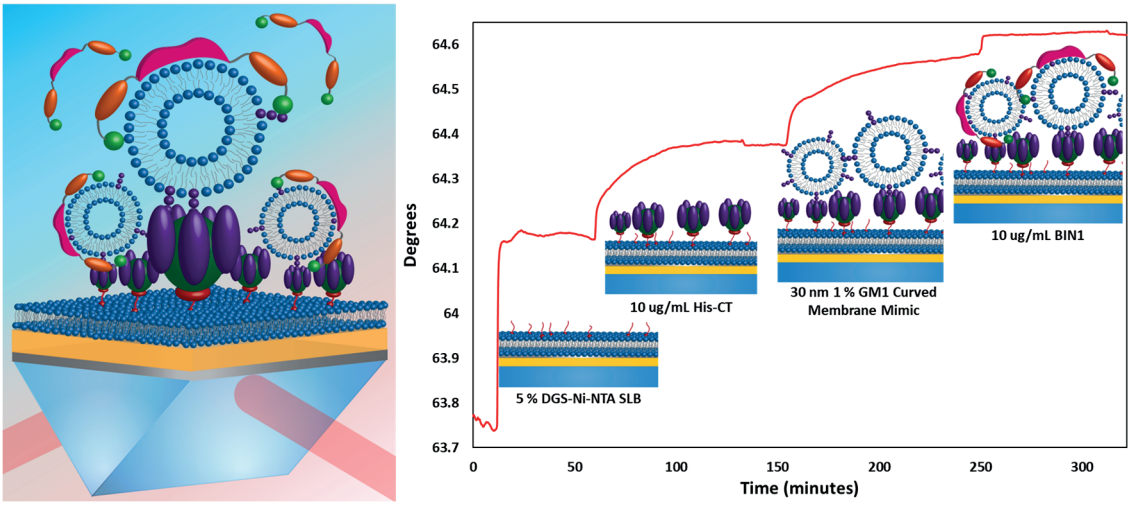


Figure 5.1. Graphic illustration and SPR spectroscopy sensorgram BIN1 binding to a curved membrane mimics. The SPR sensorgram is of 10 ug/mL BIN1 binding to a 1 % GM_1 ganglioside containing 30 nm curved membrane mimic.

BIN1 was selected as a model system as it has been linked to several highly disruptive diseases and is integral for various cellular functions, especially those related to the central nervous system (CNS).[39, 40] Cellular functions that BIN1 is affiliated with include the formation of exosomes, membrane electrical stability, transporting material inside and out of cells, and the remyelination process.[39, 40] While BIN1 has been affiliated with these functions, the way that curvature and membrane composition affect these functions is not well understood. Notable diseases that BIN1 has been affiliated with due to its role in these cellular functions include: ovarian, breast, skin, prostate, and lung cancers as well as neurological, cardiovascular, skin, and musculoskeletal diseases.[41-47] Recently, it was reported that the presence of BIN1 in urine can be used as an indicator of muscular dystrophy,[48] making it a strong biomarker candidate.[49-52] However, a disease relevant concentration in urine for BIN1 to act as a biomarker has not been identified to date.[35, 53] This missing information can be attributed to the lack of a reliable method to investigate BIN1's biophysical interactions in a quantifiable manner.

Here an investigation into the significance of membrane curvature and composition regarding BIN1's biophysical interactions is explored, and we show that BIN1 can be reliably detected in urine as a means to aid in the detection of muscular dystrophy. To investigate the significance of membrane composition and curvature, GA_1 and GM_1 gangliosides were incorporated into 30 nm, 100 nm, and planar membrane mimics. The combination of different gangliosides and curvature sizes facilitated the quantitative investigation of BIN1's biophysical interactions in a highly controllable format. GA_1 and GM_1 gangliosides were selected as they are integral components of various cell membranes throughout the CNS that BIN1 is known to interact with. The incorporation of gangliosides also allowed us to account for membrane charge, as GM_1 has a negative charge due to the presence of a sialic acid in its' antigenic head group, whereas GA_1 is neutral due to the lack of a sialic acid.[54] 30 nm and 100 nm tethered vesicles were selected as these are approximate to the average size of curved membranes that BIN1 interacts with in the CNS. A planar SLB was used to validate BIN1's preference for highly curved membranes compared to the tethered vesicles. In addition, we investigated the reliability of our platform for the detection of BIN1 in urine. To achieve this, the ideal curved membrane platform was identified and implemented. The developed platform was shown to be able to quantifiably investigate the significance of membrane curvature and composition in a manner that has not been achieved before. This platform also demonstrates a feasible method to screen for muscular dystrophy by monitoring the concentration of BIN1 in urine.

5.2 Materials and Methods:

Materials and Reagents: Asialoganglioside GA_1 was acquired from Sigma-Aldrich (St. Louis, MO). Monosialoganglioside GM_1 was purchased from Matreya (Pleasant Gap, PA). B-sub-unit histidine tagged cholera toxins was purchased from Sigma-Aldrich (St. Louis, MO). 1,2-dioleoyl-sn-glycero-3-[(N-(5-amino-1-carboxypentyl)iminodiacetic acid)succinyl] (nickel salt) DGS-Ni-NTA, 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine POPC, mini stainless steel extruder, 1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine-N-(7-nitro-2-1,3-benzoxadiazol-4-yl) (ammonium salt) NBD, and 30 nm, 100 nm, and 200 nm polycarbonate thin film membranes were all purchased from Avanti Polar Lipids (Alabaster, Al). Premium Plain BK-7 glass microscope slides and phosphate buffered saline (PBS) concentrate were purchased from Fisher Scientific (Pittsburgh, PA). Bridging integrator-1 protein (Bin-1) was purchased from Abcam (Cambridge, MA). Non histidine tagged Cholera Toxin Subunit B was purchased from Fischer Scientific (Waltham, MA).

Lipid Vesicle Preparation: Stock solutions of 5 mg/mL POPC and 5 mg/mL GA_1 and GM_1 were diluted in 1:9 methanol chloroform solution to the designated concentration and stored in a −80 °C freezer. Lipid vesicle formation was accomplished by aliquoting the desired volume of each lipid stock solution to obtain the desired % ganglioside POPC ratio into small glass vials which were then dried under nitrogen to form a thin lipid film. The dried lipid film was further left to dry for 24 hours under vacuum in desiccator. The dried lipids were then resuspended in 1× PBS (10 mM Na2HPO4, 1.8 mM KH2PO4, 137 mM

NaCl, 2.7 mM KCL, pH 7.4) to a final concentration of 1 mg/mL. After resuspension the solutions were vortexed until the solution was cloudy after which the lipid solution underwent a sonication bath for 30 minutes at a constant temperature of 60 °C. The lipid vesicle solutions were then extruded with the polycarbonate thin film filter (30 nm and 100 nm) to produce small unilamellar vesicles of uniform size. All lipid vesicles suspensions were stored at 4 °C and used within 2 days of preparation to ensure consistent vesicle structure.

Fabrication of Surface Plasmon Resonance Sensor Chips: Each glass slide was cleaned for 30 minutes in piranha solution at a constant temperate before being functionalized into SPR spectroscopy sensor chips. Chip fabrication was achieved by following the previously developed protocol in our lab under clean room protocols in a nanofabrication facility.[55] 2 nm of chromium (0.5 Å/s) followed by 48 nm of gold (2.0 Å/s), were deposited on to the cleaned glass slides via electron beam physical vapor deposition (EBPVD) (Temescal, Berkeley, CA). Once the two metals were deposited onto the glass slides they were removed from the EBPVD so that 1-3 nm of SiO_2 could be added onto the thin gold layer via plasma enhanced chemical vapor deposition (PECVD) using a Unaxis Plasmatherm 790 system (Santa Clara, CA).

SPR Analysis: SPR spectroscopy experiments were conducted with a NanoSPR5–321 (NanoSPR, Chicago, IL), a dual-channel SPR spectrometer with a GaAs semiconductor laser light source set at a wavelength of 670 nm. All experiments used the same prism with refractive index of n=1.61 and a 30 μL flow cell chamber. The running

buffer for all experiments was PBS (phosphate buffered saline) at a pH of 7.4 and a flow rate of 5 mL/hr.

Statistical Analysis: Principal component analysis (PCA) was achieved via the use of the prcomp function in R and visualized through the ggbiplot package with a set ellipse probability of 95 % using the collected endpoint data with. Analysis of variance (ANOVA) was achieved in Excel with through the Analysis ToolPak add-in and used the same end point data utilized in PCA. All of the collected data was found to be statistically relevant.

Urine Preparation: The artificial urine used to spike BIN1 in for biological matrix biosensing experiments were prepared according to a previously published protocol.[56] Each of the component chemicals were added as solids at the desired concentrations and dissolved in ultrapure DI H_2O at 38 ° C under continuous stirring. The solutions pH was measured to 6.0 ± 0.1 by a UB-5 pH meter (Denver Instruments, Arvada, CO), and solutions were kept no longer than 5 days.

5.3 Results and Discussion:

Membrane Curvature: Because of BIN1's preference for negatively charged membranes,[39] as well as its association with the CNS,[46] an investigation into its interactions with different types of membrane mimics containing GM_1 was pursued first. GM_1 is one of the most common gangliosides present throughout the CNS,[12, 54] and offers the ability to investigate BIN1's preference for negatively charged membranes due to the presence of a sialic acid in GM_1's antigenic head group.[55] Investigating BIN1's interactions with

different membrane curvatures relied on the platform that was discussed, characterized, and investigated in Chapter 4. Figures 5.1 and 5.2 both show the formation of the curved membrane platform and interactions of BIN1 with 30 nm, Figure 5.1, and 100 nm, Figure 5.2 (red), 1 % GM_1 99 % POPC tethered vesicles, as well as with a 1 % GM_1 99 % POPC SLB, Figure 5.2 (blue).

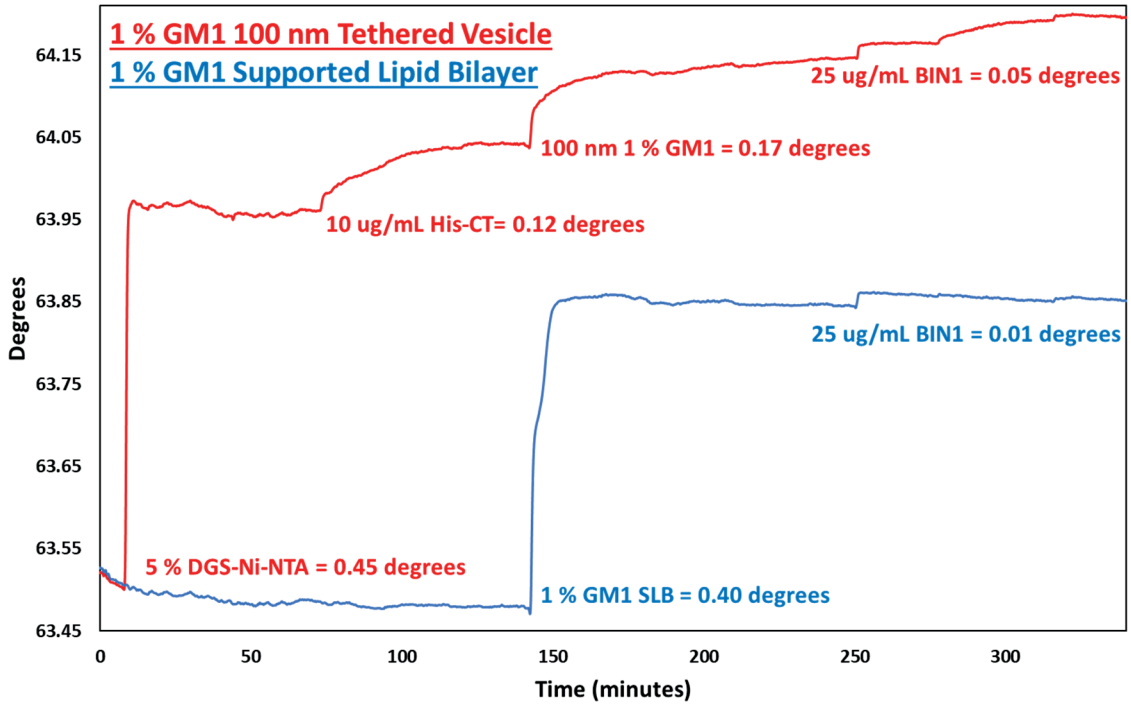


Figure 5.2. Comparison of 25 ug/mL BIN1 interactions with a 1 % GM_1 100 nm tethered curved membrane mimic (red) and a 1 % GM_1 supported lipid bilayer (blue).

As can be seen in Figure 5.2, a comparison of 25 ug/mL BIN1 binding with a 1 % GM_1 100 nm curved membrane surfaces and a 1 % GM_1 SLB is shown. BIN1 has a clear preference for the curved membrane surface as it had a fivefold larger signal response compared to the planar SLB. BIN1's interactions with a 30 nm curved membrane surface, Figure 5.1, was found to have a twenty fold larger signal response compared to the GM_1

SLB. A calibration curve of BIN1 binding to the three different biomimetic membrane surfaces is shown in Figure 5.3.

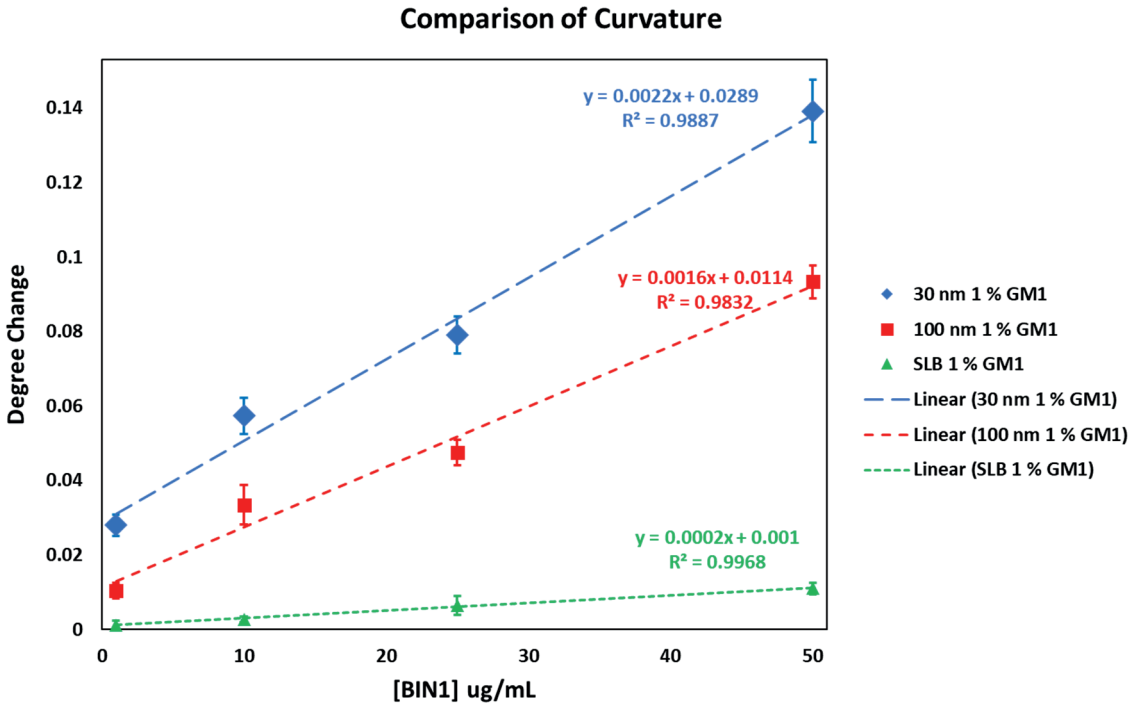


Figure 5.3. Calibration curves for BIN1 at concentrations ranging from 1 to 50 ug/mL in PBS interacting with 30 nm (blue), 100 nm (red), and supported lipid bilayer (green) 1 % GM_1 99 % POPC containing biomimetic membrane mimics.

Based upon the calibration curves, and the calculated limits of detection (LOD), determined by the 3σ method, which were 1.33 ug/mL (21 nM), 3.84 ug/mL (62 nM), and 17.23 ug/mL (278 nM) for the 30 nm, 100 nm, and SLB membrane mimics respectively, three major observations can be made. First, BIN1 has a clear preference for membranes with higher curvature, which agrees well with literature.[39] Second, based upon the calibration curve for 1 % GM_1 SLB it appears that BIN1's signal response between 25 ug/mL and 50 ug/mL reaches a saturation point. While this is most likely due to electrostatic interactions between BIN1 with GM_1 and/or the zwitterionic lipid POPC, it is possible that BIN1 has preferential binding to GM_1. Third, the tethered vesicles maintain

their curvature to some degree. This is substantial as in Chapter 4 there was only qualitative evidence, through recovery after photobleaching, that a second SLB did not form on top of the his-CT.

While the collected data appears to confirm that BIN1 has a clear preference for highly curved membranes further validation was needed. The major concern was that the observed binding interactions were not due to differences in curvature of the tethered vesicles, but due to distance from the SPR active substrate. As was discussed previously in this Dissertation, the SPR evanescent wave extends roughly 300 nm off of the SPR active substrate. While in theory both the 30 nm and 100 nm tethered vesicles should be well within this 300 nm detection range, experimental confirmation to determine the sensitivity of the developed platform was needed.

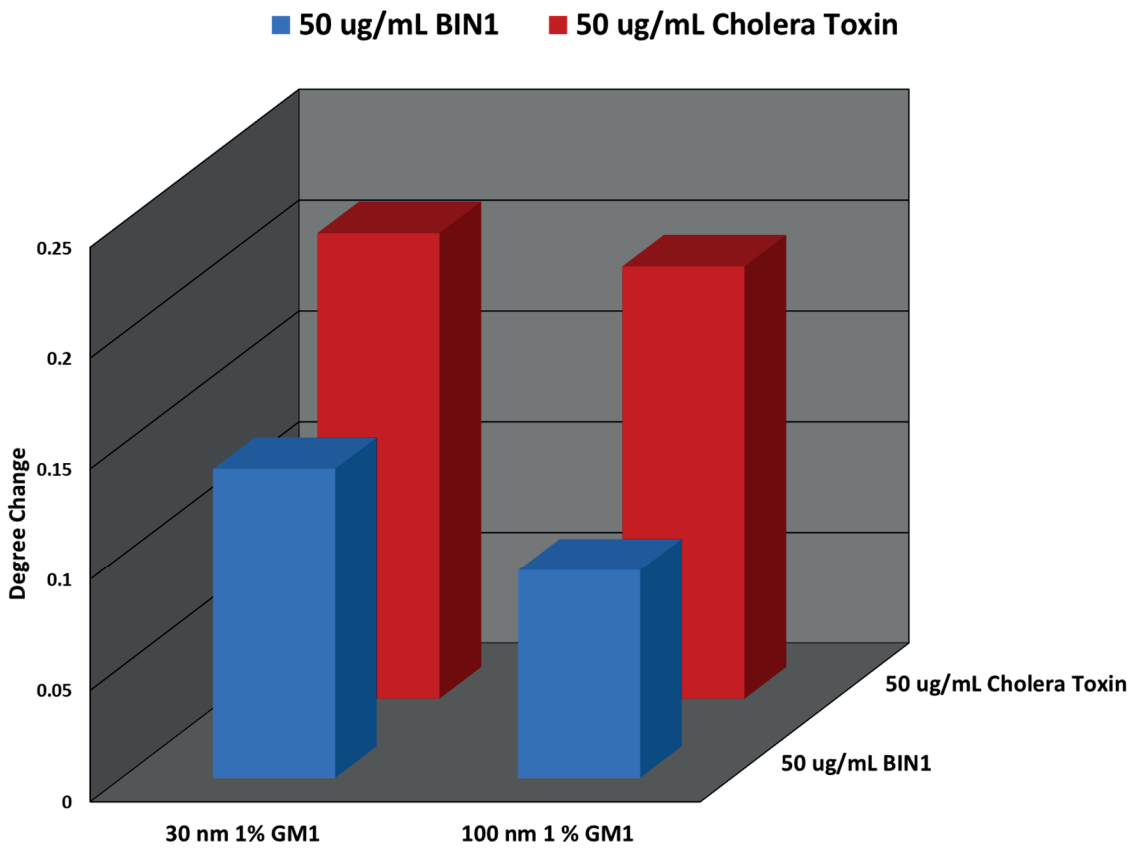


Figure 5.4. 3D bar graph of the endpoint data for 50 ug/mL BIN1 (blue) and 50 ug/mL cholera toxin (red).

To confirm that the observed differences in BIN1's binding were due to its preference for more highly curved membranes and not distance from the surface, a protein that's binding is independent of curvature was needed. Because the tethered vesicles already had GM_1 present in them, the use of a protein that had high specificity and binding affinity to GM_1 was ideal. Based upon these desirable features, cholera toxin (CT) was selected. By comparing the interactions of the 30 nm and 100 nm curved membrane mimics with CT we found that there was a 7 % reduction in signal due to distance from the substrate. In contrast, when a comparison between BIN1 with the 30 nm and 100 nm curved membrane mimicking surface was investigated, it was found there was an observable

difference of 35 %. As such, it can be confidently assumed that 7 % of the reduced SPR signal response for BIN1 binding to the 100 nm tethered vesicles is due to distance, and the rest can be attributed to BIN1's preference for highly curved membranes. While these experiments showed that 35 % of the observed change could be contributed to variation in curvature, more robust statistical analysis strategies, such as principal component analysis, would need to be implemented to validate this claim. Before that could be investigated, the significance of membrane composition needed to be determined.

Membrane Composition: Investigating how changing membrane composition impacts the observed biophysical interactions with BIN1 was pursued with the 30 nm tethered vesicles. The 30 nm curved membrane mimics were selected as BIN1 had a clear preference for them and there would be no concern of a loss in sensitivity due to distance from the SPR active substrate. A comparison between GM_1 and GA_1 was selected, as these two gangliosides would allow for the investigation of how membrane composition changes the observed biophysical interactions and give insights into the importance of charge. This was possible as GA_1 has the same antigenic structure as GM_1, except that it lacks a sialic acid group, thus charge, in its antigenic head group. [57]

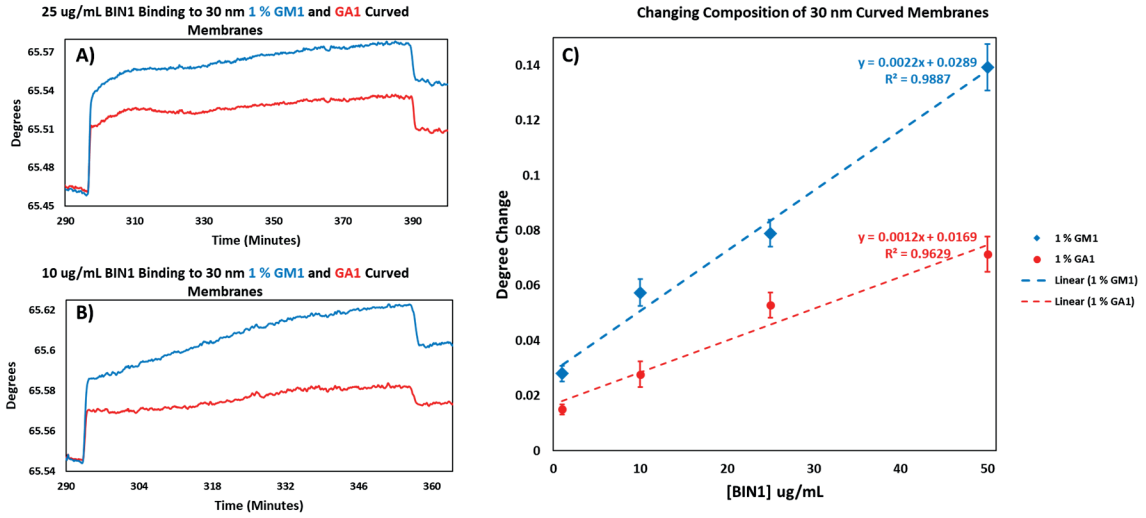


Figure 5.5. A) SPR spectroscopy sensorgram of 25 ug/mL BIN1 binding to 30 nm 1 % GM_1 (blue) and 1 % GA_1 (red) curved membrane mimics, B) 10 ug/mL BIN1 binding to 30nm 1 % GM_1 (blue) and 1 % GA_1 (red) curved membrane mimics, and C) calibration curve of BIN1 binding to GM_1 (blue) and 1 % GA_1 (red) curved membrane mimics at concentrations from 1 to 50 ug/mL.

Figure 5.5 A and B show the interactions of 25 ug/mL and 10 ug/mL BIN1 respectively binding to a 1 % GM_1 (blue) and 1 % GA_1 (red) 30 nm curved membranes. Based upon the two shown sensorgrams there is a clear preference for the vesicles containing GM_1 at both concentrations. This trend was found to be present for all of the investigated concentrations of BIN1, as shown in Figure 5.5 C. As previously stated the LOD for the 30 nm 1 % GM_1 tethered vesicles was 1.33 ug/mL (21 nM) and for the 30 nm 1 % GA_1 tethered vesicles it was found to be 4.82 ug/mL (77.8 nM). Of note was that even with a loss of 7 % detectable signal, due to distance from the SPR substrate, the 100 nm 1 % GM_1 tethered vesicles still had a lower LOD then the 30 nm 1 % GA_1 tethered vesicles. The preference for both the 30 nm and 100 nm 1 % GM_1 containing vesicles over the 30 nm 1 % GA_1 containing vesicles clearly showed BIN1's preference for negatively charged surfaces, even if a neutral membrane has more curvature. This indicates that our platform

agrees well with what has qualitatively been observed in literature,[39, 46] but facilitates quantitative investigations in highly controllable environments, which is something that has not be achieved before.

Statistical Analysis: To date an in-depth statistical analysis to determine the significance of curvature and membrane composition for BAR-domain containing proteins has not been pursued. To help decipher these complex interactions and determine the significance of membrane curvature and composition in regards to BIN1, principal component analysis (PCA) was implemented. Before PCA could be used, analysis of variance (ANOVA) was applied to ensure that all of the collected data was statistically relevant, as shown in Table 5.1.

Table 5.1. ANOVA of BIN1 interactions in PBS with all investigated ganglioside containing curved membrane mimics.

Summary for 50 ug/mL BIN1

Groups	*Count*	*Sum*	*Average*	*Variance*
30 nm 1 % GM1	5	0.6962	0.13924	7.14E-05
30 nm 1 % GA1	5	0.357	0.0714	0.0000413
100 nm 1 % GM1	5	0.4668	0.09336	0.000020048
SLB 1 % GM1	5	0.055	0.011	0.0000025

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.042582886	3	0.014194295	419.837775	2.04E-15	3.238871517
Within Groups	0.000540944	16	0.000033809			
Total	0.04312383	19				

Summary for 25 ug/mL BIN1

Groups	*Count*	*Sum*	*Average*	*Variance*
30 nm 1 % GM1	5	0.395	0.079	0.0000245
30 nm 1 % GA1	5	0.26443	0.052886	2.04E-05
100 nm 1 % GM1	5	0.2375	0.0475	0.0000115
SLB 1 % GM1	5	0.0319	0.00638	0.000006372

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.013538159	3	0.00451272	287.6841176	3.99E-14	3.238871517
Within Groups	0.000250982	16	1.57E-05			
Total	0.01378914	19				

Summary for 10 ug/mL BIN1

Groups	*Count*	*Sum*	*Average*	*Variance*
30 nm 1 % GM1	5	0.287	0.0574	0.0000238
30 nm 1 % GA1	5	0.1389	0.02778	0.000021142
100 nm 1 % GM1	5	0.1673	0.03346	0.000014348
SLB 1 % GM1	5	0.0138	0.00276	0.000000653

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.007545938	3	0.002515313	167.8469657	2.66E-12	3.238871517
Within Groups	0.000239772	16	1.50E-05			
Total	0.00778571	19				

Summary for 1 ug/mL BIN1

Groups	*Count*	*Sum*	*Average*	*Variance*
30 nm 1 % GM1	5	0.14	0.028	0.0000085
30 nm 1 % GA1	5	0.0749	0.01498	0.000003462
100 nm 1 % GM1	5	0.052	0.0104	0.0000043
SLB 1 % GM1	5	0.006	0.0012	0.0000017

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.001866282	3	0.000622094	138.535538	1.17E-11	3.238871517
Within Groups	0.000071848	16	4.49E-06			
Total	0.00193813	19				

Summary for all ug/mL [BIN1]

Groups	*Count*	*Sum*	*Average*	*Variance*
30 nm 1 % GM1	20	1.5182	0.07591	0.001779151
30 nm 1 % GA1	20	0.83523	0.0417615	0.000522092
100 nm 1 % GM1	20	0.9236	0.04618	0.000976279
SLB 1 % GM1	20	0.1067	0.005335	1.73E-05

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.050227753	3	0.016742584	20.3256886	9.04E-10	2.72494392
Within Groups	0.062602377	76	0.000823715			
Total	0.11283013	79				

With the knowledge that each of the investigated datasets were statistically relevant and differentiable, PCA could confidently be implemented. The first investigated relation is shown in Figure 5.6, which examined BIN1 at varying concentrations binding with the different 1 % GM_1 biomimetic lipid membranes. Using the data from Figure 5.3, PCA concluded that the concentrations of BIN1, PC 2, had a weight of 34.2 %, whereas the observed degree change, PC 1, had a weight of 60.8 %. PC1 can be directly correlated to the biophysical interactions, through binding kinetics, between BIN1 and the different curved membrane mimics. The weight observed for the curvature of the lipid membrane agrees well with what has been proposed in literature.[22-24, 37, 39, 46] However, to identify the true significance of membrane curvature, will require substantially more analysis of BIN1 binding to other curved membrane sizes.

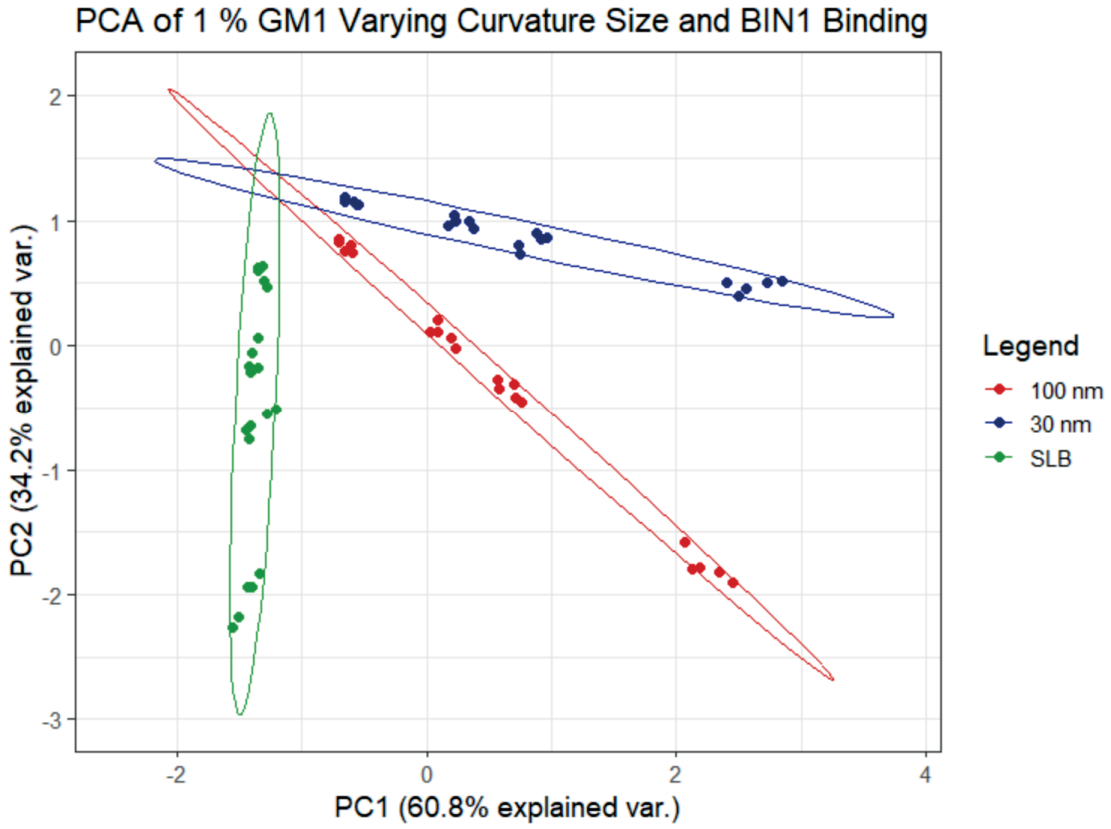


Figure 5.6. PCA of the calibration curves from Figure 5.3 for 100 nm (red) tethered, 30 (blue) nm tethered, and supported lipid bilayer (green) 1 % GM1 99 % POPC biomimetic membrane mimics.

The second investigation using PCA was to determine the significance of membrane composition on the observed biophysical interactions with BIN1, as shown in Figure 5.7. To achieve this, the data shown in Figure 5.5C was used. Notably, the concentration of BIN1, PC 2, had a weight of 28.9 % which is very similar to the weight of PC 2 for the PCA shown in Figure 5.6. This similarity in weight can be attributed to the use of similar datasets and concentrations, but further investigation into this is necessary, as it may prove to be a means to relate curvature and composition analysis. PC 1 in Figure 5.7, represents the weight of membrane composition, through the observed degree changes, and was determined to have a weight of 71.1 %. The large amount of weight associated

with composition can be attributed to BIN1's preference for GM_1 containing vesicles, negatively charged, over GA_1 containing vesicles, neutral. The finding of this comparison agrees well with what has been proposed in literature to date,[33, 39, 46] but needs further investigation with more membrane components including those that are not gangliosides.

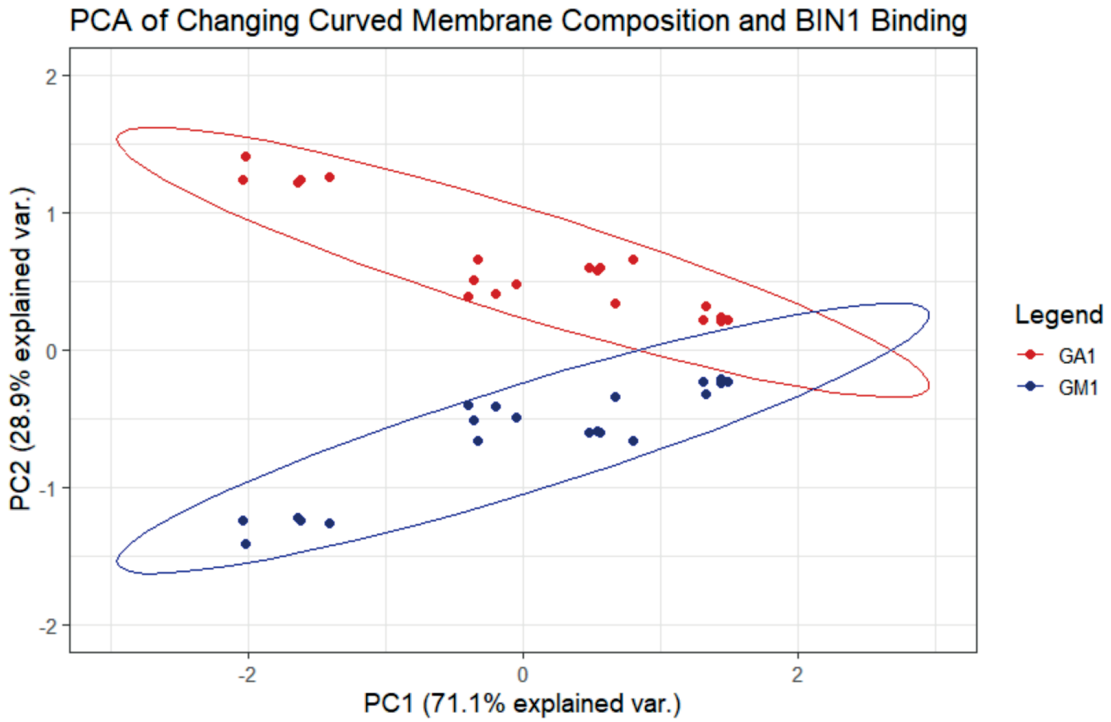


Figure 5.7. PCA of calibration curves from Figure 5.5 to identify the significance of 30 nm containing 1 % GM_1 (blue) and 1 % GA_1 (red) curved membrane mimic compositions binding interactions with BIN1.

While PCA has allowed for the first investigation regarding the significance of membrane curvature and composition in regards to BIN1, more robust statistical analysis methods are needed. In the future, partial least squares discriminant analysis and machine learning algorithms will be implemented to verify and characterize these interactions to untangle the significance of these membrane properties. This is integral as previous work

in our lab has shown that relying solely on endpoint data analysis can lead to a loss of key information to decipher various interactions.[58]

Detection of BIN1 in Urine: Knowing that the 1 % GM_1 30 nm curved membrane platform was ideal for monitoring BIN1, it was selected to demonstrate that detecting BIN1 in urine was feasible. While a specific concentration range for BIN1's presence in urine has not been identified in relation to muscular dystrophy, it has been shown that BIN1's presence in urine is linked to the disease.[48, 59] The utilization of this platform can help determine the disease relevant concentration range for BIN1. Obtaining this information is integral if BIN1 is to be used as a biomarker as a means to detect muscular dystrophy.

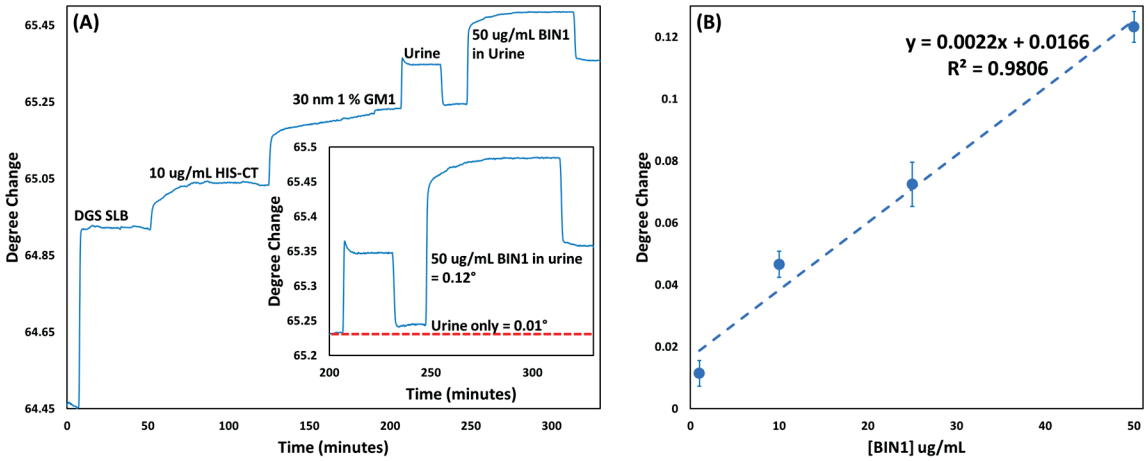


Figure 5.8. (A) overall sensorgram of 50 ug/mL BIN1 spiked into urine samples showcasing the detection capabilities of the 30 nm 1 % GM_1 tethered biomimetic curved membrane mimicking platform in urine. Inserted in (A) is a close up of the urine and spiked urine analysis urine has a 0.01 degree shift but the spiked urine sample has a 0.12 degree shift. (B) Calibration curved of BIN1 at varying concentrations spiked into urine.

Figure 5.8A showcases a representative SPR sensorgram of 50 ug/mL BIN1 spiked into urine interacting with the 30 nm 1 % GM_1 curved membrane mimic. As shown in the insert of Figure 5.8A, the urine saturation step causes a 0.01 degree shift. This shift can largely be attributed to small molecular ions present in urine interacting with the hydration

layer of the curved membrane mimic.[60, 61] When BIN1 spiked in urine is introduced, there is a large shift until the rinse cycle is initiated, causing the majority of bulk change to be removed. The left over signal can be attributed to BIN1's binding to the curved membrane, which agree well with observations in Chapter 3 regarding antibodies spiked in whole serum.[58]

BIN1 was further investigated at a diverse range of concentrations spiked into urine, as shown in Figure 5.8B. A notable loss in signal was present for all investigated concentrations when compared to BIN1 in PBS. This loss in sensitivity is most likely due to the presence of the ions in the urine samples that are likely interfering with BIN1's interaction with the curved membranes. Two hypothesis were proposed to explain the substantial loss in sensitivity. First, the ions present in urine are changing the charge of the tethered vesicle due to their interaction with the hydration layer, thus weaking BIN1's affinity. Second, the disturbance to the hydration layer may lessen the fluidity and malleability of the tethered vesicles,[62] thus making it more difficult for BIN1 to effectively bind to the tethered vesicles. Nevertheless, BIN1 was still able to more effectively bind to the 30 nm 1 % GM_1 platform in urine, LOD of 5.72 ug/mL (92 nM), over the 1 % GM_1 SLB platform in PBS, LOD of 17.23 ug/mL. This observation further showcases the importance of curved membranes in regards to BIN1-membrane interactions.

5.4 Conclusion:

Here we reported the first quantitative characterization of BIN1's biophysical interactions in regards to membrane composition and curvature size, as well as demonstrated that the developed platform can detect BIN1 in urine. Investigating the

biophysical interactions between a single protein and membrane is outside the capabilities of SPR spectroscopy. To address this, we relied on identifying patterns in the average bulk refractive index changes to investigate these highly complex biophysical interactions. While our work demonstrated that BIN1 has a clear preference for highly curved and negatively charged membranes, further validation is needed. This can be achieved by applying mathematical simulations, such as the combination of Monte Carlo simulations and Fresnel Equations. In addition, the inclusion of other membrane components besides gangliosides, such as positively charged lipids, like 1,2-dioleoyl-sn-glycero-3-ethylphosphocholine (EPC+), and negatively charged lipids, like 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (POPG-) should be pursued, as they will allow for more reliable investigations into the role that charge plays in BIN1's interactions. While the developed platform was able to reliably detect BIN1 in urine, the LOD is likely outside the currently unidentified disease relevant concentration range for muscular dystrophy. To address this, we can enhance the sensitivity of our platform through the use of amplification methods with proteins that have high specificity for BIN1. Proteins of interest include: Dynamin 2, caveolin and caveolae proteins, and Phospholipase D2, all of which are known to target specific auxiliary domains of BIN1 that are only available after BIN1 has bound to a curved membrane. These proteins can be labeled with plasmonically active materials, such as gold nanoparticles, through chemical linking methods such as EDC/NHS to drastically amplify the signal of BIN1. In addition to enhancing the sensitivity of the platform, many of these BIN1 targeting proteins themselves are affiliated with other diseases, and investigating them is a goal worth pursuing.

5.5. References:

1. Szlasa, W.; Zendran, I.; Zalesińska, A.; Tarek, M.; Kulbacka, J., Lipid composition of the cancer cell membrane. *Journal of bioenergetics biomembranes* **2020,** *52*, 321-342.

2. Reindl, M.; Waters, P., Myelin oligodendrocyte glycoprotein antibodies in neurological disease. *Nature Reviews Neurology* **2019,** *15* (2), 89-102.

3. Zhong, S.; Li, L.; Shen, X.; Li, Q.; Xu, W.; Wang, X.; Tao, Y.; Yin, H., An update on lipid oxidation and inflammation in cardiovascular diseases. *Free Radical Biology Medicine* **2019,** *144*, 266-278.

4. Feingold, K. R., Thematic review series: skin lipids. The role of epidermal lipids in cutaneous permeability barrier homeostasis. *Journal of lipid research* **2007,** *48* (12), 2531-2546.

5. Murphy, C.; Withrow, J.; Hunter, M.; Liu, Y.; Tang, Y. L.; Fulzele, S.; Hamrick, M. W., Emerging role of extracellular vesicles in musculoskeletal diseases. *Molecular aspects of medicine* **2018,** *60*, 123-128.

6. Ammendolia, D. A.; Bement, W. M.; Brumell, J. H., Plasma membrane integrity: implications for health and disease. *BMC biology* **2021,** *19*, 1-29.

7. Beseničar, M.; Maček, P.; Lakey, J. H.; Anderluh, G., Surface plasmon resonance in protein–membrane interactions. *Chemistry physics of lipids* **2006,** *141* (1-2), 169-178.

8. Mulgrew-Nesbitt, A.; Diraviyam, K.; Wang, J.; Singh, S.; Murray, P.; Li, Z.; Rogers, L.; Mirkovic, N.; Murray, D., The role of electrostatics in protein–membrane interactions. *Biochimica et Biophysica Acta-Molecular Cell Biology of Lipids* **2006,** *1761* (8), 812-826.

9. Sezgin, E.; Schwille, P., Model membrane platforms to study protein-membrane interactions. *Molecular membrane biology* **2012,** *29* (5), 144-154.

10. Auluck, P. K.; Caraveo, G.; Lindquist, S., α-Synuclein: membrane interactions and toxicity in Parkinson's disease. *Annual review of cell developmental biology* **2010,** *26*, 211-233.

11. Safari, F.; Suetsugu, S., The BAR domain superfamily proteins from subcellular structures to human diseases. *Membranes* **2012,** *2* (1), 91-117.

12. Wanleenuwat, P.; Iwanowski, P.; Kozubski, W., Antiganglioside antibodies in neurological diseases. *Journal of the neurological sciences* **2020,** *408*, 116576.

13. Peetla, C.; Stine, A.; Labhasetwar, V., Biophysical interactions with model lipid membranes: applications in drug discovery and drug delivery. *Molecular pharmaceutics* **2009,** *6* (5), 1264-1276.

14. White, S. H., Biophysical dissection of membrane proteins. *Nature* **2009,** *459* (7245), 344-346.

15. Sych, T.; Levental, K. R.; Sezgin, E., Lipid–Protein Interactions in Plasma Membrane Organization and Function. *Annual Review of Biophysics* **2022,** *51*, 135-156.

16. Simunovic, M.; Evergren, E.; Callan-Jones, A.; Bassereau, P., Curving cells inside and out: roles of BAR domain proteins in membrane shaping and its cellular implications. *Annual review of cell developmental biology* **2019,** *35*, 111-129.

17. Bozelli Jr, J. C.; Epand, R. M., Membrane shape and the regulation of biological processes. *Journal of molecular biology* **2020,** *432* (18), 5124-5136.

18. Has, C.; Das, S. L., Recent developments in membrane curvature sensing and induction by proteins. *Biochimica et Biophysica Acta -General Subjects* **2021,** *1865* (10), 129971.

19. Has, C.; Sivadas, P.; Das, S. L., Insights into membrane curvature sensing and membrane remodeling by intrinsically disordered proteins and protein regions. *The Journal of Membrane Biology* **2022,** *255* (2-3), 237-259.

20. Kessels, M. M.; Qualmann, B., Different functional modes of BAR domain proteins in formation and plasticity of mammalian postsynapses. *Journal of Cell Science* **2015,** *128* (17), 3177-3185.

21. Rao, Y.; Haucke, V., Membrane shaping by the Bin/amphiphysin/Rvs (BAR) domain protein superfamily. *Cellular Molecular Life Sciences* **2011,** *68*, 3983-3993.

22. Mim, C.; Unger, V. M., Membrane curvature and its generation by BAR proteins. *Trends in biochemical sciences* **2012,** *37* (12), 526-533.

23. Antonny, B., Mechanisms of membrane curvature sensing. *Annual review of biochemistry* **2011,** *80*, 101-123.

24. McIlhinney, J.; Hooper, N.; Gallop, J. L.; McMahon, H. T. In *BAR domains and membrane curvature: bringing your curves to the BAR*, Biochemical Society Symposia, Portland Press: **2005**; pp 223-231.

25. Carman, P. J.; Dominguez, R., BAR domain proteins—a linkage between cellular membranes, signaling pathways, and the actin cytoskeleton. *Biophysical reviews* **2018,** *10* (6), 1587-1604.

26. Liu, S.; Xiong, X.; Zhao, X.; Yang, X.; Wang, H., F-BAR family proteins, emerging regulators for cell membrane dynamic changes—from structure to human diseases. *Journal of hematology oncology* **2015,** *8* (1), 1-14.

27. Chen, Y.; Aardema, J.; Misra, A.; Corey, S. J., BAR proteins in cancer and blood disorders. *International journal of biochemistry molecular biology* **2012,** *3* (2), 198.

28. Madsen, K. L.; Bhatia, V. K.; Gether, U.; Stamou, D., BAR domains, amphipathic helices and membrane-anchored proteins use the same mechanism to sense membrane curvature. *FEBS letters* **2010,** *584* (9), 1848-1855.

29. Mim, C.; Cui, H.; Gawronski-Salerno, J. A.; Frost, A.; Lyman, E.; Voth, G. A.; Unger, V. M., Structural basis of membrane bending by the N-BAR protein endophilin. *Cell* **2012,** *149* (1), 137-145.

30. Sánchez-Barrena, M. J.; Vallis, Y.; Clatworthy, M. R.; Doherty, G. J.; Veprintsev, D. B.; Evans, P. R.; McMahon, H. T., Bin2 is a membrane sculpting N-BAR protein that influences leucocyte podosomes, motility and phagocytosis. *PloS one* **2012,** *7* (12), e52401.

31. Bassereau, P.; Sorre, B.; Lévy, A., Bending lipid membranes: experiments after W. Helfrich's model. *Advances in colloid interface science* **2014,** *208*, 47-57.

32. Lipowsky, R., Spontaneous tubulation of membranes and vesicles reveals membrane tension generated by spontaneous curvature. *Faraday discussions* **2013,** *161*, 305-331.

33. Prévost, C.; Zhao, H.; Manzi, J.; Lemichez, E.; Lappalainen, P.; Callan-Jones, A.; Bassereau, P., IRSp53 senses negative membrane curvature and phase separates along membrane tubules. *Nature communications* **2015,** *6* (1), 8529.

34. Hsieh, W.-T.; Hsu, C.-J.; Capraro, B. R.; Wu, T.; Chen, C.-M.; Yang, S.; Baumgart, T., Curvature sorting of peripheral proteins on solid-supported wavy membranes. *Langmuir* **2012,** *28* (35), 12838-12843.

35. Chugh, V.; Vijaya Krishna, K.; Pandit, A., Cell membrane-coated mimics: a methodological approach for fabrication, characterization for therapeutic applications, and challenges for clinical translation. *ACS nano* **2021,** *15* (11), 17080-17123.

36. Ray, P.; Chakraborty, R.; Banik, O.; Banoth, E.; Kumar, P., Surface Engineering of a Bioartificial Membrane for Its Application in Bioengineering Devices. *ACS omega* **2023**.

37. Simunovic, M.; Voth, G. A.; Callan-Jones, A.; Bassereau, P., When physics takes over: BAR proteins and membrane curvature. *Trends in cell biology* **2015,** *25* (12), 780-792.

38. Tsai, F.-C.; Simunovic, M.; Sorre, B.; Bertin, A.; Manzi, J.; Callan-Jones, A.; Bassereau, P., Comparing physical mechanisms for membrane curvature-driven sorting of BAR-domain proteins. *Soft Matter* **2021,** *17* (16), 4254-4265.

39. Fu, Y.; Hong, T., BIN1 regulates dynamic t-tubule membrane. *Biochimica et Biophysica Acta -Molecular Cell Research* **2016,** *1863* (7), 1839-1847.

40. Klingseisen, A.; Lyons, D. A., Axonal regulation of central nervous system myelination: structure and function. *The Neuroscientist* **2018,** *24* (1), 7-21.

41. Falcone, S.; Roman, W.; Hnia, K.; Gache, V.; Didier, N.; Lainé, J.; Auradé, F.; Marty, I.; Nishino, I.; Charlet-Berguerand, N., N-WASP is required for Amphiphysin-2/BIN 1-dependent nuclear positioning and triad organization in skeletal muscle and is involved in the pathophysiology of centronuclear myopathy. *EMBO molecular medicine* **2014,** *6* (11), 1455-1475.

42. Zhao, M.; Maani, N.; Dowling, J. J., Dynamin 2 (DNM2) as cause of, and modifier for, human neuromuscular disease. *Neurotherapeutics* **2018,** *15*, 966-975.

43. Lo, H. P.; Lim, Y.-W.; Xiong, Z.; Martel, N.; Ferguson, C.; Ariotti, N.; Giacomotto, J.; Rae, J.; Floetenmeyer, M.; Moradi, S. V., Cavin4 interacts with Bin1 to promote T-tubule formation and stability in developing skeletal muscle. *Journal of Cell Biology* **2021,** *220* (12), e201905065.

44. Cohen, A. W.; Hnasko, R.; Schubert, W.; Lisanti, M. P., Role of caveolae and caveolins in health and disease. *Physiological reviews* **2004,** *84* (4), 1341-1379.

45. Ghim, J.; Chelakkot, C.; Bae, Y.-S.; Suh, P.-G.; Ryu, S. H., Accumulating insights into the role of phospholipase D2 in human diseases. *Advances in biological regulation* **2016,** *61*, 42-46.

46. Prokic, I.; Cowling, B. S.; Laporte, J., Amphiphysin 2 (BIN1) in physiology and diseases. *Journal of molecular medicine* **2014,** *92*, 453-463.

47. Furuhashi, M., Fatty acid-binding protein 4 in cardiovascular and metabolic diseases. *Journal of atherosclerosis thrombosis* **2019,** *26* (3), 216-232.

48. Gargan, S.; Dowling, P.; Zweyer, M.; Swandulla, D.; Ohlendieck, K., Identification of marker proteins of muscular dystrophy in the urine proteome from the mdx-4cv model of dystrophinopathy. *Molecular Omics* **2020,** *16* (3), 268-278.

49. Silva-Rojas, R.; Nattarayan, V.; Jaque-Fernandez, F.; Gomez-Oca, R.; Menuet, A.; Reiss, D.; Goret, M.; Messaddeq, N.; Lionello, V. M.; Kretz, C., Mice with muscle-specific deletion of Bin1 recapitulate centronuclear myopathy and acute downregulation of dynamin 2 improves their phenotypes. *Molecular Therapy* **2022,** *30* (2), 868-880.

50. Fugier, C.; Klein, A. F.; Hammer, C.; Vassilopoulos, S.; Ivarsson, Y.; Toussaint, A.; Tosch, V.; Vignaud, A.; Ferry, A.; Messaddeq, N., Misregulated alternative splicing of BIN1 is associated with T tubule alterations and muscle weakness in myotonic dystrophy. *Nature medicine* **2011,** *17* (6), 720-725.

51. Hansson, O., Biomarkers for neurodegenerative diseases. *Nature medicine* **2021,** *27* (6), 954-963.

52. Sztretye, M.; Szabó, L.; Dobrosi, N.; Fodor, J.; Szentesi, P.; Almássy, J.; Magyar, Z. É.; Dienes, B.; Csernoch, L., From mice to humans: an overview of the potentials and limitations of current transgenic mouse models of major muscular dystrophies and congenital myopathies. *International Journal of Molecular Sciences* **2020,** *21* (23), 8935.

53. Zhou, K.; Hong, T., Cardiac BIN1 (cBIN1) is a regulator of cardiac contractile function and an emerging biomarker of heart muscle health. *Science China Life Sciences* **2017,** *60*, 257-263.

54. Sipione, S.; Monyror, J.; Galleguillos, D.; Steinberg, N.; Kadam, V., Gangliosides in the brain: physiology, pathophysiology and therapeutic applications. *Frontiers in neuroscience* **2020,** *14*, 572965.

55. Malinick, A. S.; Lambert, A. S.; Stuart, D. D.; Li, B.; Puente, E.; Cheng, Q., Detection of multiple sclerosis biomarkers in serum by ganglioside microarrays and surface plasmon resonance imaging. *ACS sensors* **2020,** *5* (11), 3617-3626.

56. Sarigul, N.; Korkmaz, F.; Kurultak, İ., A new artificial urine protocol to better imitate human urine. *Scientific reports* **2019,** *9* (1), 20159.

57. McDaniel, R. V.; McLaughlin, A.; Winiski, A. P.; Eisenberg, M.; McLaughlin, S., Bilayer membranes containing the ganglioside GM1: models for electrostatic potentials adjacent to biological membranes. *Biochemistry* **1984,** *23* (20), 4618-4624.

58. Malinick, A. S.; Stuart, D. D.; Lambert, A. S.; Cheng, Q., Surface plasmon resonance imaging (SPRi) in combination with machine learning for microarray analysis of multiple sclerosis biomarkers in whole serum. *Biosensors Bioelectronics: X* **2022,** *10*, 100127.

59. Petzold, G. C.; Marcucci, M.; Butler, M. H.; Van Landeghem, F. K.; Einhäupl, K. M.; Solimena, M.; Valdueza, J. M.; De Camilli, P., Rhabdomyolysis and paraneoplastic stiff-man syndrome with amphiphysin autoimmunity. *Annals of Neurology: Official Journal of the American Neurological Association the Child Neurology Society* **2004,** *55* (2), 286-290.

60. Navrátilová, I.; Skládal, P., Immunosensor for the measurement of human serum albumin in urine based on the Spreeta surface plasmon resonance sensor. *Supramolecular Chemistry* **2003,** *15* (2), 109-115.

61. Song, J.; Franck, J.; Pincus, P.; Kim, M. W.; Han, S., Specific ions modulate diffusion dynamics of hydration water on lipid membrane surfaces. *Journal of the American Chemical Society* **2014,** *136* (6), 2642-2649.

62. Zimmermann, R.; Küttner, D.; Renner, L.; Kaufmann, M.; Werner, C., Fluidity modulation of phospholipid bilayers by electrolyte ions: Insights from fluorescence microscopy and microslit electrokinetic experiments. *The Journal of Physical Chemistry A* **2012,** *116* (25), 6519-6525.

Chapter 6: Conclusion and Future Perspectives

6.1 Summary of Dissertation Work

The presented Dissertation focused on the development and application of novel biosensing and bioanalysis strategies for surface plasmon resonance (SPR) techniques. The scope of these advancements included: the development, characterization, and application of biomimetic membrane platforms, methods to account for and reduce cross reactive and nonspecific interactions, and the application of robust post data acquisition analysis strategies.

The biomimetic membrane platforms presented here included a myelin sheath mimicking microarray and a highly tunable curved membrane mimic. The myelin sheath mimicking microarray enabled the detection and differentiation of several multiple sclerosis specific biomarkers in serum at disease relevant concentrations. The curved membrane mimicking platform was used to probe the biophysical interactions between ganglioside containing vesicles with both cholera toxin (CT) and Bridging Integrator-1 (BIN1). In addition, the feasibility of detecting muscular dystrophy by monitoring the presence of BIN1 in urine was demonstrated.

The antifouling strategies reported in this Dissertation facilitated a substantial reduction in background signal due to nonspecific binding and/or cross reactivity. Applying these strategies highlighted that the developed platforms can easily be incorporated into currently established disease screening test. The applications of statistical analysis and machine learning demonstrated that SPR can be used to investigate various interactions more effectively than traditional end point data analysis strategies.

The combination of the presented advancements allowed for the detection, differentiation, and interpretation of biological interactions in a manner that was previously considered unattainable for SPR studies. While several major advancements for SPR techniques were presented here, that can also easily be incorporated into other biosensing strategies, further enhancements for SPR and biosensing in general should be pursued.

Surface Plasmon Resonance Imaging for Disease Biomarker Detection

The presented carbohydrate microarray used in Chapters 2 and 3 can be expanded to include various other myelin sheath components. Additional lipids present in the myelin sheath, such as sulfolipids, cerebrosides, other gangliosides, globosides, glycophosphosphingolipids, and saccarolipids,[1-6] can easily be incorporated into the PFDTS SAM via the hydrophobic-hydrophobic interactions. In addition to more lipids, the pseudo myelin sheath microarray can be expanded to include proteins.

By incorporating DGS-Ni-NTA into the PFDTS SAM histidine tagged proteins (his-proteins) can easily be incorporated via the well-established Ni-his interaction.[7] Many proteins can be purchased or expressed with histidine tags, but for the investigation of multiple sclerosis, the focus should be on the incorporation of transmembrane proteins prevalent throughout the myelin sheath.[8] Notable proteins to investigate include myelin associated glycoprotein,[9] myelin basic protein,[10] and proteolipid protein.[11] By incorporating both lipids and proteins in the pseudo myelin sheath microarray, the reliability of the platform will increase dramatically.

Detection of disease specific biomarkers can further be enhanced by replacing the underlying gold substrate with aluminum. Recently, we reported the benefits of combining

aluminum based microarrays (A-chips) with SPRi.[12] The findings of this study showed a notable increase in sensitivity for SPRi, as the A-chips had a 61.6 % higher sensitivity to changes in refractive index compared to the more traditional Au based biochips (K-chips). The A-chips can easily be functionalized with perfluorinated or other alkyl chains with either the naturally occurring thin layer of aluminum oxide,[12, 13] or via depositing a thin layer of SiO_x.[14, 15] The substantial increase in sensitivity the A-chips offer could further aid in the detection and characterization of low abundant biomarkers and various analyte-antigen interactions.

In addition to changing the underlying SPR substrate to increase the sensitivity of SPRi, the use of amplification methods can be pursued.[16] Various plasmonically active nanoparticles,[17, 18] have reliably been used for the amplification of a diverse number of biomarkers via a plethora of functionalization strategies.[19-23] Incorporating amplification methods may be necessary for expanding the developed platform to the detection of biomarkers with small molecular masses, weak binding affinities, and/or those at concentrations outside the limit of detection in serum.[24-27]

Curved Membrane Mimics for the Investigation of Complex Biophysical Interactions

Chapter 4 and 5 presented the development, characterization, and application of a biomimetic curved cellular-membrane mimicking platform. Further investigation of the biophysical interactions between various proteins can benefit from the inclusion of other base lipid component besides the net neutral zwitterionic lipid POPC.[28] Of most interest are the positively charged lipid 1,2-dioleoyl-sn-glycero-3-ethylphosphocholine (EPC^+) and the negatively charged lipid 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol)

(POPG$^-$).[29] These lipids can facilitate the change in overall charge of the vesicles without relying on the antigenic binding sites of the biologically relevant lipids. In addition to changing the overall charge of the vesicles, other membrane components can be included to modify the fluidity of the tethered vesicles, such as by incorporating cholesterol,[30] to make the membranes more rigid.[31, 32]

New platforms can also be explored to generate curved membrane mimics, such as through the use of solid support systems via silica nanoparticles, shown in Figure 6.1.[33] As was demonstrated in this Dissertation and other studies, lipid vesicles will unravel in the presence of SiO_x and naturally form supported lipid bilayers (SLB).[29, 33] Various strategies have been developed to adhere silica nanoparticles to various surfaces, including electrostatic adhesion,[29] carbodiimide crosslinking,[34] alkyne-azide click reactions,[35] and sintering.[36] Silica nanoparticles are also widely available commercially at a diverse range of sizes, specifically from 1 nm to 500 nm. This wide range of nanoparticles sizes could facilitate the investigation of highly controllable membrane curvatures effects on protein-membrane biophysical interactions.

Utilizing solid support platforms for the investigation of curvature sensing proteins, like BIN1 and alpha-syn,[37] may offer unique benefits that tethered membranes lack. Most notably is that the solid support systems would facilitate the implementation of highly defined geometries in simulations. This would allow for far more accurate models compared to those for tethered vesicles, which rely heavily on assumptions. However, solid support systems have the potential to alter the observed protein-membrane biophysical interactions, such as through differences in membrane fluidity and malleability. [33, 38-41]

These differences would likely alter the observed biophysical interactions, which could lead to false conclusions.

Future studies should also incorporate other curvature sensing proteins, like alpha synuclein, and proteins that target the curvature sensing proteins. Many curvature sensing proteins, such as BIN1, have been found to act as recruitment sites for other proteins. Recently it was reported that BIN1 undergoes confirmation changes as it binds to curved membranes,[42] allowing for its auxiliary domains to unfurl and become accessible. Specific auxiliary domains known to act as recruitment sites and are only available after BIN1 has bound to a curved surface include the Src homology 3 (SH3) domain,[43] clathrin and AP2 binding domain (CLAP),[44] and Myc-binding domain (MDP).[45] However, the role and importance of the membranes curvature and charge in the exposure of these domains is not fully understood. Investigating these auxiliary domains orientation and fundamental biophysical interactions may lead to new disease detection and drug development strategies.[39, 45-47]

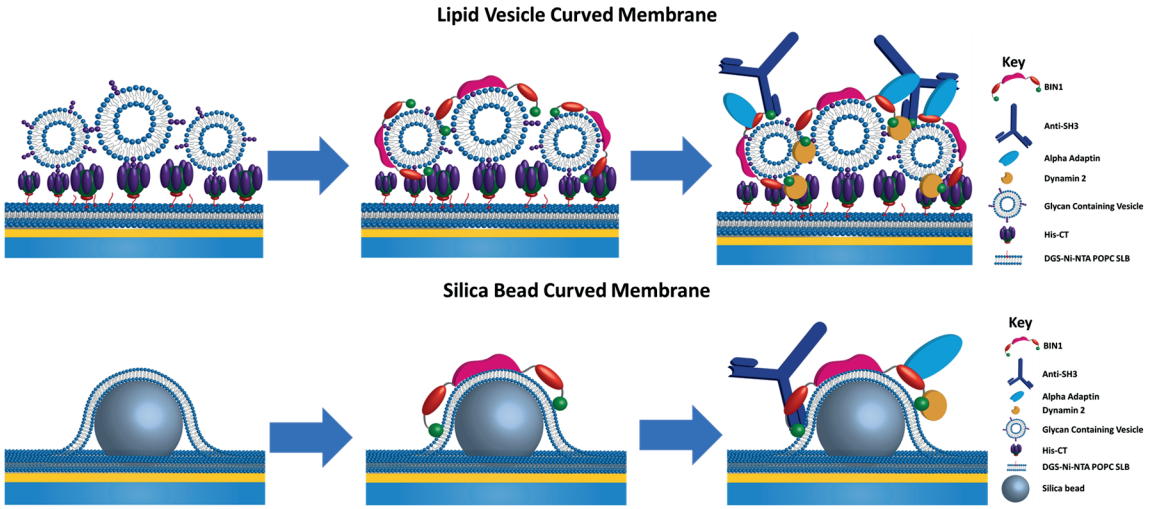


Figure 6.1. Illustration of potential future investigation strategy of using both tether vesicles and silica beads for the investigation of BIN1 and its auxiliary domains with a diverse range of proteins of interest.

Figure 6.1 showcases the proposed strategies for future studies. Many of the proteins that target the SH3, AP2, and MDP binding domains of BIN1 have been associated with various diseases, such as cancer, neurological, skin, cardiovascular, and musculoskeletal diseases.[45, 48-50] Proteins of interest that target the SH3 domain and are notably worth investigating include dynamin 2, caveolin and caveolea proteins, Wiskott–Aldrich syndrome protein, and Phospholipase D2.[48-52] For the CLAP and AP2 binding domains, proteins such as clathrin and fatty acid-binding protein 4, have been found to have high specificity for their respective domains.[45, 53] Each of the listed proteins that target BIN1's auxiliary domains have been reported to be of great interest as drug targets, disease biomarkers, or both due to the diverse roles and importance they play in many biological systems related to diseases.[39, 43-53]

Statistical analysis and Machine Learning

While the ML and statistical analysis tools used in this Dissertation offered immense benefits, these approaches are just a few of the diverse and powerful post data acquisition techniques available to researchers. Incorporating other models into future studies will further validate and prove the reliability of the models used for routine data analysis, especially in the context of disease detection where high variability is always a major concern. Other ML algorithms worth exploring include but are not limited to support vector machines and random forest models.[54] Both of these models may offer unique benefits and pattern recognition capabilities that the presented models were not able to detect. Future studies utilizing the ML and statistical models presented here, regardless of the analytical methods being used, should focus on incorporating larger datasets for further validation.[55]

In addition to including other statistical and ML models, incorporating more variables in the data analysis and training/testing of the models should be explored. Neural networks in particular have the ability to analyze images,[56] thus a program can be written to monitor the observed binding kinetics with the CCD camera used for SPRi investigations in real time. This could lead to new advancements in real time detection as the nnet may be able to predict and provide conclusions throughout the experiment to the user.

6.3 References:

1. Schmitt, S.; Castelvetri, L. C.; Simons, M., Metabolism and functions of lipids in myelin. *Biochimica et Biophysica Acta -Molecular Cell Biology of Lipids* **2015,** *1851* (8), 999-1005.

2. Schnaar, R. L., Brain gangliosides in axon–myelin stability and axon regeneration. *FEBS letters* **2010,** *584* (9), 1741-1747.

3. Tracey, T.; Kirk, S.; Steyn, F.; Ngo, S. In *The role of lipids in the central nervous system and their pathological implications in amyotrophic lateral sclerosis*, Seminars in Cell & Developmental Biology, Elsevier: **2021**; pp 69-81.

4. Dietschy, J. M.; Turley, S. D., Thematic review series: brain Lipids. Cholesterol metabolism in the central nervous system during early development and in the mature animal. *Journal of lipid research* **2004,** *45* (8), 1375-1397.

5. Sarbu, M.; Ica, R.; Zamfir, A. D., Gangliosides as biomarkers of human brain diseases: trends in discovery and characterization by high-performance mass spectrometry. *International Journal of Molecular Sciences* **2022,** *23* (2), 693.

6. Podbielska, M.; Ariga, T.; Pokryszko-Dragan, A., Sphingolipid Players in Multiple Sclerosis: Their Influence on the Initiation and Course of the Disease. *International Journal of Molecular Sciences* **2022,** *23* (10), 5330.

7. Cawley, J. L.; Jordan, L. R.; Wittenberg, N. J., Detection and characterization of vesicular gangliosides binding to myelin-associated glycoprotein on supported lipid bilayers. *Analytical Chemistry* **2020,** *93* (2), 1185-1192.

8. Han, H.; Myllykoski, M.; Ruskamo, S.; Wang, C.; Kursula, P., Myelin-specific proteins: A structurally diverse group of membrane-interacting molecules. *Biofactors* **2013,** *39* (3), 233-241.

9. Quarles, R. H., Myelin-associated glycoprotein (MAG): past, present and beyond. *Journal of neurochemistry* **2007,** *100* (6), 1431-1448.

10. Boggs, J., Myelin basic protein: a multifunctional protein. *Cellular Molecular Life Sciences CMLS* **2006,** *63*, 1945-1961.

11. Greer, J. M.; Lees, M. B., Myelin proteolipid protein—the first 50 years. *The international journal of biochemistry cell biology* **2002,** *34* (3), 211-215.

12. Lambert, A. S.; Valiulis, S. N.; Malinick, A. S.; Tanabe, I.; Cheng, Q., Plasmonic biosensing with aluminum thin films under the Kretschmann configuration. *Analytical chemistry* **2020,** *92* (13), 8654-8659.

13. Luschtinetz, R.; Oliveira, A. F.; Duarte, H. A.; Seifert, G., Self-assembled Monolayers of Alkylphosphonic Acids on Aluminum Oxide Surfaces–A Theoretical Study. Wiley Online Library: **2010**.

14. Phillips, K. S.; Wilkop, T.; Wu, J.-J.; Al-Kaysi, R. O.; Cheng, Q., Surface plasmon resonance imaging analysis of protein-receptor binding in supported membrane arrays on gold substrates with calcinated silicate films. *Journal of the American Chemical Society* **2006,** *128* (30), 9590-9591.

15. Malinick, A. S.; Lambert, A. S.; Stuart, D. D.; Li, B.; Puente, E.; Cheng, Q., Detection of multiple sclerosis biomarkers in serum by ganglioside microarrays and surface plasmon resonance imaging. *ACS sensors* **2020,** *5* (11), 3617-3626.

16. Szunerits, S.; Spadavecchia, J.; Boukherroub, R., Surface plasmon resonance: Signal amplification using colloidal gold nanoparticles for enhanced sensitivity. *Reviews in Analytical Chemistry* **2014,** *33* (3), 153-164.

17. Zhou, C.; Zou, H.; Sun, C.; Ren, D.; Chen, J.; Li, Y., Signal amplification strategies for DNA-based surface plasmon resonance biosensors. *Biosensors Bioelectronics* **2018,** *117*, 678-689.

18. Lambert, A.; Yang, Z.; Cheng, W.; Lu, Z.; Liu, Y.; Cheng, Q., Ultrasensitive detection of bacterial protein toxins on patterned microarray via surface plasmon resonance imaging with signal amplification by conjugate nanoparticle clusters. *ACS sensors* **2018,** *3* (9), 1639-1646.

19. Chang, C.-C., Recent advancements in aptamer-based surface plasmon resonance biosensing strategies. *Biosensors* **2021,** *11* (7), 233.

20. Ding, X.; Yan, Y.; Li, S.; Zhang, Y.; Cheng, W.; Cheng, Q.; Ding, S., Surface plasmon resonance biosensor for highly sensitive detection of microRNA based on DNA super-sandwich assemblies and streptavidin signal amplification. *Analytica chimica acta* **2015,** *874*, 59-65.

21. Lei, P.; Tang, H.; Ding, S.; Ding, X.; Zhu, D.; Shen, B.; Cheng, Q.; Yan, Y., Determination of the invA gene of Salmonella using surface plasmon resonance along with streptavidin aptamer amplification. *Microchimica Acta* **2015,** *182*, 289-296.

22. Liu, Y.; Cheng, Q., Detection of membrane-binding proteins by surface plasmon resonance with an all-aqueous amplification scheme. *Analytical chemistry* **2012,** *84* (7), 3179-3186.

23. Liu, Y.; Dong, Y.; Jauw, J.; Linman, M. J.; Cheng, Q., Highly sensitive detection of protein toxins by surface plasmon resonance with biotinylation-based inline

atom transfer radical polymerization amplification. *Analytical chemistry* **2010,** *82* (9), 3679-3685.

24. Masson, J.-F., Surface plasmon resonance clinical biosensors for medical diagnostics. *ACS sensors* **2017,** *2* (1), 16-30.

25. Linman, M. J.; Abbas, A.; Cheng, Q., Interface design and multiplexed analysis with surface plasmon resonance (SPR) spectroscopy and SPR imaging. *Analyst* **2010,** *135* (11), 2759-2767.

26. Phillips, K. S.; Cheng, Q., Recent advances in surface plasmon resonance based techniques for bioanalysis. *Analytical bioanalytical chemistry* **2007,** *387*, 1831-1840.

27. Malinick, A. S.; Stuart, D. D.; Lambert, A. S.; Cheng, Q., Surface plasmon resonance imaging (SPRi) in combination with machine learning for microarray analysis of multiple sclerosis biomarkers in whole serum. *Biosensors Bioelectronics: X* **2022,** *10*, 100127.

28. Seantier, B.; Breffa, C.; Felix, O.; Decher, G., Dissipation-enhanced quartz crystal microbalance studies on the experimental parameters controlling the formation of supported lipid bilayers. *The Journal of Physical Chemistry B* **2005,** *109* (46), 21755-21765.

29. Sundh, M.; Svedhem, S.; Sutherland, D. S., Formation of supported lipid bilayers at surfaces with controlled curvatures: Influence of lipid charge. *The Journal of Physical Chemistry B* **2011,** *115* (24), 7838-7848.

30. Zhang, Y.; Li, Q.; Dong, M.; Han, X., Effect of cholesterol on the fluidity of supported lipid bilayers. *Colloids Surfaces B: Biointerfaces* **2020,** *196*, 111353.

31. Los, D. A.; Murata, N., Membrane fluidity and its roles in the perception of environmental signals. *Biochimica et Biophysica Acta -Biomembranes* **2004,** *1666* (1-2), 142-157.

32. Glazier, R.; Salaita, K., Supported lipid bilayer platforms to probe cell mechanobiology. *Biochimica et Biophysica Acta -Biomembranes* **2017,** *1859* (9), 1465-1482.

33. Tanaka, M.; Komikawa, T.; Yanai, K.; Okochi, M., Proteomic exploration of membrane curvature sensors using a series of spherical supported lipid bilayers. *Analytical Chemistry* **2020,** *92* (24), 16197-16203.

34. Moore, C.; Monton, H.; O'Kennedy, R.; Williams, D.; Nogues, C.; Crean, C.; Gubala, V., Controlling colloidal stability of silica nanoparticles during bioconjugation

reactions with proteins and improving their longer-term stability, handling and storage. *Journal of Materials Chemistry B* **2015,** *3* (10), 2043-2055.

35. Lutz, J.-F.; Zarafshani, Z., Efficient construction of therapeutics, bioconjugates, biomaterials and bioactive surfaces using azide–alkyne "click" chemistry. *Advanced drug delivery reviews* **2008,** *60* (9), 958-970.

36. Silencieux, F.; Bouchoucha, M.; Mercier, O.; Turgeon, S.; Chevallier, P.; Kleitz, F.; Fortin, M.-A., Mesoporous silica nanoparticles under sintering conditions: a quantitative study. *Langmuir* **2015,** *31* (47), 13011-13021.

37. Antonny, B., Mechanisms of membrane curvature sensing. *Annual review of biochemistry* **2011,** *80*, 101-123.

38. Bozelli Jr, J. C.; Aulakh, S. S.; Epand, R. M., Membrane shape as determinant of protein properties. *Biophysical Chemistry* **2021,** *273*, 106587.

39. Simunovic, M.; Evergren, E.; Callan-Jones, A.; Bassereau, P., Curving cells inside and out: roles of BAR domain proteins in membrane shaping and its cellular implications. *Annual review of cell developmental biology* **2019,** *35*, 111-129.

40. Chan, Y.-H. M.; Boxer, S. G., Model membrane systems and their applications. *Current opinion in chemical biology* **2007,** *11* (6), 581-587.

41. Castellana, E. T.; Cremer, P. S., Solid supported lipid bilayers: From biophysical studies to sensor design. *Surface Science Reports* **2006,** *61* (10), 429-444.

42. Daumke, O.; Roux, A.; Haucke, V., BAR domain scaffolds in dynamin-mediated membrane fission. *Cell* **2014,** *156* (5), 882-892.

43. Dawson, J. C.; Legg, J. A.; Machesky, L. M., Bar domain proteins: a role in tubulation, scission and actin assembly in clathrin-mediated endocytosis. *Trends in cell biology* **2006,** *16* (10), 493-498.

44. Kaksonen, M.; Roux, A., Mechanisms of clathrin-mediated endocytosis. *Nature reviews Molecular cell biology* **2018,** *19* (5), 313-326.

45. Prokic, I.; Cowling, B. S.; Laporte, J., Amphiphysin 2 (BIN1) in physiology and diseases. *Journal of molecular medicine* **2014,** *92*, 453-463.

46. Jiang, X. X.; Zhu, Y. R.; Liu, H. M.; Chen, S. L.; Zhang, D. M., Effect of BIN1 on cardiac dysfunction and malignant arrhythmias. *Acta Physiologica* **2020,** *228* (3), e13429.

47. Safari, F.; Suetsugu, S., The BAR domain superfamily proteins from subcellular structures to human diseases. *Membranes* **2012,** *2* (1), 91-117.

48. Falcone, S.; Roman, W.; Hnia, K.; Gache, V.; Didier, N.; Lainé, J.; Auradé, F.; Marty, I.; Nishino, I.; Charlet-Berguerand, N., N-WASP is required for Amphiphysin-2/BIN 1-dependent nuclear positioning and triad organization in skeletal muscle and is involved in the pathophysiology of centronuclear myopathy. *EMBO molecular medicine* **2014,** *6* (11), 1455-1475.

49. Zhao, M.; Maani, N.; Dowling, J. J., Dynamin 2 (DNM2) as cause of, and modifier for, human neuromuscular disease. *Neurotherapeutics* **2018,** *15*, 966-975.

50. Lo, H. P.; Lim, Y.-W.; Xiong, Z.; Martel, N.; Ferguson, C.; Ariotti, N.; Giacomotto, J.; Rae, J.; Floetenmeyer, M.; Moradi, S. V., Cavin4 interacts with Bin1 to promote T-tubule formation and stability in developing skeletal muscle. *Journal of Cell Biology* **2021,** *220* (12), e201905065.

51. Cohen, A. W.; Hnasko, R.; Schubert, W.; Lisanti, M. P., Role of caveolae and caveolins in health and disease. *Physiological reviews* **2004,** *84* (4), 1341-1379.

52. Ghim, J.; Chelakkot, C.; Bae, Y.-S.; Suh, P.-G.; Ryu, S. H., Accumulating insights into the role of phospholipase D2 in human diseases. *Advances in biological regulation* **2016,** *61*, 42-46.

53. Furuhashi, M., Fatty acid-binding protein 4 in cardiovascular and metabolic diseases. *Journal of atherosclerosis thrombosis* **2019,** *26* (3), 216-232.

54. Sarker, I. H., Machine learning: Algorithms, real-world applications and research directions. *SN computer science* **2021,** *2* (3), 160.

55. Mahesh, B., Machine learning algorithms-a review. *International Journal of Science Research* **2020,** *9*, 381-386.

56. Anwar, S. M.; Majid, M.; Qayyum, A.; Awais, M.; Alnowami, M.; Khan, M. K., Medical image analysis using convolutional neural networks: a review. *Journal of medical systems* **2018,** *42*, 1-13.

Appendix

A.1 Example Code for Principal Component Analysis

```
library(devtools)
library(dplyr)
data
mydata.pca <- prcomp(data, center = T, scale = T)
summary(mydata.pca)
library(ggbiplot)
ggbiplot(mydata.pca)
mydata.antibody <- c(rep("GT1B", “# of data points”), rep("GM1", “# of data points”), rep("GA1", “# of data points”))
ggbiplot(mydata.pca, ellipse=TRUE, ellipse.prob=.95, varname.adjust=.5, obs.scale= 1, var.scale=1, var.axes=FALSE, groups=mydata.antibody)+
  xlim(“best fit for data”, “best fit for data”)+
  ylim(“best fit for data”, “best fit for data”)+
  scale_colour_discrete(name="Legend")+
  geom_point(size=2, aes(colour = factor(mydata.antibody)))+
  theme_bw(base_size = 14, base_family = "")+
  scale_colour_manual(name="Legend", values= c("DARK GREEN", "red", "blue"))+
  ggtitle("PCA of Ganglioside Array")“depend on dataset”
ggbiplot(mydata.pca, ellipse=TRUE, ellipse.prob=.95, varname.adjust=.5, obs.scale= 1, var.scale=1, var.axes=FALSE, groups=mydata.antibody)+
  xlim(“best fit for data”, “best fit for data”)+
  ylim(“best fit for data”, “best fit for data”)+
  scale_colour_discrete(name="Legend")+
  geom_point(size=2, aes(colour = factor(mydata.antibody)))+
  theme_bw(base_size = “depend on dataset”, base_family = "")+
  scale_colour_manual(name="Legend", values= c("forest green", "red3", "dark blue"))+
  ggtitle("PCA of Ganglioside Array")
```

A.2 Example Code for Partial Least Squares Discriminant Analysis

```
library(DiscriMiner)
?plsDA
GA1 <- cbind(GA1, rep('ga1', dim(GA1)[1]))
GT1b <- cbind(GT1b, rep('gt1b', dim(GT1b)[1]))
GM1 <- cbind(GM1, rep('gm1', dim(GM1)[1]))
names(GA1) <- c('Time', rep('Repeated', 1), 'group')
names(GT1b) <- c('Time', rep('Repeated', 1), 'group')
names(GM1) <- c('Time', rep('Repeated', 1), 'group')
new_ga1<- GA1 [, -(“best fit for data”: “best fit for data”)]
dat <- rbind(GM1, GT1b)
names(new_ga1) <- names(dat)
dat<-rbind(dat, GA1)
dat
plsDA(dat[, “best fit for data”: “best fit for data”], dat$group, comps = 3, autosel = F)
fit <- plsDA(dat[, “best fit for data”: “best fit for data”], dat$group, comps = 3, autosel = F)
plot(fit)
plot(fit, ncomp = NULL, nc = 3, show.legend = TRUE, show.labels = T)
install.packages('mixOmics')
if (!requireNamespace("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
BiocManager::install("mixOmics")
library(mixOmics)
fit1 <- splsda(dat[, “best fit for data”: “best fit for data”], dat$group)
plotIndiv(fit1)
plotIndiv(fit1, ind.names = FALSE, legend = TRUE, ellipse = TRUE, star = TRUE)
background <- background.predict(fit1, comp.predicted=2, dist = "max.dist")
plotIndiv(fit1, comp = 1:2, group = dat$group, ind.names = FALSE,
         title = "Maximum distance",background = background)
plotIndiv(fit1, ind.names = FALSE, style = '3d', ncomp = 3) “depend on dataset”
fit1 <- splsda(dat[, 2:5], dat$group, ncomp =3)
plotVar(fit1)
```

A.3 Example Code for K Nearest Neighbor

```
library(DiscriMiner)
library(lattice)
library(ggplot2)
library(naivebayes)
ga1 <- cbind(ga1, rep('ga1', dim(ga1)[1]))
gt1b <- cbind(gt1b, rep('gt1b', dim(gt1b)[1]))
gm1 <- cbind(gm1, rep('gm1', dim(gm1)[1]))
names(ga1) <- c('Time', rep('Repeated', 1), 'group')
names(gt1b) <- c('Time', rep('Repeated', 1), 'group')
names(gm1) <- c('Time', rep('Repeated', 1), 'group')
new_ga1<- ga1
dat <- rbind(gm1, gt1b)
names(new_ga1) <- names(dat)
dat<-rbind(dat, new_ga1)
dat
library(class)
?knn
ntraing <-sample(nrow(dat), “# number of data points”)
training <- dat[ntraing, “depend on dataset”]
testing <- dat[-ntraing, “depend on dataset”]
label <- training[, “depend on dataset”]
test1 <- knn(training[,“depends on dataset”], testing[,“depends on dataset”], label, k=“best fit for dataset”)
table(test1, testing[ , “depends on dataset”])
(“best fit for data”)/nrow(testing)
?knn
library(caret)
x <- testing[ , c(1, 2, 3 “or more depending on dataset”)]
x$group <- as.factor(x$group)
model <- knn3(group ~ ., data= x, k = find best value)
decisionplot(model, x, class = "group", main = "kNN (“depends on dataset”)")
decisionplot <- function(model, data, class = NULL, predict_type = "class",
                         resolution = 100, showgrid = TRUE, ...) {
  if(!is.null(class)) cl <- data[,class] else cl <- 1
 data <- data[,1:2] “depend on dataset”
 k <- length(unique(cl))
  plot(data, col = as.integer(cl)+1L, pch = as.integer(cl)+1L, ...)
 r <- sapply(data, range, na.rm = TRUE)
 xs <- seq(r[1,1], r[2,1], length.out = resolution)
 ys <- seq(r[1,2], r[2,2], length.out = resolution)
 g <- cbind(rep(xs, each=resolution), rep(ys, time = resolution))
 colnames(g) <- colnames(r)
 g <- as.data.frame(g)
  p <- predict(model, g, type = predict_type)
 if(is.list(p)) p <- p$class
 p <- as.factor(p)
  if(showgrid) points(g, col = as.integer(p)+1L, pch = ".")
  z <- matrix(as.integer(p), nrow = resolution, byrow = TRUE)
 contour(xs, ys, z, add = TRUE, drawlabels = t, lwd = 2, levels = (1:(k-1))+.5) invisible(z)}
```

A.4 Example Code for Neural Network

```
data
library(nnet)
?nnet
xm<- data[,1:3] “depend on dataset”
ym<- data[,4] “depend on dataset”
ym_temp <- ym[-1]
ym_2 <-ym[-2]"
# ga1_count <- 25 “depends on dataset”
# gm1_count <- 25 “depends on dataset”
# gt1b_count <- 25  “depends on dataset”
# ga1 <- (pnn_temp == "ga1" && ym[i] !="ga1")
decay_function <- function(d){
result <- rep(0, 75) “depend on dataset”
c1_wrong <- 0
c2_wrong <- 0
c3_wrong <- 0
c1_right <- 0
c2_right <- 0
c3_right <- 0
for(i in 1:75)
  {
  ym_temp <- ym[-i]
xm_temp <- xm[-i,]
model_temp <- nnet(x=xm_temp, y=class.ind(ym_temp), size=2, softmax=T, entropy=T,
maxit=65000,
                decay = d, trace=F) “depend on dataset”
pnn_temp <- predict(model_temp, newdata=xm[i,], type="class")
pnn_temp
confusion_2 <- table(ym[i], pnn_temp)
result[i] <- (pnn_temp==ym[i])
c1_wrong <- c1_wrong + ((pnn_temp=="GT1B") && (ym[i] != "GT1B") )#
c2_wrong <- c2_wrong + ((pnn_temp=="GM1") && (ym[i] != "GM1") )#
c3_wrong <- c3_wrong + ((pnn_temp=="GA1") && (ym[i] != "GA1") )#
c1_right <- c1_right + ((pnn_temp=="GT1B") && (ym[i] == "GT1B") )#
c2_right <- c2_right + ((pnn_temp=="GM1") && (ym[i] == "GM1") )#
c3_right <- c3_right + ((pnn_temp=="GA1") && (ym[i] == "GA1") )#
}
c1_count <- sum(ym == "GT1B") #
c2_count <- sum(ym == "GM1") #
c3_count <- sum(ym == "GA1") #
fpr_c1 <- c1_wrong/(c2_count+c3_count)
fpr_c2 <- c2_wrong/(c1_count+c3_count)
fpr_c3 <- c3_wrong/(c1_count+c2_count)
tpr_c1 <- c1_right/(c1_count)
tpr_c2 <- c2_right/(c2_count)
tpr_c3 <- c3_right/(c3_count)
return(matrix(c(mean(result), fpr_c1, tpr_c1, fpr_c2, tpr_c2, fpr_c3, tpr_c3), nrow=1))
}
decay_function(0.91)
decay_range <- seq(from=0.9, to=1.1, by=0.01) “depend on dataset”
```

```
decay_range
success_rates <- sapply(decay_range, decay_function)
success_rates
plot(x=success_rates[6,], y=success_rates[7,])
model <- nnet(x=xm, y=class.ind(ym), size=2, softmax=T, entropy=T, maxit=65000
            decay =0.91) “depend on dataset”
class.ind(ym)
summary(model)
predict(model)
pnn <- predict(model, type="class")
confusion <- table(ym, pnn)
confusion
data[“start”, “end”] “depend on dataset”
data[“start”, “end”] “depend on dataset”
xmz <- matrix(c(data[, “start”], data[, “end”), ncol = “depends on dataset”)
ymz <- c(rep("GA1", “depends on # of experiments must at least exceed 5”), rep("GM1", “depends
on # of experiments must at least exceed 5”), rep("GT1B", “depends on # of experiments must at
least exceed 5”)
decay_function <- function(d){
  result <- rep(0, “total of data input if at least 5 for each then 15”)
  c1_wrong <- 0
  c2_wrong <- 0
  c3_wrong <- 0
  c1_right <- 0
  c2_right <- 0
  c3_right <- 0
  for(i in 1:18 “depends on dataset”)
  {
    ymz_temp <- ymz[-i]
    xmz_temp <- xmz[-i,]
    model_temp  <-  nnet(x=xmz_temp,  y=class.ind(ymz_temp),  size=2,  softmax=T,  entropy=T,
maxit=65000 “depends on dataset”,
                    decay = d, trace=F) “depend on dataset”
    pnn_temp <- predict(model_temp, newdata=xmz[i,], type="class")
    pnn_temp
    confusion_2 <- table(ymz[i], pnn_temp)
    result[i] <- (pnn_temp==ym[i])
    c1_wrong <- c1_wrong + ((pnn_temp=="GT1B") && (ymz[i] != "GT1B") )#
    c2_wrong <- c2_wrong + ((pnn_temp=="GM1") && (ymz[i] != "GM1") )#
    c3_wrong <- c3_wrong + ((pnn_temp=="GA1") && (ymz[i] != "GA1") )#
    c1_right <- c1_right + ((pnn_temp=="GT1B") && (ymz[i] == "GT1B") )#
    c2_right <- c2_right + ((pnn_temp=="GM1") && (ymz[i] == "GM1") )#
    c3_right <- c3_right + ((pnn_temp=="GA1") && (ymz[i] == "GA1") )#
  }
  c1_count <- sum(ymz == "GT1B") #
  c2_count <- sum(ymz == "GM1") #
  c3_count <- sum(ymz == "GA1") #
  fpr_c1 <- c1_wrong/(c2_count+c3_count)
  fpr_c2 <- c2_wrong/(c1_count+c3_count)
  fpr_c3 <- c3_wrong/(c1_count+c2_count)
  tpr_c1 <- c1_right/(c1_count)
  tpr_c2 <- c2_right/(c2_count)
  tpr_c3 <- c3_right/(c3_count)
```

```
  return(matrix(c(mean(result), fpr_c1, tpr_c1, fpr_c2, tpr_c2, fpr_c3, tpr_c3), nrow=1))
}
decay_function(0.91)

decay_range <- seq(from=.2, to= 0.25, by=0.0001) “depend on dataset”
decay_range
success_rates <- sapply(decay_range, decay_function)
success_rates
plot(x=success_rates[6,], y=success_rates[7,])
success_rates[,success_rates[1,]==max(success_rates[1,])]
decay_range[success_rates[1,]==max(success_rates[1,])]
modelz <- nnet(x=xmz, y=class.ind(ymz), size=2, softmax=T, entropy=T, maxit=65000,
            decay =0.91) “depends on dataset”
class.ind(ymz)
summary(modelz)
predict(modelz)
pnn <- predict(modelz, type="class")
confusion <- table(ymz, pnn)
confusion
xmz
q <- seq(“start point”, “end point”, by = “changes based upon data 5 was used for model”)
d <- seq(“start point”, “end point”, by = “changes based upon data 5 was used for model”)
xmz_test <- expand.grid(x=q, y=d)
pnn_test <- predict(modelz, newdata = xmz_test, type = "class")
pnn_test[pnn_test == "GA1"] <- "red"
pnn_test[pnn_test == "GM1"] <- "green" “red and green were switch by accident in publication :/”
pnn_test[pnn_test == "GT1B"] <- "blue"
plot(xmz_test, col = pnn_test)
modelz
xmz[,1] <- (xmz[,1] - mean(xmz[,1]))/sd(xmz[,1])
xmz[,2] <- (xmz[,2] - mean(xmz[,2]))/sd(xmz[,2])
colMeans(xmz)
apply(xmz, 2, sd)
current_value <- Inf
for (i in 1:500) {
  w <- runif(15, -7, 7)
  model_current <- nnet(x = xmz, y = class.ind(ymz), Wts= w, size=2, softmax=T, entropy=T,
maxit=65000,
                                    decay =0.23) “depend on dataset”
  if (model_current$value < current_value) {
    nnet.save <- model_current
    current_value <- model_current$value
  }
}
nnet.save$value
pnn_opt <- predict(nnet.save)
```

A.5 Example Code for Monte Carlo Simulations

The developed code was adapted from Daniel Stuart and can be found at https://danieldstuart.github.io/LipidVesicleModeling.html

For those of you who have read my Dissertation and made it to this final page, I want to say thank you. This document is an accumulation of my scientific accomplishments to date, and sharing it with you has been a great honor. I want to emphasize "It's a weird feeling, scientific breakthroughs. There's no Eureka moment. Just a slow, steady progression toward a goal. But man, when you get to that goal it feels good."[1] – Ryland Grace, and now that this chapter of my life is complete "I think I'm quite ready for another adventure"[2] – Bilbo Baggins.

References:

1. Weir, A., *Project Hail Mary: A Novel*. Ballantine Books: **2021**.

2. Jackson, P., The Lord of the Rings: The Return of the King. *New Line Cinema* **2003**.