Characterizing and Predicting the Interaction of Proteins with Nanoparticles

by

Karsten M. Poulsen

Department of Mechanical Engineering and Materials Science
Duke University

Date:_____

Approved:

_____
Christine K. Payne, Supervisor

_____
Gaurav Arya

_____
Samira Musah

_____
Mary Clair Thompson

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Mechanical Engineering and Materials Science in the Graduate School
of Duke University
2023

ABSTRACT

Characterizing and Predicting the Interaction of Proteins with Nanoparticles

by

Karsten M. Poulsen

Department of Mechanical Engineering and Materials Science
Duke University

Date: _____

Approved:

_____
Christine K. Payne, Supervisor

_____
Gaurav Arya

_____
Samira Musah

_____
Mary Clair Thompson

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the Department of
Mechanical Engineering and Materials Science in the Graduate School of
Duke University
2023

Abstract

Nanoparticles are being used or implemented in a wide array of applications including health care, cosmetics, automotive, food, beverage, coatings, consumer electronics, and coatings. Despite this widespread use, we are unable to predict how nanoparticles will interact with biological systems. This is important for both nanotoxicity resulting from human exposure to nanomaterials and the development of new nano-based biotechnologies. The first step in the interaction of nanoparticles with biological systems is often with blood, for biomedical applications, or lung fluid, when nanoparticles are inhaled. In both cases, the nanoparticles are exposed to proteins that form a "corona" by adsorbing on the nanoparticle surface. The subsequent cellular response is determined by this protein corona rather than the bare nanoparticle.

Our goal is to develop a predictive capability for protein-nanoparticle interactions. This requires lab automation, large datasets, machine learning, and mechanistic studies. We first developed and validated a semi-automated approach to generate, purify, and characterize protein coronas using a liquid handling robot and low-cost proteomics. Using this semi-automated approach, we characterized the formation of protein coronas with increasing incubation time and serum concentration. Incubation time was found to be an important parameter for corona composition and

concentration at high incubation concentrations but yielded only a small effect at low serum incubation concentrations.

To better understand how the protein corona affects biological responses, we investigated the response of macrophage cells to titanium dioxide nanoparticles as a function of the protein corona. As in our previous work with serum proteins, we measured the concentration and composition of murine lung fluid proteins adsorbed on the surface of titanium dioxide nanoparticles. We found that a low concentration of lung fluid corona, relative to fetal bovine serum and bovine serum albumin coronas, led to an increased expression of cytokines (interleukin 6 (IL-6), tumor necrosis factor-alpha (TNF-α), and macrophage inflammatory protein 2 (MIP-2)), indicating an inflammation response. This underscores the importance of understanding how the composition and concentration of the protein corona governs organism responses to nanoparticle exposures.

Our validated high-throughput lab automation work allowed us to synthesize a library of magnetic nanoparticles and characterize their resulting protein coronas. The resulting nanoparticle dataset has 12 unique NP surfaces, seven surface charges, two core sizes, and two core materials. We used this dataset to generate and characterize, via proteomics, a variety of protein coronas varying incubation concentration and purification methods. We used the resulting proteomic dataset in conjunction with a database of protein physicochemical properties to build a machine learning model that

identifies the most important nanoparticle and protein properties for protein corona formation. The model was tested using external datasets and found that it can predict human serum and lung fluid coronas on varying nanoparticle surfaces. Overall, this combination of lab automation, machine learning, and mechanistic analysis suggests that a generalizable understanding of the protein corona formation and its effects is forthcoming.

Statement of Co-Authorship

Chapters throughout this dissertation have been drawn from collaborative published studies. The co-authors contributions have been essential in producing these manuscripts. I note the co-authors at the begging of each chapter but in this section I will briefly detail author involvement.

In Chapter 2, my co-authors were Thomas Pho and Professors Julie Champion and Christine Payne. I was responsible for developing automated protocols for characterizing protein coronas, while Thomas Pho was responsible for synthesizing and characterizing the OVA-nanoparticles. The data visualizations and analysis related to the protein coronas are my own. Christine and Julie provided essential context for the impact of the research that was being performed.

In Chapter 3, my co-author was Professor Christine Payne. I was responsible for nanoparticle and corona characterization, experimental design, analysis, and data visualization. Christine was invaluable in advice on experimental directions and helping contextualize our results.

In Chapter 4, my co-authors were Michaela Albright, Nicholas Niemuth, and Professors Rob Tighe and Christine Payne. I was responsible for analyzing and characterizing the formation of the protein coronas. All data visualizations and statistical testing is my own. Nicholas treated the cells and Michaela performed the PCR

experiments. Rob provided essential context and understanding for lung inflammation response.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I am so fortunate to have such great family members, friends, advisors, and mentors.

I am especially grateful to my advisor, Christine Payne, for choosing me to be a part of their lab and for being an amazing mentor. Their guidance and support have been invaluable in my research journey. They have always pushed me to tackle interesting problems and showed me how to craft a story out of the answers. I am truly grateful for their belief in my abilities and their dedication to my growth as a researcher. Working with them has been a privilege, and I am thankful for the lasting impact they have had on my academic and personal development.

Along with all the advising Christine did she also introduced me to Mary Clair Thompson who has been invaluable in helping me figure out how to start working on the machine learning side of things. From helping break down the problems to clarifying my understanding of the systems, she has been both a helpful and insightful advisor. I cannot imagine where I would be today without her help. I would also like to thank my other committee members, Samira Musah and Gaurav Arya for their support, feedback and advice which helped shape my project into what it is today.

While I came in with some wet lab experience, I had limited to no experience with biological systems. Thanks to Dhanya Jayaram and Darbi Griffith for showing me the ropes when I first joined the Payne Lab. It would have been a much harder transition

without their support and training. I am especially grateful for Nathan Rayens and Xu Han who have been here with me undergoing similar stresses, keeping me sane when in the lab late at night or giving ideas and support when I felt stuck in a rut. I am so fortunate to have had them as lab mates since the beginning of this adventure. Thank you to all of the Payne Lab members I have worked with (Dhanya Jayaram, Darbi Griffith, Nathan Rayens, Xu Han, Gustavo Sosa Macias, Judith Dominguez, Faisal Anees, Samantha Holmes, Morgan Heckman, Mia Malden, David Scarlett, Emi Yuan, and Diego Montoya) for the years of conversation, support, community, coffee chats, and friendship.

None of this would have been possible without the support of my entire family. Starting with my parents way back when, guiding me towards science and engineering and fostering my love for understanding the world around me. Their continued support and love have helped keep me going even when things felt hard. Another crucial role was convincing my partner in life, Maren, to move across the country away from our families and support. I couldn't have done this without you, and I am so grateful that you are in my life. Not only has your support been invaluable, but you brought Nora and Nalu into my life who have put everything I do into perspective and given me all the motivation (with only a little bit of distraction) in the world. I am all too fortunate to have such a wonderful group of friends and family who have supported me throughout

this journey by helping with house projects, brainstorming potential problems with my research, going hiking, puppy play dates, proofreading documents, and so much more.

Finally, I want to thank the Mechanical Engineering and Materials Science (MEMS) department and Pratt School of Engineering for making an inclusive open community that helped me find helpful peers and friends whenever I needed them. I have been very fortunate to find great friends and communities through the student groups I was involved with including MEMS graduate student committee and Engineering Graduate Student Council.

1. Introduction

According to ISO nanomaterials are typically, but not exclusively, a material with any external dimension that is between 1-100 nm [1-4]. There is a marked interest with materials on this scale due to their high surface area to volume ratio and their capability to exhibit unique optical, magnetic, conductive, and mechanical properties [2]. People have been interacting with nanomaterials since antiquity. These nanomaterials were commonly used as additives for artwork to make sparkly glazes and eventually led to the creation of stained glass [2-4]. The modern understanding of nanoparticles (NPs), including their size, has been around since the 1950s with the synthesis and characterization of colloidal gold NPs [5, 6]. In parallel, the development of characterization techniques like transmission electron microscopy, scanning electron microscopy, atomic force microscopy, and dynamic light scattering have spurred dramatic advances in our understanding of NP structure and functionalization [5]. With the increasing popularity and usage of nanomaterials, there is clear motivation for understanding how they interact with organisms, both for developing new applications in nanomedicine and preventing toxic exposures to workers and consumers.

In this dissertation, I provide a semi-automated set of methods for analyzing NP protein interactions. To go along with these methods, I have analyzed NP-protein interactions from multiple perspectives to get an understanding of what biophysical properties mediate these interactions. I worked with collaborators to understand how NP protein complexes can modulate the inflammatory response in cells in an inhalation

model study. All of these motivate my final work of working towards predicting NP protein interactions with machine learning. The introduction will provide a background of NP applications, the current field of NP protein interactions, how NPs interact with cells, and the machine learning methods used to predict NP protein interactions.

1.1 Nanoparticle applications and exposures

NPs are used in a wide range of applications spanning large industrial applications like paint and pesticide formulations to drug delivery [7-10]. NPs are of interest predominantly because they have extremely high surface area, are slow to sediment relative to their larger counterparts. In addition, they are essentially limitless in their physical composition and can be synthesized for case-specific properties. At present, researchers and manufacturers have established commercial availability and synthesis methods for NPs that are soft (polymeric, lipid, protein, exosomes) and hard (gold, silver, copper, iron oxide, silica, titanium dioxide) [9, 11-17]. Along with structural properties, there is also significant literature on how to functionalize NP surfaces to produce tailored performance in any given application. These functionalization techniques include changing the surface charge, modifying hydrophobicity, and conjugation of structures for effective shielding from molecules through steric hindrance or controlled self-assembly [9, 18-26]. Taken together, these synthesis and functionalization techniques point to limitless problems that customized NPs can be

used to solve. Within this dissertation, I will focus primarily on how NPs interact with biofluids in the context of downstream effects on organisms.

1.1.1 Drug Delivery

NPs are an ideal candidate as a drug delivery platform for a plethora of reasons, particularly because of increased drug efficacy and decreased side effects through targeted drug delivery [9, 27-30]. For many medications, the drug is administered in a way where it is systemically delivered to the entire body. While this is not always a problem, when drugs are being administered systemically that cause cell death, such as chemotherapies, they have serious side effects when they hit any off-target cells. In these scenarios the ideal solution is to design a drug delivery system that will only release the drug locally at a desired target site, such as only around cancer cells. NPs are an appealing choice for this as they are small enough to travel throughout the bloodstream, can be functionalized to target specific cells, and can be designed to deliver payloads with specific stimuli, such as temperature, magnetic, ultrasound, light, and pH [31-35]. The stimulation can either be external, as in applied locally by an instrument or endogenous, as in when a NP enters a specific cell structure.

There are of course some issues with this idealized picture. The moment NPs come into contact with any sort of biofluid (blood, lung lining, saliva, etc.) proteins adsorb onto the surface forming a protein "corona" [18, 36-40]. The protein corona can counteract designed functionalization like targeted delivery and can potentially lead to

reduced lifetime within the bloodstream, further decreasing the NPs likelihood to reach its desired target [27, 30, 40, 41]. Although the protein corona represents a challenge for targeted drug delivery, it has the potential to be mediate or prevent immune response by cloaking the foreign NP in recognizable, native proteins [42, 43]. With every NP drug application, it is necessary to understand the protein corona that will form. Consequently, predicting the formation of the corona prior to designing a novel drug delivery system has the potential to improve development capabilities, reduce costs and lead to more effective delivery. Understanding what factors dictate the formation of specific coronas is the primary motivation for our experiments in Chapter 3.

1.1.2 Inhalation exposures

NPs are increasingly being used on large scale in the production of industrial materials as well as for large commercial applications like agriculture [44-50]. With this increased use there is a growing likelihood of uncontrolled NP exposures for humans. The predominant exposure risk is inhalation, either from manufacturing and agricultural settings, where fine NP dusts may be aerosolized, or when NP-containing materials degrade [51-54]. While NP inhalation is not well understood, it is expected to have similar effects, if not more severe, to inhaled dusts, which have been shown to cause inflammation and debilitating lung diseases [55, 56]. In particular, titanium dioxide inhalation is a major concern due to its ubiquity in paints, plastics, coatings, paper, and food additives (for whitening/brightening) [57-59].

We understand that the protein corona is deterministic in the cellular response and biological fate of NPs. Most of the research on the protein corona has focused on blood serum proteins more relevant to nanomedicines and sensors [60-62]. This leaves significant space in the literature for understanding how NPs will interact with the complex environment of the lung lining, which includes a mixture of lipids, proteins, and mucins. Lung fluid plays a role in the bodies innate defense system with antimicrobial and oxidant resistant proteins [63-65] and it is not well understood if the NPs that interact with these immune response proteins will lead to an increased or decreased toxicity response. Previous work in this area has shown that, when introduced to human respiratory tract lining fluid, NPs have formed a corona with highly enriched immunoregulatory proteins, suggesting an increased inflammation response [66]. Specific cytotoxicity markers have been analyzed both in vivo and in vitro [67, 68]. NP cytotoxicity assays in vitro showed that zinc oxide and certain titanium dioxide NPs were cytotoxic to cells, while multiwalled carbon nanotubes were not cytotoxic [68]. This is contrasted with results from in vivo studies in mice and rats where all of the nanomaterials caused similar increases in neutrophil cell counts and inflammatory lesions [67]. These results motivate understanding how the protein corona modulates the inflammatory response of inhaled NPs (Chapter 4).

1.2 Characterizing nanoparticle protein interactions

The term protein corona was coined in 2007, but previously it was referred to as opsonization, which referred to the immune response that targets NPs for removal [40]. The current study of the protein corona relies on the field of proteomics, which is the study of proteins. The gold standard in proteomics today is a combination of liquid chromatography and mass spectrometry [69-73]. This can be done in a couple of ways, but the most common way is called shotgun proteomics, wherein the protein is digested by a specific enzyme, usually trypsin because its cleavage rules are well understood. The peptides are separated with a liquid chromatography column and analyzed by mass spectrometry twice (LC-MS/MS) to determine the peptide weight and sequence. The observed peptides can then be associated with specific proteins based on an understanding of what peptides should exist from an original library of protein sequences and the enzyme cleaving rules. Although relatively slow and convoluted, this method provides the most reliable information about specific proteins and their relative quantities, when looking at complex protein mixtures. Along with LC-MS/MS, gel electrophoresis can be used to separate proteins by their molecular weight, giving a qualitative composition of a protein sample. Finally, protein concentration can be measured broadly, without specificity to protein identities, with a variety of colorimetric assays. Using these methods, we can characterize the composition and concentration of the protein corona.

A problem that has loomed in the protein corona field is the dynamics of the corona formation. The general idea around the corona is that its composition is dynamic and that higher affinity proteins will eventually displace the lower affinity more abundant proteins known as the Vroman effect [74]. This effect is not always observed and seems likely to depend on the specific surface of the NP creating a hard to generalize system. Pisani et al, observed that the protein binding could be broken into three separate groups including adsorption and desorption, early enrichment, and late enrichment, showing that, at least for their specific silica NPs, the corona is dynamic over time [75]. This question of corona dynamics was one motivation for our work in Chapter 3.

Another specific interest in the field is determining what surface interactions drive the adsorption of proteins. From a physicochemical perspective it would seem likely that electrostatics, where a charged surface interacts with a partially charged protein, would be the dominant interacting force, especially since it is well understood that many proteins carry a charge at physiological pH. Even on a smaller scale, we know which specific amino acids are charged, or *polar*, and we can predict which amino acids are exposed. To this end, NP surface charge can be modified and has been shown to significantly affect the composition of the protein corona [23, 76, 77]. In general, charged surfaces tend to adsorb more protein while neutral surfaces tend towards decreased

protein adsorption, but how the protein corona composition changes with surface charge is still unclear [18].

Along with electrostatics, hydrophobicity is expected to play a role in protein adsorption. Generally, proteins are folded so that their hydrophobic parts are not accessible to solution, increasing solubility. There are still a few hydrophobic amino acids on the surface of the protein that could initiate an interaction and potentially denature the protein to create a large hydrophobic interaction at the NP-protein interface. To this end, preventing protein adsorption and corona formation through the use of hydrophilic linkers is one attractive and extremely well-documented methodology. The most common way to do this is to functionalize a NP surface with a long hydrophilic linker, predominantly polyethylene glycol (PEG), that essentially acts as a shield to prevent interactions between the NP and surrounding biomolecules [20, 78-88]. This preventive shield is often referred to as a stealth coating. Interestingly, hydrophilic linkers do not completely prevent protein adsorption and this could be an effect of incomplete functionalization or protein adsorption directly to PEG itself. Along with this, the relatively recent finding that anti-PEG antibodies exist [89], further challenges how PEG prevents protein adsorption and its mechanistic function in NP stealth coatings.

While electrostatics and hydrophobics are likely the strongest NP-protein interactions expected to play a role in corona formation, many research groups have also

studied other NP properties, such as size, shape, functional groups, and core material [23, 83, 90-92]. Outside of NP-specific properties, even experimental methods have been shown to have an effect. For example, something as simple as purifying the NP-protein complexes using magnets as opposed to centrifugation leads to varied corona compositions [12]. Along these lines, temperature, shaking/flow, time, protein source, and incubation concentration all seem to play a role in the formation of the protein corona [36, 39, 60, 75, 93, 94]. It is expected that how the protein corona is isolated from the NP along with *specifically* how the protein corona composition is evaluated will also play a role in determined corona composition. All of these factors motivate standard, reproducible methods for analyzing the protein corona across experiments and researchers. This, in part, drives our use of an automated approach to generating and analyzing the protein corona.

In summary, the protein corona is complex and the interactions that facilitate adsorption are not fully understood. We expect there is a complex relationship between protein/NP properties and the observed protein corona. Given the dramatic number of contributing variables in this problem, we hypothesize that clarifying this complex relationship will require computational techniques, such as machine learning, over a large dataset of well documented protein coronas and NP surfaces. The goal with building a machine learning model is to further understand what mechanisms drive

protein adsorption and to be able to predict protein corona formation for development of new novel nanomaterials and drug delivery systems.

1.3 Predicting the protein corona with machine learning

Machine learning has rapidly gained popularity in a variety of applications for its ability to elucidate nuanced and complicated relationships from large-scale data that is too complicated for direct human analysis. This is especially useful for systems with many properties that are expected to have a role and one specific outcome or dependent variable. Since libraries like Scikit-learn, Caret, PyTorch, and TensorFlow were built, machine learning has become much more accessible and understandable even to non-computational researchers [95-98]. It is important to note that Caret and Scikit-learn are focused on what is commonly referred to as machine learning, while PyTorch and TensorFlow focus on deep learning, a subset of machine learning.

Machine learning is often focused on solving two main problems: regression and classification. Regression involves the use of a continuous label, while classification uses a binary label. Both methods can take in a variety of independent variables, often referred to as features. In the case of predicting protein coronas the features can be related to the NP, protein, biofluid, or experimental characteristics. With any of these features it is important to try and capture all the important physical characteristics of the system. For instance, with NPs the features should encapsulate properties like hydrophobicity, charge, size, ligand functionalization, and core material. The machine

learning methods can take these features and build a model in the process, "learning" how the features are related to the dependent variable.

For the problem of predicting the protein corona, the most commonly used method in the literature is random forest [99-101]. Random forest is popular because it is easy to interpret and not as prone to overfitting as some comparable methods. In brief, random forest uses a "forest" of decision trees to determine the most likely outcome. Each decision in the decision tree is based off of splitting the data along a feature (Fig 1-1) so that the model can most closely match the true data labels. The decisions made at each split in the decision trees are average across the "forest" and an intuitive, macroscopic decision is made based on the average of output tree decisions. Although these trees are constructed using observed, "training" data, they can then be used to predict labels on unknown, new data.

There are two common ways of looking at the protein corona commonly referred to as protein abundance and enrichment. Enrichment is the relative abundance of the protein on the corona relative to the abundance in the biofluid. Generally, classification can provide answers to important mechanistic questions when thresholding for proteins that have high enrichment as that suggest high affinity [100, 101]. In the case of the protein corona, the classifier label is set at a certain enrichment threshold. Using classification, it has been seen that using just protein properties, enrichment in the corona can be predicted for a different proteome than originally trained on [101]. Using

regression, where the label is continuous like actual protein abundance in the corona (rather than a threshold), relatively high accuracy has been observed, with the most

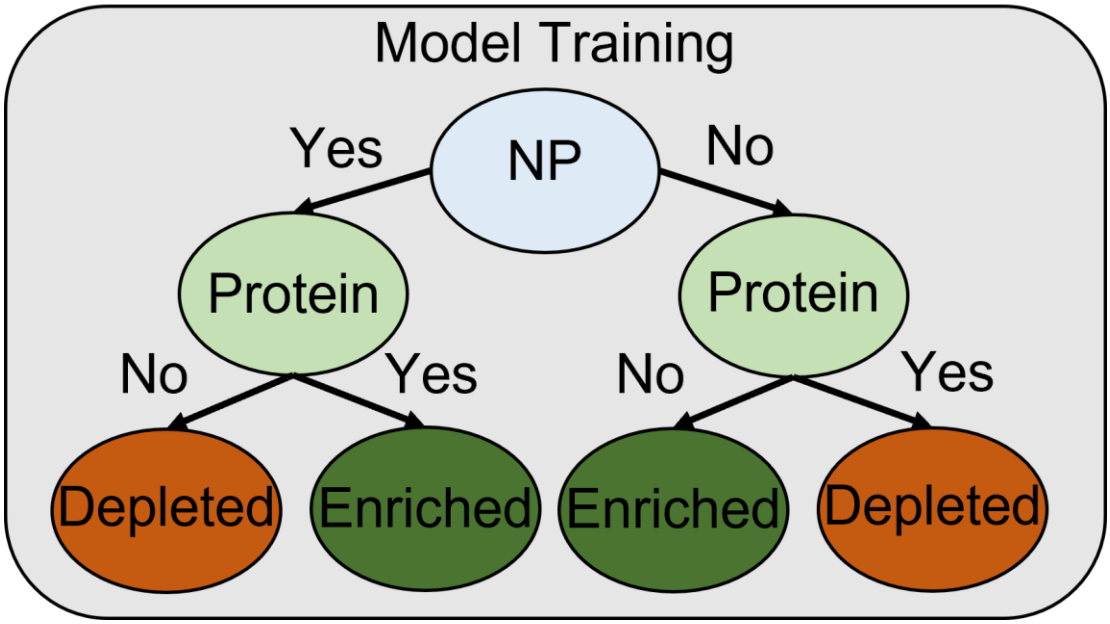

Figure 1-1 Example decision tree for training a model on protein corona data.

important features found to be NP size and core material [102]. Both of these results have not been expanded to larger datasets or higher throughput models that can generalize and predict protein-NP interaction. This lays out the motivation for the work done in appendix A.

1.4 Objectives and outline

This introduction has provided an overview of the key methods used in this dissertation and major historical results in the protein corona field. Significant work is still ahead in further understanding protein corona formation. Specifically, predictive capabilities and generalizable trends remain elusive. The main objective of this work is

to build methods for learning these generalizable trends, including predicting the formation of the protein corona based on protein and NP characteristics, and understanding how the protein corona affects cellular responses. These methods were developed using automated methodologies with magnetic NPs, low-cost proteomics, and mechanistic studies and can be easily adopted by researchers and manufacturers in the field.

Chapter 2 showcases the use of automated liquid handling and low-cost proteomics to highlight the potential of automated methodologies. To validate our low-cost proteomics, we ran two protein corona samples in technical triplicate and confirmed that the low-cost methods produced the same results for the top 25 most abundant proteins. We also validated the automated capabilities of making and purifying protein coronas with multiple replicates in a 96 well plate format using a liquid handling robot. This chapter is drawn from my first author publication in Analytical and Bioanalytical Chemistry [103] with coauthors Thomas Pho and professors Julie Champion and Christine Payne.

Chapter 3 further builds on the liquid handling automation and low-cost proteomics to characterize how serum incubation concentration and incubation time affect the formation of the protein corona. In this chapter we learned that protein corona concentration increases with increasing serum concentration. We also found that incubation time does not play a large role in the most abundant proteins of the protein

corona within our magnetic NP system. These results further highlighted the capabilities of the automated methods creating a framework for high throughput data. This chapter is adapted from my first author publication in Analytical and Bioanalytical Chemistry [60] coauthored with professor Christine Payne.

Chapter 4 moves into a new direction and works toward understanding how the protein corona affects biological responses. In this chapter we characterized the interaction of titanium dioxide NPs with bronchoalveolar lavage fluid (BALF), for the first time, and its subsequent response with macrophage cells. We found that the BALF corona elicited an increased inflammation response relative to the bare NPs, NPs with a fetal bovine serum (FBS) corona, and NPs with an albumin corona. This is surprising and interesting because the BALF corona was still predominantly albumin similar to the other two coronas, suggesting that a lower abundance higher enrichment protein played a role in the increased inflammation response. This chapter is drawn from my first author paper that is currently in review with coauthors Michaela Albright and Professors Rob Tighe and Christine Payne.

Chapter 5 summarizes my work for this dissertation and provides avenues for future study. Some of this future work includes my development of a NP library and machine learning model for protein corona characterization, which are shown in Appendix A. For example, a large magnetic NP dataset with varied sizes, charges, and ligands has been synthesized and characterized. In conjunction those NP-coronas have

been characterized via proteomics to build a machine learning database which random forest regression and classification have been employed on. These results point towards new mechanistic questions to answer and allow us to understand what the predictive capabilities can look like.

2. Automation and low cost proteomics for characterization of the protein corona: Experimental methods for big data[1]

Nanoparticles used in biological settings are exposed to proteins that adsorb on the surface forming a "protein corona." These adsorbed proteins dictate the subsequent cellular response. A major challenge has been predicting what proteins will adsorb on a given nanoparticle surface. Instead, each new nanoparticle and nanoparticle modification must be tested experimentally to determine what proteins adsorb on the surface. We propose that any future predictive ability will depend on large data sets of protein-nanoparticle interactions. As a first step towards this goal, we have developed an automated workflow using a liquid handling robot to form and isolate protein coronas. As this workflow depends on magnetic separation steps, we test the ability to embed magnetic nanoparticles within a protein nanoparticle. These experiments demonstrate that magnetic separation could be used for any type of nanoparticle in which a magnetic core can be embedded. Higher-throughput corona characterization will also require lower cost approaches to proteomics. We report a comparison of fast, low-cost, and standard, slower, higher-cost, liquid chromatography coupled with mass spectrometry to identify the protein corona. These methods will provide a step forward

[1] This chapter comes from my published first-author paper (Poulsen et al. 2020) with OVA NP synthesis and characterization performed by Thomas Pho in the Champion lab at Georgia Institute of Technology. Co-authors on this paper include Thomas Pho, and Professors Julie Champion and Christine Payne.

in the acquisition of the large data sets necessary to predict nanoparticle-protein interactions.

2.1 Introduction

Nanoparticles (NPs) used in any "real world" application encounter a range of molecules that will adsorb on the surface of the NP [104-113]. These adsorbed molecules, referred to as a "corona," determine how the NP functions in downstream applications. The majority of studies in this area have focused on the protein corona that forms when NPs, often developed for nanomedicine applications, interact with blood serum proteins. This protein corona is responsible for subsequent NP-cell interactions including NP clearance from circulation and accumulation in the liver and spleen [30, 114-118]. More recently, the study of biomolecular coronas has been expanded to include lipids and metabolites [119-121]. Beyond these nanomedicine applications, NPs used as anti-microbials, fertilizers, pesticides, and agricultural sensors directly interact with the molecules found in water and soil environments [122-128].

Currently lacking in this research area is the ability to predict what molecules will adsorb on the surface of a NP. Instead, each new NP and surface modification requires an experiment to identify the adsorbed proteins. These experiments consist of incubating NPs and proteins to form a protein corona, removing unbound and weakly bound proteins from the solution, and then carrying out proteomics to identify the adsorbed proteins. This workflow has been described in detail recently [113].

Computational approaches have made progress in terms of predicting the structure of individual proteins as they interact with NP surfaces [129-142], but the ability to predict which protein within a protein mixture will adsorb on the surface of specific NP is not yet possible.

The ability to predict what proteins will adsorb on the surface of NPs would have major implications for the nanomedicine community by saving the time and money associated with experiments. In addition, any methods developed for this community could be extended to the environmental and other NP communities, facilitating work in these areas. To be able to predict what proteins will adsorb on the surface of NPs, we need much larger data sets than are typically used. Most research, including our own [143-146], has tried to predict protein adsorption by working with groups of 2-5 NPs with, for example, varied surface charge, but the small sample size has made it difficult to predict protein adsorption. W.C.W Chan and co-workers synthesized and characterized a library of 105 gold NPs using 3 different gold cores (15 nm, 30 nm, 60 nm) and 67 different ligands (small molecules, polymers, peptides, surfactants) [147]. This size of data set is ideal for corona prediction, but this level of effort and the cost of proteomics for this many samples is outside the scope of many labs.

We identified two steps in the protein corona experimental workflow where time and/or costs could be reduced to enable higher-throughput, lower-cost, experiments that

would allow for the acquisition of larger data sets. The first is the use of an automated liquid handling robot to form and isolate protein coronas. Using a relatively low-cost liquid handler (~$5000 USD), this allows for a significant decrease in human experimental time. The second area is proteomics cost, working with the proteomics core facility to identify lower cost options that provide the necessary sensitivity for corona experiments. Experimental methods and outcomes for each are described below with a comparison between high-throughput and lower cost approaches with standard lab protocols.

We expect that these approaches can be used to generate large data sets that will be essential to recent machine learning efforts in the protein corona community [148, 149]. Robotics have been used previously for the automated synthesis of colloidal NPs [150]. Used in combination with the approaches described below, one could imagine a completely automated workflow with embedded machine learning for optimization [151].

2.2 Materials and Methods

2.2.1 Nanoparticles (NPs)

Three types of NPs were used in the course of experiments. Magnetic NPs (magNPs, 200 nm, carboxylate-modified, #SC0202, Ocean NanoTech, San Diego, CA) were used for automation experiments. Small magnetic NPs, referred to as iron oxide NPs (IONPs, 10 nm, methyl-modified, #IO-A10-5, Cytodiagnostics, Burlington, ON,

Canada), were used as core NPs in the protein NPs, described below. Polystyrene NPs (200 nm, carboxylate-modified, #F8806, Life Technologies, Carlsbad, CA) were used for proteomics experiments. The polystyrene NPs were chosen for proteomics as well-characterized, well-behaved NPs, for testing new proteomics protocols [143, 145, 152-155].

2.2.2 Protein NP synthesis

Ovalbumin (OVA) NPs were prepared using a modified desolvation method [156, 157]; 0.4 mL ethanol was added at a constant rate of 0.5 mL/min to 0.1 mL of 22 mg/mL OVA (#A5503, Thermo Fisher Scientific, Waltham, MA) in phosphate buffered saline (PBS, pH 7.4) under constant stirring at 600 rpm. The protein NPs were collected by centrifugation and resuspended by sonication in 0.5 mL PBS. 2.5 μL of a 50-fold dilution of glutaraldehyde (#G7651, Thermo Fisher) was added to stabilize the protein NPs and stirred for one hour at room temperature. The protein NPs were centrifuged at 14,000 xg and resuspended in 0.750 mL PBS.

2.2.3 Iron oxide nanoparticle (IONP)-doped protein NP synthesis

OVA protein NPs loaded with IONPs were prepared following a similar desolvation method. Methyl-functionalized IONP were sonicated and 10 μL of 5 mg/mL IONP were added to 0.1 mL of 22 mg/mL OVA and stirred briefly. The ethanol desolvation and glutaraldehyde crosslinking were performed exactly as described above

for the undoped protein NPs. IONP-protein NPs were collected by centrifugation at 18,000 xg, resuspended in 0.5 mL PBS and separated using a neodymium magnet (#NB041, Applied Magnets, Plano, TX).

2.2.4 NP characterization

Transmission electron microscopy (TEM) for magNP and polystyrene NP characterization used a Tecnai G² TWIN (FEI, Hillsboro, OR) at the Shared Materials Instrumentation Facility at Duke University. MagNP images were obtained at 160 kV with 25 kX magnification and polystyrene NPs at 120 kV with 20 kX magnification. TEM for protein NPs was carried out on a CX-100 (JEOL, Akishima, Japan) TEM, part of the core facilities at the Georgia Institute of Technology. Protein NP images were obtained at 100 kV with 80 kX magnification. All samples were prepared by drop casting on 400 mesh copper grids (#CF400-Cu, Electron Microscopy Sciences, Hatfield Township, PA) and dried at room temperature. Protein NPs were also stained with 5 μL of 1% sodium phosphotungstate (#496626, Thermo Fisher) solution and washed twice with Milli-Q water. Particle diameters were measured with ImageJ [158]. Average and standard deviation are reported for all measurements.

Hydrodynamic diameter, polydispersity index, and zeta potential of the magNPs (400 μg/mL in water), polystyrene NPs (8 pM in water), and protein NPs (2.93 mg/mL in PBS for size, 0.293 mg/mL in 10% PBS for zeta potential) were measured using dynamic light scattering (DLS; Malvern Zetasizer, Nano-Z, Malvern Instruments,

Worchestershire, England). Measurements were carried out in triplicate with three distinct samples. Each measurement consisted of 12 - 30 runs. Average and standard deviation are reported for all measurements. Electrophoretic mobility was converted to a zeta potential using the Smoluchowski approximation.

2.2.5 Liquid handling robot

A liquid handling robot (OT-2, Opentrons, Brooklyn, NY) with additional electromagnet and temperature modules was used to automate protein corona formation and isolation. Experimental protocols were written in python using Opentrons API v2.1. Pipette tips (20 μl, single and 300 μl, multi) and tipracks used were purchased from Opentrons to verify compatibility and calibration. To run each protocol, the robot was calibrated in the x, y and z directions, followed by calibrating the location of each piece of labware. The locations of each reagent and sample were designated in the script and appropriately positioned after the robot was calibrated. Labware not provided from Opentrons includes a tube rack with base and a water trough. The tube rack and base were 3D printed (3 mm polylactic acid, Ultimaker 2, Utlimaker, Geldermalsen, Netherlands) and the g-code was prepared using Cura (v4.4.1, Ultimaker). The designs for the tube rack were published on thingiverse by *thehair*.

2.2.6 Protein analysis

The protein corona was formed by incubating NPs, both magNPs and polystyrene NPs, with 10% fetal bovine serum (FBS, #10437028, Thermo Fisher). For

automation of experiments, the protein concentration in each wash step was determined by absorbance at 280 nm using a plate reader (SpectraMax iD3, Molecular Devices). UV transparent plates were used for all experiments in order to measure absorption at 280 nm. In practice, less expensive, non-transparent, plates could be used once the number of wash steps had been determined. Absorbance at 280 nm was converted to protein concentration using an extinction coefficient of 43,824 M^{-1} cm^{-1}, the value for bovine serum albumin. While FBS is a mixture of many different proteins, albumin is the most abundant (55%). As a technical note, bicinchoninic (BCA) assays are not directly compatible with the magNPs due to interaction of the copper reagent with the iron oxide of the NPs [159, 160]. Average and standard deviation are shown. P values were calculated using a two-tailed Student's *t*-test.

Gel electrophoresis was used to image individual proteins present in the hard corona of the magNPs. MagNPs with hard coronas (16 μL) were heated in loading buffer (Laemmli loading buffer, #BP-110R, Boston BioProducts, Ashland, MA) for 5 min at 95 ^{o}C and then loaded onto a gel (tris-glycine sodium dodecyl sulfate (SDS) gel, #4561093, Bio-Rad, Hercules, CA) for SDS-polyacrylamide gel electrophoresis (PAGE; 230 V, 35 min). A 10 to 250 kDa molecular weight marker (Precision Plus Protein Dual Color Standards, #1610374, Bio-Rad) was included. Gels were rinsed by microwaving in deionized water (1 min heat, 1 min rocking at room temperature, replace water, x3) and then stained (SimplyBlue Safe Stain, #LC6060, Thermo Fisher) by microwaving until

near boiling (1 min) and then rocked for 5 min. The gel was rinsed in deionized water (10 min, rocking) and NaCl solution (20% w/V, >5 minutes, rocking) and then imaged (PhotoDoc-It, Analytik Jena, Jena, Germany). ImageJ was used for densitometric analysis (Gel Analyzer) and profile plots (Plot Profile) [158].

Proteomic analysis was carried out in the Proteomics and Metabolomics Core Facility, part of the Duke Center for Genomics and Computational Biology. A protein corona was formed on polystyrene NPs using previously published protocols [143, 145, 152-155]. A BCA assay (#23250, Thermo Fisher) was used to determine protein concentration. The protein-NP sample was processed by the Core Facility including in-solution trypsin digest and addition of an internal standard of yeast alcohol dehydrogenase (#186002328, Waters Corporation, Milford, MA) for quantification. This single sample was then analyzed, in triplicate, using two different liquid chromatography (LC) instruments, referred to as MicroFlow and NanoFlow for a total of 6 experiments. MicroFlow LC used a 1 mm x 100 mm ultra-performance liquid chromatography (UPLC, M-Class, Waters Corporation; 80 μL/min) column with a 17 minute total elution time. NanoFlow LC used a 75 μm x 250 mm UPLC column (nanoAcquity, Waters Corporation; 400 nL/min) with a 60 minute total elution time. Both methods use an acetonitrile gradient (5-40%) with 0.1% formic acid. Peptide fragments were analyzed using in-line tandem mass spectrometry (Orbitrap Fusion Lumos, Thermo Fisher). Data was analyzed using Mascot (v. 2.5.1, Matrix Science,

Boston, MA) and Scaffold (v. 4.10.0, Proteome Software, Portland, OR). A detailed description of the Core protocol is provided in Supporting Information.

2.3 Results and Discussion

2.3.1 NP Characterization

DLS (3 distinct samples) and TEM were used to characterize the NPs used for experiments (Table 2-1 and Figure 2-1).

Table 2-1: Characterization of NPs

NPs	d_h (nm)	PDI	ZP (mV)	TEM (nm)
magNPs #SC0202	260 ± 2	0.13 ± 0.02	-32 ± 4	129 ± 43 n = 85
IONPs #IO-A10-5	9.48 ± 1	0.20 ± 0.01	-38 ± 3	10.5 ± 1.3 n = 71
polystyrene NPs #F8806	307 ± 2	0.10 ± 0.02	-36 ± 0.5	241 ± 77 n=50

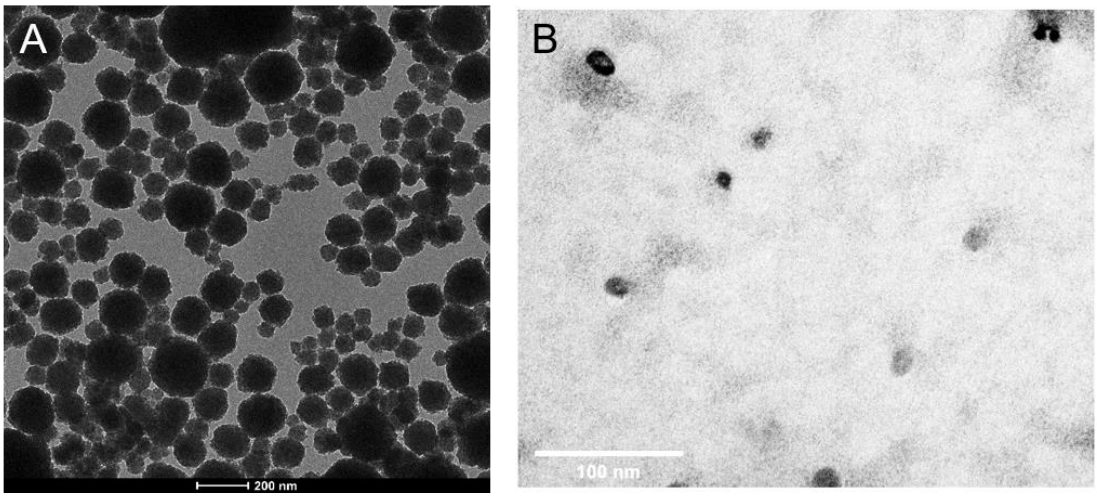

Figure 2-1: TEM images of magnetic NPs. A. MagNPs used in automation experiments. Scale bar is 200 nm. B. IONPs used as magnetic cores in protein NPs. Scale bar is 100 nm.

Carboxylate-modified polystyrene NPs, which are well-behaved NPs that have been characterized extensively in the Payne Lab [143, 145, 152-155], were used for proteomics experiments. A TEM image of the polystyrene NPs has been published previously along with the DLS and TEM values (Table 1) [155].

2.3.2 Protein Corona Formation and Characterization

The use of a liquid handling robot is expected to increase throughput and reproducibility in protein corona formation and characterization. Translating experiments done by a person to the robot includes changing the separation steps used to remove unbound and weakly bound proteins and isolate the hard corona. Our lab, and many others, uses centrifugation to pellet the NPs and remove the supernatant. The NPs are then resuspended in water or buffer and this step is repeated until no protein is detected in the supernatant, typically measured by BCA assay. The protein that remains bound to the NPs is defined as the hard corona. Moving samples to a centrifuge, even in a 96-well format, is a bottleneck in the workflow. Instead, we use a magnetic base within the robot frame to separate NPs from proteins using a similar washing process, but with a magnetic pull down step in place of centrifugation. This does require using magNPs either as stand-alone NPs or as a magnetic core embedded within an NP of interest, as described below.

The protein corona was formed by incubating magNPs (3.2 g/L) with 10% FBS in PBS for a total volume of 250 μL. FBS was chosen as the recommended protein nutrient source for many cell lines, making these results immediately relevant to *in vitro* experiments. All protocols described in this manuscript could also be used for human, or other, serums. Experiments were carried out in 96-well plates working with 8 samples simultaneously. Corona formation, resulting in protein-magNP complexes, was carried out at room temperature for 30 minutes with automated mixing by pipetting up and down every 3 minutes during the 30 minute incubation. Removal of unbound and weakly bound proteins, typically referred to as "washing" in the protein corona community, was carried out using a magnetic base module to separate the protein-magNPs from the proteins in solution. The washing process was initiated by engaging the magnetic base, removing 80% of the volume (200 μL) in each well and replacing it with an equal volume of PBS. For each resuspension the solution was mixed by pipetting up and down (150 μL x3) by the robot. For the first wash, we waited 10 min prior to removing solution to allow time for pull down of the magNPs. Subsequent wash steps used 5 min incubations. The solution that was removed at each wash step was saved for analysis to measure protein concentration (Figure 2-2).

The goal of the wash process is to identify the wash step at which all unbound and weakly bound protein has been removed, with only NP-bound protein remaining. The protein concentration in each wash step was determined by absorbance at 280 nm

using a plate reader. Absorbance at 280 nm was converted to protein concentration (ε = 43,824 M^{-1} cm^{-1}) using Beer's Law. After 4 washes, there is no statistical difference between the subsequent wash solutions (Figure 2-2). The concentration of the magNPs is also confirmed at this point, as it is possible to lose NPs during wash steps. We measure magNP concentration by absorption at 330 nm and comparison to a calibration curve with known concentrations. We observed a 4.2 ± 2.4% (n=8) loss after 10 wash steps. After 5 wash steps, we used gel electrophoresis to characterize the proteins present in the hard corona of the magNPs (Figure 2-3 and 2-6). Densitometric analysis of the two prominent bands at ~70 kDa and ~125 kDa shows that these two bands have average intensities of 3.0 ± 0.5 and 1.3 ± 0.3, respectively. While the standard deviation is relatively low, 18% and 2%, respectively, future experiments will determine the source of this variability.

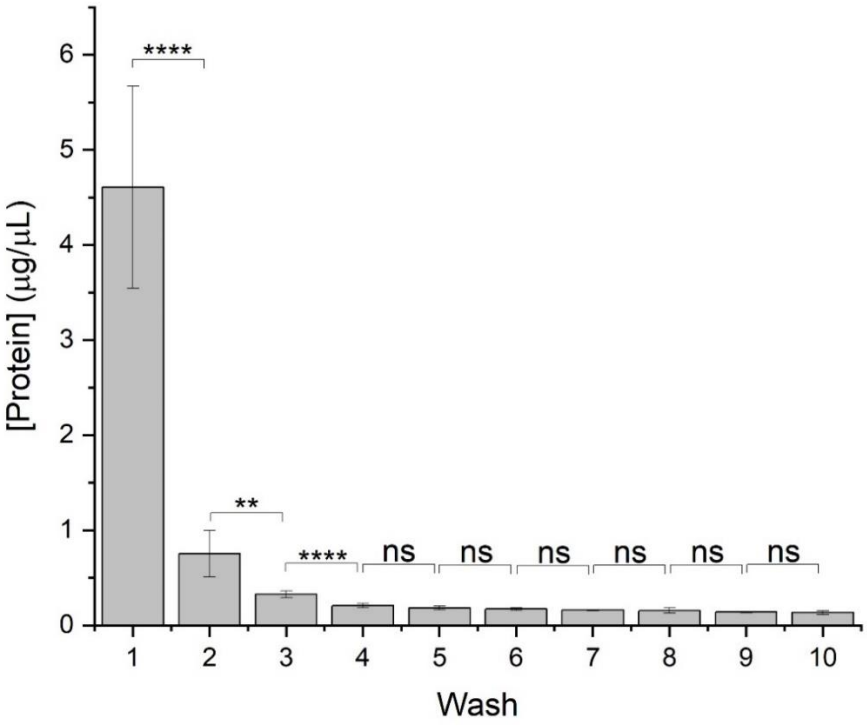

Figure 2-2: Protein concentration (µg/µL) measured at each wash step. n=8; ****p<0.0001, **p<0.01, ns=non-significant.

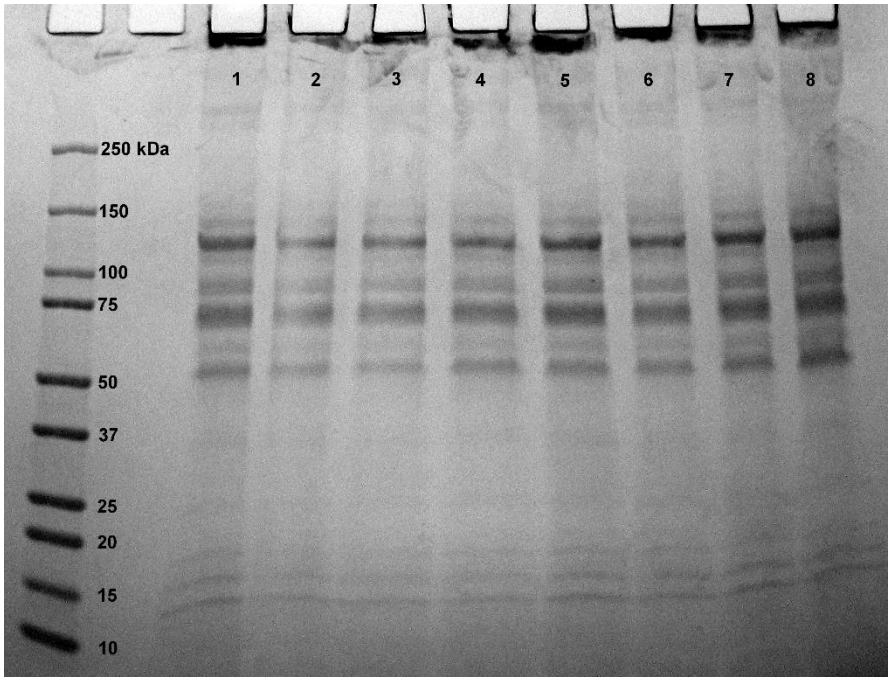

Figure 2-3: Gel electrophoresis of the hard corona adsorbed on magNPs. A profile plot of each lane is shown in Fig. 2-6.

An immediate advantage of the use of the liquid handling robot is the time saved by the human researcher. Preparing the robot for a run includes placement of labware and calibration of tips, which takes ~3 minutes of human time. To complete the 5 wash steps to isolate a hard corona requires 66 minutes of robot time, which includes protein-NP incubation, pipetting to mix wells, pull downs, and transferring samples between wells. In comparison, we estimate that preparing 8 samples by hand, working with a multi-channel pipettor and centrifugation, would require 30 minutes of active human time and 75 minutes of waiting (30 min incubation, 15 minute centrifugation x3). Using magnetic pull down instead of centrifugation would decrease waiting times, but does not change the time necessary for a human to mix and transfer samples.

2.3.3 Protein NPs with Magnetic Cores

Although this automated workflow does require magnetic NPs, magnetic NPs can be modified to vary surface properties or can be embedded within another material. Surface properties, rather than the composition of the NP core, control protein corona formation [146, 147, 161]. For example, our previous work compared the corona-mediated cellular response to iron oxide and polystyrene NPs functionalized with the same carboxymethyl dextran ligand and found that the ligand, rather than the NP core, was the key feature [146]. To demonstrate that magnetic NPs could be embedded within

another material, we fabricated OVA protein NPs by desolvation with and without magnetic iron oxide NPs (IONPs) embedded inside. OVA NPs were chosen since these are model vaccine nanoparticles. This platform is highly tunable and has been used for a variety of proteins, such as influenza antigens [157, 162, 163].

IONPs, ~20x smaller than the protein NPs, were selected so they would be trapped within the protein NPs instead of simply adsorbing the OVA proteins on their surface. For sufficient magnetic strength, multiple IONPs need to be embedded in each NP. IONPs were incorporated within OVA NPs, as indicated by TEM (Figure 2-4). The IONPs appeared to be clustered within the NP.

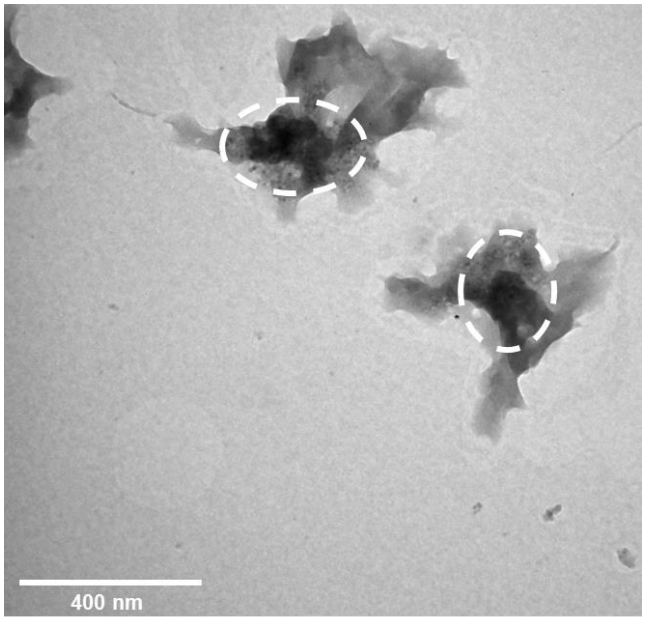

Figure 2-4: TEM image of OVA-IONPs containing clusters of 10 nm IONPs (white ovals). Scale bar 400 nm.

Statistical analysis of DLS and zeta potential measurements indicated that there was no difference in OVA NP size and surface charge with the embedded IONPs (Table 2-2, Figure 2-7 and 2-8). The hydrodynamic diameters for OVA and IONP-OVA NPs were 274 ± 16 and 251 ± 27 nm, respectively ($p = 0.5363$). Zeta potentials were -27 ± 2 mV and -26 ± 1 mV, respectively ($p = 0.8955$). In comparison, the IONPs had a hydrodynamic diameter of 9.48 ± 0.80 nm and a zeta potential of -38.3 ± 3.0 mV (Table 2-1). The IONP-OVA NPs were easily separated from solution with a small magnet, indicating they can be used in the same high-throughput protein corona workflow. All characterization of IONP-OVA NPs was performed after magnetic separation and resuspension.

Table 2-2: Characterization of protein NPs.

NPs	d_h (nm)	PDI	ZP (mV)
OVA NP	274 ± 16	0.31 ± 0.05	-27 ± 2
IONP-OVA NP	251 ± 27	0.22 ± 0.06	-26 ± 1

2.3.4 Low Cost Proteomic Analysis of the Protein Corona

Duke's Proteomics Core Facility, and likely that of other core facilities, uses the time necessary for an experiment to set the cost for staff support and reagents. This means that faster experiments lead to lower user fees. In the case of the liquid chromatography (LC) and mass spectrometry (MS) used for protein corona analysis (LC-MS/MS), the use of high flow LC systems (MicroFlow) can decrease the time, and price,

of the experiments by a factor of 2 compared to NanoFlow LC, which is the standard method in our Core. The subsequent mass spectrometry is identical for both separation methods. Our analysis of the corona formed on polystyrene NPs using both LC systems, with identical samples, showed that of the 25 most abundant proteins identified in the corona, 23 were identified using both LC methods.

The most abundant protein identified by both LC methods was complement C3. Despite being the most abundant protein in serum, serum albumin was only the 14th most abundant protein adsorbed on the surface of the polystyrene NPs. This result, with complement C3 dominating the corona, is similar to previous proteomics analysis of nearly identical carboxylate-modified polystyrene NPs (200 nm) carried out in 2015 [152], which differed only in the fluorophore. Previous results also showed that the corona increases the hydrodynamic diameter of carboxylate-modified polystyrene NPs and decreases the zeta potential, reflecting the charge of the adsorbed protein [143]. The relative enrichment of complement C3 and depletion of albumin in the corona has also been observed by our lab with titanium dioxide NPs [152], and by others working with gold nanoparticles [147], iron oxide nanoparticles [164], and iron oxide nanoworms [165]. While this observation points towards the importance of complement C3 as an immune protein recognizing an invader, the underlying biophysical mechanism has not been determined.

The list of top-25 proteins identified with MicroFlow was used to generate a histogram of MicroFlow and NanoFlow results ordered by total spectral counts (Figure 2-5). The histogram shows the average and standard deviation of 3 repeats on a single sample. The four proteins that differed between MicroFlow and NanoFlow were both identified by the other method, but were not within the top-25 most abundant proteins. For example, carboxypeptidase B2 and alpha-2-HS-glycoprotein were the 23th and 25th most abundant proteins identified by MicroFlow and the 27th and 28th most abundant identified by NanoFlow. The proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner respository with the dataset identifier PXD01911 [166].

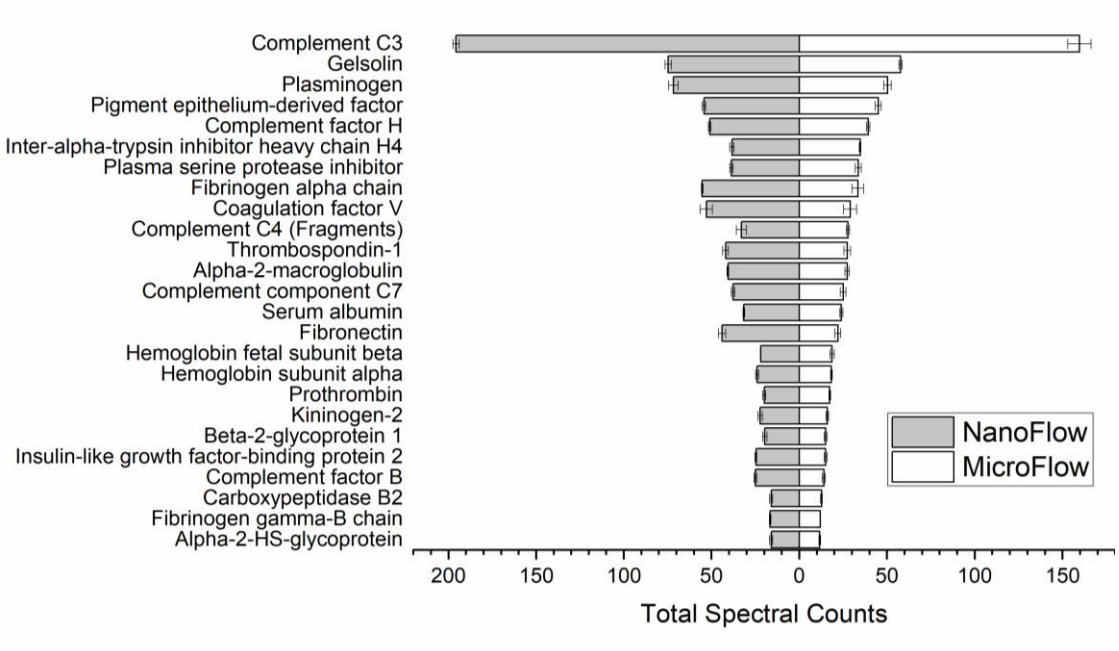

Figure 2-5: Comparison of total spectral counts obtained with low-cost MicroFlow LC (white) and standard NanoFlow LC (gray). The same biological sample

was used for each LC method. The mean and standard deviation are shown for n=3 technical replicates. The complete data sets have been uploaded to the ProteomeXchange with the identifier PXD019118.

The downside of the faster flow system is the need for larger amounts of sample and identification of fewer proteins. Our proteomics core facility, and likely others, requests 5 μg of protein/sample for NanoFlow compared to 20 μg of protein/sample for MicroFlow. The slower flow rate of NanoFlow LC does mean that more proteins will be identified. For the polystyrene NPs, MicroFlow LC identified an average of 87 proteins, compared to 137 with standard NanoFlow LC. The 50 proteins that were not detected with MicroFlow LC had relatively low spectral counts. For example, NanoFlow LC results in 196 total spectral counts for complement C3, the most abundant protein identified in the polystyrene NP hard corona, and 32 spectral counts for serum albumin. Ezrin, with 7.3 spectral counts, was the most abundant protein identified by NanoFlow that was not observed with MicroFlow. In our previous work with albumin and plasminogen, we observed that the most abundant proteins dominate the cellular interaction [143, 146, 154].

2.4 Conclusion

In conclusion, the ability to predict what proteins will be present in the corona of a given NP would save considerable time and money. The first step towards this goal is obtaining significantly larger data sets, which will require automation. We report the use of a liquid handling robot to form and isolate the protein corona (Figure 2-2 and 2-3).

Protein NPs with embedded magnetic cores demonstrate that this automated workflow can be extended to other NPs (Figure 2-4 and Table 2-2). With the large data sets, there is still the challenge of the cost associated with proteomic analysis. To address that, we compared two LC separation methods. The faster speed of MicroFlow LC translates to lower cost, while still successfully identifying abundant corona proteins (Figure 2-5). We hope this work will aid other researchers in the acquisition of large protein corona data sets.

2.5 Supplementary Information

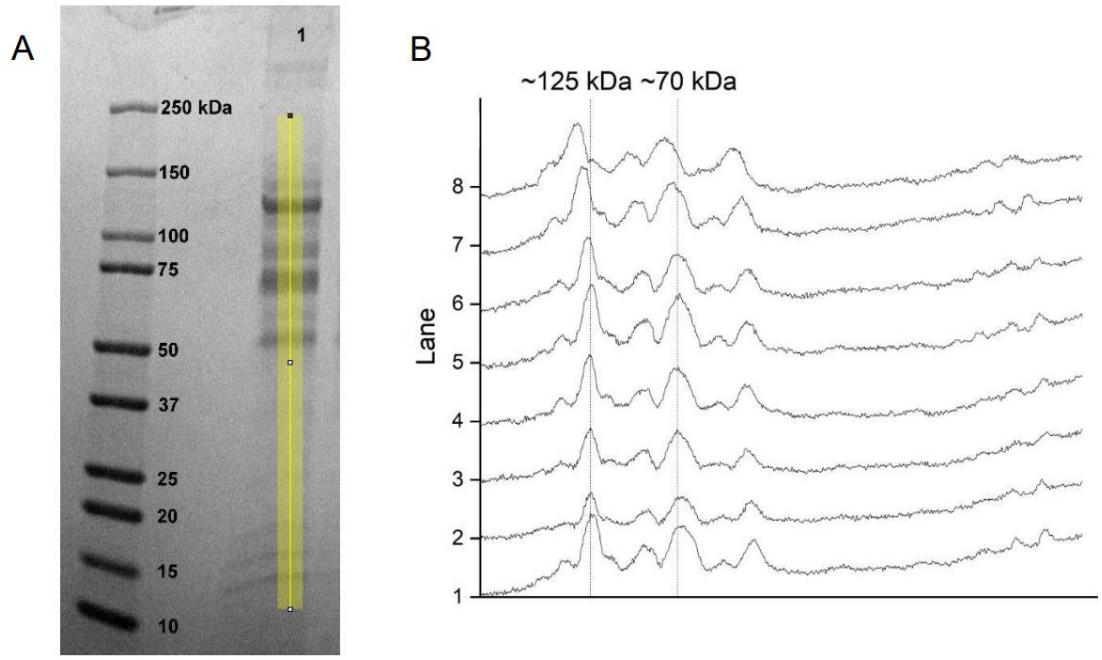

Figure 2-6: Lane analysis (ImageJ, Plot Profile) of the gel shown in Fig. 2-3. A. Example of the line (yellow) used for profile analysis. Lane 1 is shown, all 8 lanes were analyzed. B. Profile intensity plot of all 8 lanes (y-axis). The vertical lines at ~70 kDa and ~125 kDa show the bands that were analyzed by densitometry in the main text.

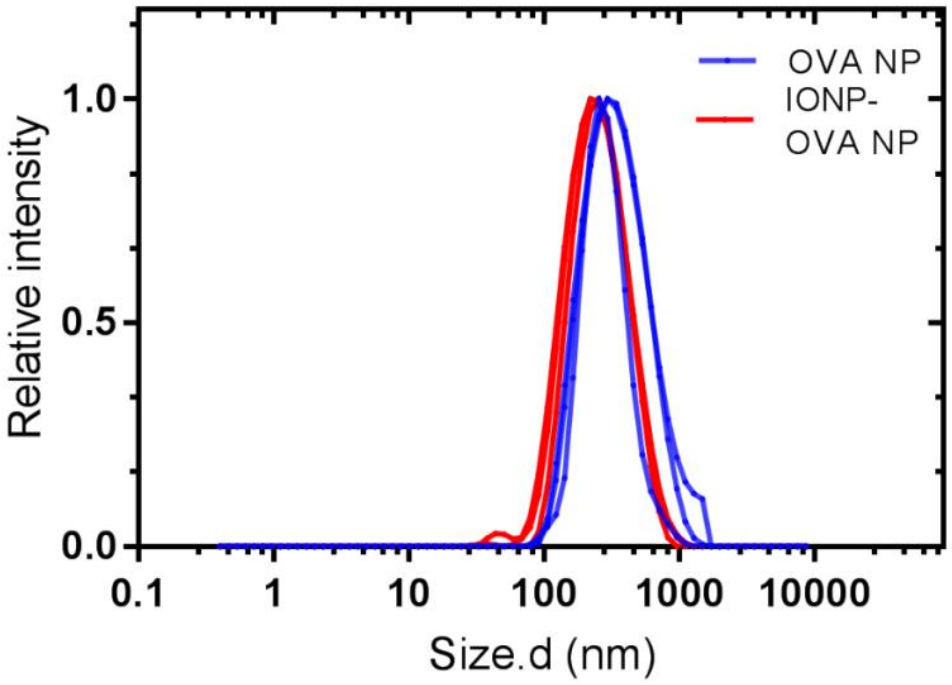

Figure 2-7: Size distribution of individual samples (n=3) of OVA NPs (blue) and IONPOVA NPs (red) measured by DLS.

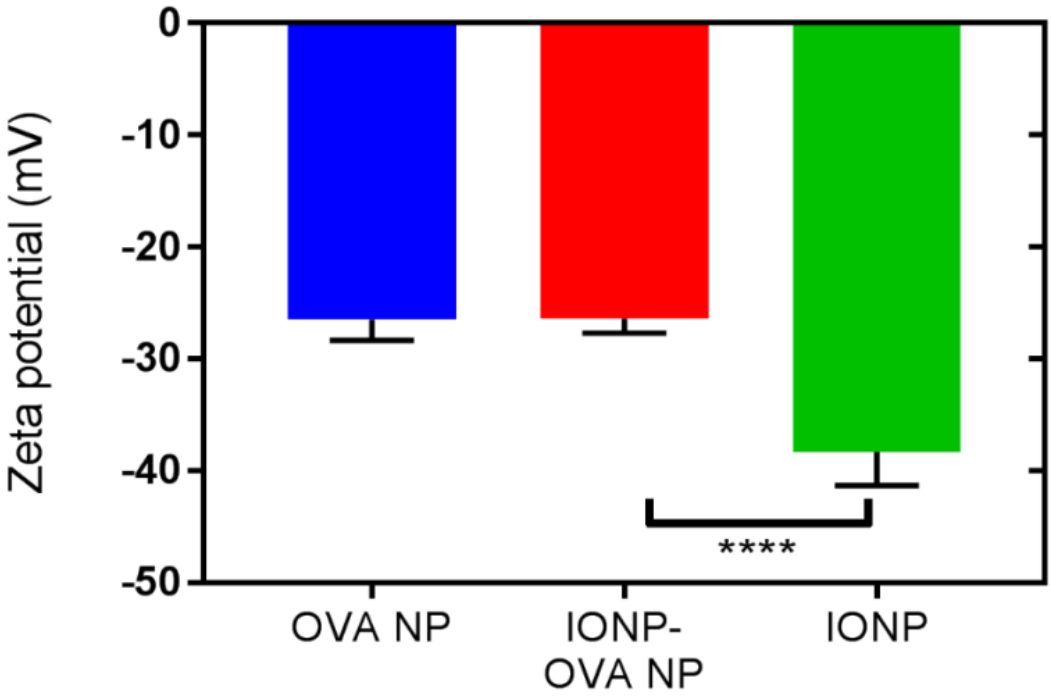

Figure 2-8: Zeta potential of OVA NPs (blue), IONP-OVA NP (red), and IONPs (green). Student's t-test was performed with GraphPad, ** indicates p value < 0.0001.**

2.5.1 LC-MS/MS Proteomics Analysis

Protein-nanoparticle samples in 50 mM ammonium bicarbonate were supplemented with 0.2% RapiGest (Waters Corporation) acid labile surfactant. They were then reduced with 10 mM dithiolthreitol at 40 °C for 30 min and then alkylated with 20 mM iodoacetamide for 45 min at room temperature. Approximately 500 ng of sequencing grade trypsin was added directly to the samples and digestion was allowed

to occur overnight at 37 °C. Following digestion, samples were acidified for 2 hrs with 1.0% trifluoroacetic acid (TFA) to hydrolyze the RapiGest surfactant.

Digested peptides were lyophilized to dryness and resuspended in 15 uL of 0.2% formic acid/2% acetonitrile. Each sample was subjected to chromatographic separation either on a Waters nanoAcquity UPLC equipped with a 1.7 μm high-strength silica (HSS) T3 C18 75 μm I.D. X 250 mm reversed-phase column (NanoFlow data) or a Waters M-Class UPLC equipped with a 1.7 um ethylene bridged hybrid (BEH) C18 1mm I.D. x 10 cm reversed phase column (MicroFlow data). The mobile phase consisted of (A) 0.1% formic acid in water and (B) 0.1% formic acid in acetonitrile. For NanoFlow, 3 μL was injected and peptides were trapped for 3 min on a 5 μm Symmetry C18 180 μm I.D. X 20 mm column at 5 μl/min in 99.9% A. The analytical column was then switched in-line and a linear elution gradient of 5% B to 40% B was performed over 30 min at 400 nL/min. For MicroFlow, 10 uL was injected directly onto the column with a linear elution gradient of 5% B to 40% B was performed over 15 min at 80 uL/min. The analytical column was connected to a Fusion Lumos mass spectrometer (Thermo Fisher) through an electrospray interface operating in a data-dependent mode of acquisition. The instrument was set to acquire a precursor MS scan from m/z 375-1500 at R=120,000 (target AGC 2e5, max IT 50 ms) with MS/MS spectra acquired in the ion trap (target AGC 5e3, max IT 100 ms). For all experiments, higher energy collisional dissociation

(HCD) energy settings were 30 V and a 20 s dynamic exclusion was employed for previously fragmented precursor ions.

Raw LC-MS/MS data files were processed in Proteome Discoverer (Thermo Scientific) and then submitted to independent Mascot searches (Matrix Science) against a SwissProt (Bos Taurus taxonomy) containing both forward and reverse entries of each protein (8,654 forward entries). Search tolerances were 5 ppm for precursor ions and 0.6 Da for product ions using trypsin specificity with up to two missed cleavages. Carbamidomethylation (+57.0214 Da on C) was set as a fixed modification, whereas oxidation (+15.9949 Da on M) was considered a dynamic mass modifications. All searched spectra were imported into Scaffold (v4.4, Proteome Software) and scoring thresholds were set to achieve a peptide false discovery rate of 1% using the PeptideProphet algorithm.

3. Concentration and composition of the protein corona as a function of incubation time and serum concentration: An automated approach to the protein corona[2]

Nanoparticles in contact with proteins form a "corona" of proteins adsorbed on the nanoparticle surface. Subsequent biological responses are then mediated by the adsorbed proteins rather than the bare nanoparticles. The use of nanoparticles as nanomedicines and biosensors would be greatly improved if researchers were able to predict which specific proteins will adsorb on a nanoparticle surface. We use a recently developed automated workflow with a liquid handling robot and low-cost proteomics to determine the concentration and composition of the protein corona formed on carboxylate-modified iron oxide nanoparticles (200 nm) as a function of incubation time and serum concentration. We measure the concentration of the resulting protein corona with a colorimetric assay and the composition of the corona with proteomics, reporting both abundance and enrichment relative to the fetal bovine serum (FBS) proteins used to form the corona. Incubation time was found to be an important parameter for corona concentration and composition at high (100% FBS) incubation concentrations, with only a slight effect at low (10%) FBS concentrations. In addition to these findings, we describe two methodological advances to help reduce the cost associated with protein corona

[2] This chapter comes from my published first-author paper (Poulsen and Payne, 2022) with co-author professor Christine Payne.

experiments. We have automated the digest step necessary for proteomics and measured the variability between triplicate samples at each stage of the proteomics experiments. Overall, these results demonstrate the importance of understanding the multiple parameters that influence corona formation, provide new tools for corona characterization, and advance bioanalytical research in nanomaterials.

This chapter comes from my published first-author paper (Poulsen and Payne, 2022).

3.1 Introduction

Nanoparticles (NPs) present in any biological environment are exposed to proteins that adsorb on the NP surface forming a protein corona. This protein corona determines the biological identity of the NP, controlling cellular internalization, immune response, biodistribution, and circulation time [18, 94, 110, 167-175]. A better understanding of the protein corona is essential for both effective development of nanomedicines and nanosensors and determination of the toxicity associated with human exposure to NPs [7, 8, 176, 177]. The varied properties of NPs (diameter, composition, zeta potential, surface functionalization), the complexity of biological systems (protein structure, hydrophobicity, and charge; concentration of biomolecules), and the effect of the methods used to prepare and analyze the corona makes it difficult to unravel the relevant parameters that control the NP-protein interaction [18, 94, 178, 179]. For example, recent work has highlighted that even the specific type of mixing

process used to form the corona; rocking, vortexing, or flow, will alter the protein corona [93]. Similarly, separation of the protein-NP complexes from unbound protein by centrifugation or magnetic separation will result in a different protein corona [12, 180].

To advance the field, we need to move to larger data sets and improved reproducibility, while also considering cost and accessibility of experiments [103, 181]. For example, liquid chromatography tandem mass spectrometry (LC-MS/MS) is necessary for corona analysis, but it is an expensive technique that requires access to a proteomics core facility. Gel electrophoresis is much less expensive, but often requires LC-MS/MS or western blotting for verification. Obtaining the maximum amount of information from gel electrophoresis experiments would help reduce the costs of protein corona experiments [181]. Machine learning offers the exciting possibility of being able to predict the composition of the protein corona on a given NP, thereby reducing the costs associated with experiments. Existing efforts have focused on protein properties [100-102]. Future work focusing on NP properties would further advance the field. The large data sets required for these experiments will require automated formation, isolation, and analysis of the protein corona. In addition to increasing the number of samples, automation will save time and costs, and improve reproducibility [103, 182].

We describe the concentration and composition of the protein corona as a function of incubation time (30 min and 1 day) and the concentration of fetal bovine

serum (FBS) used to form the corona (5%, 10%, and 100%) on 200 nm carboxylate-modified magnetic NPs, using semi-automated processes. The isolated protein-NP complexes were analyzed using dynamic light scattering (DLS), a bulk protein colorimetric assay, gel electrophoresis, and LC-MS/MS. To understand the need, and costs, for replicates in LC-MS/MS studies, sample replicates were compared at each step of the process. Our experiments show that incubation time is an important parameter for corona concentration, but only at high (100% FBS) incubation concentrations. Quantifying the enrichment and depletion of corona proteins relative to the FBS used to form the corona provides a compositional signature of the protein corona that points us toward protein features that influence corona formation, identifying individual proteins for future studies. Proteomics data shows a shifting composition of corona proteins as a function of incubation time and concentration, but mostly of low abundance proteins. In addition to this corona characterization, we describe a protocol for the automation of the digest step necessary for proteomics. We hope these results will clarify the importance of incubation time and concentration as interdependent variables in corona formation and help to advance corona experiments by reducing the costs.

3.2 Materials and Methods

3.2.1 Nanoparticles (NPs)

Magnetic nanoparticles (200 nm, carboxylate-modified, #SC0202, Ocean NanoTech, San Diego, CA) were used for all experiments.

3.2.2 NP characterization

NP diameter was measured with transmission electron microscopy (TEM) and dynamic light scattering (DLS). TEM was carried out using Tecnai G² TWIN TEM (FEI, Hillsboro, OR) at the Shared Materials Instrumentation Facility at Duke University. NP images were obtained at 160 kV with 25 kX magnification. All samples were prepared by drop casting on 400 mesh copper grids (#CF400-Cu, Electron Microscopy Sciences, Hatfield Township, PA) and drying at room temperature (RT) for 12-18 hrs. Nanoparticle diameters were measured using ImageJ [183]. Average and standard deviation are reported for all measurements.

Hydrodynamic diameter, polydispersity index, and zeta potential of the NPs (100 μg/mL in phosphate buffered saline (PBS) diluted 1:100 in ultrapure water) were measured using DLS (Zetasizer, Malvern Instruments, Worcestershire, England). Measurements were carried out in triplicate with three distinct samples. Each measurement was performed for 12 - 30 runs. The average and standard deviation are reported for all measurements. Electrophoretic mobility was converted to zeta potential using the Smoluchowski approximation.

3.2.3 Liquid handling robot

A liquid handling robot (OT-2, Opentrons, Brooklyn, NY) with a magnetic baseplate and temperature-control module was used to automate protein corona formation and isolation, as described previously [103]. Protocol scripts were written in

python using Opentrons API v2.11. Pipette tips (300 μl, single and multi) and tip racks were purchased from Opentrons to verify compatibility and calibration. The locations of each reagent and sample were designated in the script and appropriately positioned prior to running the robot. Most experiments used a 96-well plate with three or six replicates, specified in the text. Within a row of eight wells, two wells were used for background subtraction.

3.2.4 Protein corona formation and quantification

A protein corona was formed by incubating NPs (3.2 mg/mL) in 1-100% solutions of FBS (#10437028, Thermo Fisher Scientific, Waltham, MA) diluted in PBS. The incubations were performed at RT on a microplate shaker for times ranging from 30 minutes to 1 day, as specified in the text. To remove unbound proteins, the NPs were "washed" by the robot. Each wash step consisted of a magnetic pull down using the magnetic baseplate, removal of the supernatant, and then resuspension in an equal volume of PBS. The hard corona is defined as the protein that remains bound to the NPs with no unbound protein detected in the supernatant, as described below. Centrifugation is typically used for removal of unbound proteins. The use of magnetic pulldown with the liquid handling robot provides a faster workflow [103]. The number of washes required to remove unbound protein was dependent on the concentration of FBS used for corona formation (Fig. 25).

Protein concentration was measured with the Pierce 660 nm Protein Assay Reagent (referred to as a 660 nm assay; #2260, Thermo Fisher Scientific) with the addition of Ionic Detergent Compatibility Reagent (#22663, Thermo Fisher Scientific) according to the manufacturer's instructions. The concentration of protein present in the hard corona was determined by removing the proteins from the NPs by incubating with sodium dodecyl sulfate (SDS) buffer (5% w/v, #L3771, Sigma-Aldrich, Burlington, MA) for 30 minutes at RT. Protein concentration was then determined by measuring absorbance at 660 nm using a plate reader (Spectramax iD3, Molecular Devices, San Jose, CA). A residual amount of protein is resistant to SDS removal independent of duration of SDS incubation (Fig. 26). The NP concentration after the washing process was measured by absorption at 440 nm and comparison with a calibration curve of known concentrations. Protein concentration is reported as protein relative to NP concentration (μg protein/mg NPs).

3.2.5 Protein digestion for proteomics

Prior to proteomic analysis, samples were digested using a modified S-Trap mini column (Protifi, Farmingdale, NY) protocol. Proteins were removed from the NP surface by incubating with SDS buffer for 30 minutes. Protein concentration was determined using the 660 nm assay. Samples were pooled to load a minimum of 25 μg of protein on each S-Trap. Two modifications were made to the S-trap protocol: dithiothreitol (DTT; #R0861, Thermo Fisher Scientific) and iodoacetamide (IAM; #I1149, Sigma-Aldrich) were

used as the reducer (20 mM) and alkylator (40 mM), respectively. DTT and IAM are commonly used for proteomics and are recommended substitutions. Following the completion of the S-trap protocol the resulting digested proteins were lyophilized and stored at -20 °C until proteomic analysis.

3.2.6 Proteomic analysis

Proteomic analysis was carried out in the Proteomics and Metabolomics Core Facility, part of the Duke Center for Genomics and Computational Biology, as described previously [103]. In brief, digested samples were analyzed using LC-MS/MS with ≤ 25 µg of digested protein injected. MicroFlow LC was performed with an ultra-performance liquid chromatography (UPLC, 1 mm x 100 mm, M-Class, Waters Corporation; 80 µL/min) column and a 17 minute total elution time. The column was run with an acetonitrile gradient (5-40%) with 0.1% formic acid. Peptide fragments were analyzed using in-line tandem mass spectrometry (Orbitrap Fusion Lumos, Thermo Fisher).

To further develop our low-cost workflow, we analyzed the LC-MS/MS data using MaxQuant (v2.1.0, Max Plank Institute, Munich, Germany), an open source software designed to qualitatively and quantitatively analyze mass spectrometry data [184, 185]. The raw LC-MS/MS spectra were searched, using their integrated Andromeda search engine, against the Swiss-Prot Bos Taurus canonical protein database from UniProt, accessed on April 27, 2022. A custom contaminants file was used while searching, which contained a relevant subset of the Common Repository of Adventitious

Proteins (cRAP) database [186]. For protein and peptide quantification and identification, default MaxQuant parameters were used including a 0.01 false discovery rate, a minimum peptide length of 7 amino acids, a maximum peptide length of 25 amino acids, oxidation and acetyl groups as variable modifications, and carbamidomethyl as a fixed modification. The Top3 label-free protein quantification method was used, as described previously for shotgun proteomic studies of the protein corona [187-189]. This quantification method averages the integrated intensity of the three most intense peptides for each protein.

Resulting proteomic data was analyzed and filtered in Perseus (v2.0.3.1, Max Plank Institute). Proteins were excluded if they were considered contaminants, quality control standards, or were not observed in at least 70% of the samples. This 70% requirement, the default for Perseus, is appropriate for similar samples such as these. After filtering in Perseus, 205 proteins were observed across 15 different samples. Data normalization, heatmap and Venn diagram generation were performed in Python (v3.9, Python Software Foundation, Beaverton, OR). A quantitative internal standard was not used for these experiments. Low abundance proteins in serum, defined as less than 3 peptides observed, had abundances set to 3% of the lowest observed abundance, which allowed fold change to be calculated. To correct for any change in performance or differences in protein loading, each sample was normalized to itself by dividing by the mean of the interior 80% of the protein intensities [190]. Each sample was scaled to have

the same average. We report these values as percent normalized abundance. Fold change for each protein was calculated by taking the log base 2 of the normalized abundance of samples divided by serum.

Our discussion focuses on the 50 most abundant proteins present in the coronas. The total number of observed proteins is larger: 205 proteins were observed in the protein corona formed from a 30 min, 10% FBS incubation condition and 60 proteins were identified in FBS. Based on the concentration of protein per NP in the 30 min, 10% FBS sample, we estimate 1,200 proteins are present on each NP. This calculation uses the molecular weight of albumin as a representative protein. Albumin is the most abundant protein in the corona and its molecular weight is near the weighted average of the other corona proteins, making it an appropriate representative protein. Proteins with a normalized abundance of >0.3%, and therefore not one of the 50 most abundant proteins, are calculated to be present at levels less than 4 proteins per NP. At this low level, we assume these low abundance proteins are not drivers of biological responses. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the Proteomics Identification Database (PRIDE) partner repository with the dataset identifier PXD035811 and DOI10.6019/ PXD035811 [191].

3.2.7 Gel electrophoresis

Gel electrophoresis was used to visualize individual proteins present in the corona of the NPs, as described previously [103]. In brief, protein coronas were removed

from NPs using SDS, as described above. The removed proteins were heated in loading bufer (Laemmli, #BP-110R; Boston BioProducts, Ashland, MA), incubated for 5 min at 95 °C, and loaded onto a gel (tris–glycine SDS gel, #4561093, Bio-Rad, Hercules, CA) for SDS–polyacrylamide gel electrophoresis (PAGE; 230 V, 35 min). A 10 to 250 kDa molecular weight marker (Precision Plus Protein Dual Color Standards, #1610374, Bio-Rad) was included. Gels were rinsed by microwaving in deionized water (1 min heat, 1 min rocking at RT, replace water,×3), stained (SimplyBlue Safe Stain, #LC6060, Thermo Fisher) by microwaving until near boiling (1 min), and then rocked for 5 min. Gels were destained in deionized water (10 min, rocking) and NaCl solution (20% w/v,>5 min, rocking) and then imaged (PhotoDoc-It, Analytik Jena, Jena, Germany).

3.3 Results and Discussion

3.3.1 NP Characterization

Commercially-available magnetic NPs (200 nm, carboxylate-modified) were used for all experiments. They were characterized by TEM and DLS (Table 3-1) prior to experiments. The use of commercial NPs allows other researchers easy access to these materials. Magnetic NPs and a magnetic baseplate are necessary for high-throughput experiments using the liquid handling robot. Without magnetic separation, NPs require centrifugation to separate NPs from unbound proteins. Centrifugation is a bottleneck in the standard protein corona workflow as NPs are moved to the centrifuge and multiple centrifugation and resuspension steps are required.

3.3.2 Protein corona increases NP diameter and zeta potential

Protein coronas were formed by incubating NPs in FBS. FBS is used as a nutrient source for many cell lines. The concentrations of FBS used to form the protein corona ranged from 1%-100%, as specified in the text. 100% FBS, equivalent to 40 mg/mL of protein as measured by a 660 nm assay, was used to model *in vivo* protein concentrations. Since cells are typically cultured in 10% FBS, 10% FBS was used to model *in vitro* protein concentrations. A hard corona was defined as the protein adsorbed to the surface of the NPs after no unbound protein is present in the supernatant, as determined by a 660 nm protein assay. This required 6 wash steps for NPs incubated with 10% FBS and 12 wash steps for NPs incubated with 100% FBS (Figure 3-6).

Following incubation with 10% or 100% FBS (30 min, RT), the hard corona significantly increased the hydrodynamic diameter (d_h) and zeta potential of the resulting protein-NP complexes, as determined by a one-way analysis of variance (ANOVA) with a post-hoc Tukey test ($p<0.01$), in comparison to the bare NPs. These changes are in agreement with previous work in our lab [76, 192-194], and others [11, 18, 23, 195, 196]. For example, an extensive study of 105 different gold NPs (15 nm, 30 nm, and 60 nm cores; 67 different anionic, cationic, and neutral ligands) showed a similar increase in d_h following corona formation [23]. The change in zeta potential depends on the initial charge of the NP. Both cationic and anionic NPs form a protein corona. The

zeta potential of the protein-NP complex will reflect the charge of adsorbed protein. For example, comparing cationic (amine-modified) and anionic (carboxlylate-modified) polystyrene NPs (200 nm) showed that cationic NPs become anionic following formation of a protein corona and anionic NPs show an increased zeta potential, becoming less negative [76].

Table 3-1: NP diameter (d_{TEM} and d_h), polydispersity index (PDI), and zeta potential (ZP).

d_{TEM} (nm)	[FBS] (%)	d_h (nm)	PDI	Δd_h (nm)	ZP (mV)	ΔZP (mV)
129 ± 43	0	242 ± 4	0.13 ± 0.004	-	-39 ± 1	-
-	10	305 ± 17	0.17 ± 0.04	+63	-34 ± 2	+5
-	100	313 ± 15	0.19 ± 0.05	+71	-33 ± 2	+6

3.3.3 Concentration of the protein corona increases with FBS concentration

The concentration of the protein corona is measured with a colorimetric assay (660 nm assay). We report protein corona concentration (μg)/NP concentration (mg), as described in Materials and Methods. We observed that the amount of protein adsorbed on the NP surface increases as the initial concentration of FBS used to form the corona increases (1% to 100%; Fig. 3-1a). For these experiments, the incubation time was kept constant at 30 min, a standard incubation time for corona experiments [197, 198]. All experiments were carried out at RT.

Previous studies have examined corona concentration as a function of the concentration of serum used to form the corona [37, 39, 196]. While many of these studies have also observed an increase in corona concentration as a function of serum concentration, it is important to note that this result depends on NP composition and type of serum. For example, a side-by-side comparison of silica and polystyrene NPs, both 200 nm, in human plasma (1 hr incubation, 3%-80% plasma) showed an increasing corona on the polystyrene NPs, but a slightly decreasing corona on the silica NPs, with increasing plasma concentrations [196]. A comparison of corona formation on poly(lactide-co-glycolide) NPs (200 nm; 30 min incubation) using human serum and FBS showed different adsorption isotherms for the two different serums [39].

3.3.4 Increasing incubation time increases the concentration of the protein corona: Greater increase at high (100%) FBS incubation concentrations

In a second set of experiments, we varied incubation time (30 min or 1 day) and kept the initial FBS incubation concentration constant at 10% FBS or 100% FBS. For coronas formed with 10% FBS, increasing the incubation time resulted in only a small (15%) increase in the amount of protein adsorbed on the NP surface (Fig. 3-1b). In comparison, incubation with 100% FBS resulted in an increased (59%) corona concentration following a 1 day incubation in FBS (Fig. 3-1b). Much previous work has addressed the time dependence of corona formation including the concentration and composition of adsorbed proteins [36, 74, 75, 90, 195]. Despite extensive study, it has

been difficult to identify a general trend in time-dependent changes in protein corona concentration. For example, a comparison of human plasma protein adsorption on cationic polystyrene NPs (120 nm), anionic polystyrene NPs (120 nm), and anionic silica NPs (35 nm) from 1-1000 min showed an overall increase with time for the anionic NPs, both polystyrene and silica, but with different rates. Cationic polystyrene NPs showed a loss in protein over time. Separate experiments with anionic mesoporous silica NPs (100 nm) showed an increase in protein concentration when incubated with human serum or FBS over a 1-10,000 min period [75].

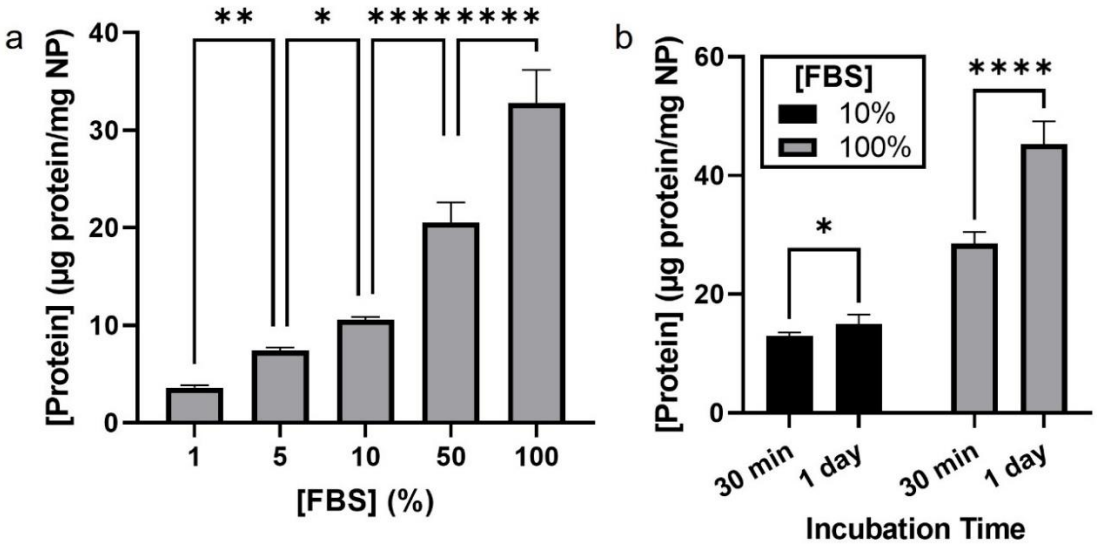

Figure 3-1: Concentration of FBS adsorbed on NPs. (a) Protein adsorption as a function of initial FBS concentration used to form the protein corona. NPs were incubated with FBS for 30 min (n=6). The mean protein per NP (µg/mg) is reported. Error bars show standard deviation. Significance was calculated using an ANOVA with a post-hoc Tukey test. (b) Protein adsorption as a function of incubation time. NPs were incubated with 10% (black) or 100% (gray) FBS for either 30 min or 1 day to form a protein corona (n=6). Error bars show standard deviation. Significance was calculated using an unpaired two-sided t-test. *p<0.05, **p<0.01, ****p<0.0001

3.3.5 Towards low-cost proteomics: Automation and replicates

The best method to determine the composition of the protein corona is LC-MS/MS. At the Duke University core facility, and likely others, the cost of a proteomic experiment scales with the time required to run the samples. Previously, our lab compared the data generated by a lower cost, high-speed method, MicroFlow, and the standard, slower, NanoFlow [103]. We found nearly identical results for the 50 most abundant proteins observed in the protein corona, with a 4x decrease in cost per sample using MicroFlow. Our next steps in reducing costs, described below, have been to develop methods for automated sample digestion and determine the need for replicate samples.

We used commercially available S-Traps to isolate and digest the corona proteins. The process is not currently fully automated as the S-Traps must be manually moved to a centrifuge, but S-Traps can maintain the 96-well plate format. Aside from the centrifugation steps, the digestion is performed by the liquid handling robot, decreasing the hands-on time for the researcher. The digestion requires three different incubations with multiple reagent loading steps. Without the robot, the protocol is estimated to include 25 minutes of hands-on pipetting time, 145 minutes of incubation time, and 5 minutes of centrifuge time for 8 samples. Using the robot removes the hands-on researcher time for the 25 minutes of pipetting. Adding this time saved to previously reported active time savings of 30 minutes [103], shows that 55 minutes of active

pipetting time is saved during automated corona formation and digestion. The time saved increases linearly with the number of samples and also increases reproducibility by removing human pipetting errors.

We tested three aspects of experimental reproducibility relevant to proteomics (Fig. 3-2). Verification of the LC-MS/MS was performed using a study pool quality control (SPQC), in which all samples in each batch are pooled (study pool) and an aliquot of the study pool is injected at the beginning, middle and end of the injection run. SPQC, which is a standard quality control measure [199], verifies that the LC-MS/MS instrument does not vary as a function of injection time. To characterize possible variation due to the digest, a single sample (30 min, 10% FBS) was aliquoted (x3) and digested individually. To measure biological variability, three separate samples (30 min, 10% FBS) were digested individually.

Venn diagrams of the 50 most abundant proteins present in the coronas are used to show the variation in these samples (Fig. 3-2). Variation is very low. As expected, the SPQC has almost no variation with only two different proteins between the three different injection times. The variation of a triplicate digestion of a single sample is similarly low. The true biological triplicate, three individual samples, also shows very low variation with 48 of 50 proteins overlapping across all 3 samples. Overall, with the low variation in the digest and biological replicates, we are confident that digest replicates are not necessary and that biological replicates can be used to focus on specific

questions of interest. Use of a single sample for proteomics has been used in other protein corona studies and may be appropriate for some *in vitro* experiments [23, 75, 200].

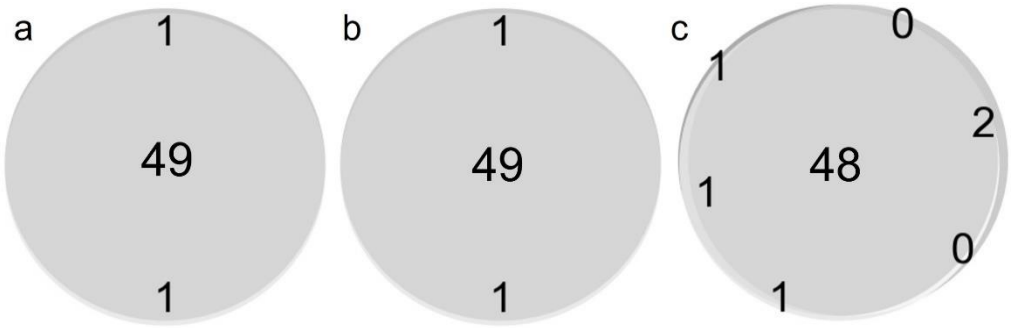

Figure 3-2: Venn diagrams showing variation in proteomic data. (a) SPQC consisting of pooled samples injected at the beginning, middle, and end of the injection run. (b) Digest triplicate obtained from a single sample digested in triplicate. (c) Biological triplicate. The Venn diagrams were created by sorting each sample by the most abundant proteins.

3.3.6 Composition of the protein corona: Abundance and enrichment

Proteomic analysis of protein coronas can be evaluated as the amount of protein detected (normalized abundance; Table 3-2) or the amount of protein in the corona relative to the amount in serum (enrichment, Fig. 3-3). For example, albumin is the most abundant protein in FBS comprising 36% of the protein in FBS (Table 3-2). It is present in the corona at high abundance (8.83 ± 2.16%). However, compared to the very high abundance of albumin in FBS, albumin is depleted in the protein corona (Fig. 3-3). In comparison, prothrombin (7.68 ± 0.63%), the 2nd most abundant protein in the corona, is

only the 45th most abundant protein in FBS (0.05%). It is highly enriched in the protein corona (Fig. 3-3).

Table 3-2: Normalized abundance (%) for the top twenty-five most abundant corona proteins (10% FBS, 30 min incubation, n=3). For comparison, the rank order of proteins in FBS is shown in parentheses.

Protein	Protein corona (%)	FBS (%)
Albumin	8.83 ± 2.16	35.69 (1)
Prothrombin	7.68 ± 0.63	0.05 (45)
Antithrombin-III	6.84 ± 0.21	0.33 (22)
Complement factor H	5.03 ± 0.41	0.11 (37)
Plasma serine protease inhibitor	4.88 ± 1.55	0.15 (33)
Alpha-2-HS-glycoprotein (fetuin-A)	4.25 ± 1.51	26.26 (2)
Tetranectin	4.14 ± 0.16	0.08 (40)
Kininogen-1	3.56 ± 0.28	0.66 (14)
Hemoglobin subunit alpha	3.43 ± 0.05	0.46 (15)
Alpha-1-antiproteinase	3.15 ± 0.57	12.09 (3)
Gelsolin	2.88 ± 0.21	0.12 (36)
Apolipoprotein E	2.77 ± 0.05	0.02 (59)
Thrombospondin-1	2.45 ± 0.14	0.08 (41)
Hemoglobin fetal subunit beta	1.8 ± 0.01	0.28 (25)
Protein AMBP	1.71 ± 0.06	0.07 (43)
Plasma kallikrein	1.56 ± 0.14	0.04 (50)
Histidine-rich glycoprotein	1.43 ± 0.06	0.04 (48)
Kininogen-2	1.42 ± 0.03	0.26 (26)

Beta-2-glycoprotein 1	1.37 ± 0.11	0.36 (18)
Complement C4	1.37 ± 0.12	0.15 (32)
C4b-binding protein alpha chain	1.36 ± 0.04	0 (61)
Hyaluronan-binding protein 2	1.16 ± 0.13	0 (62)
Plasminogen	1.13 ± 0.09	0.21 (29)
Apolipoprotein A-I	1.08 ± 0.11	0.43 (17)
Alpha-2-macroglobulin	1.02 ± 0.08	1.26 (10)

The most abundant proteins observed in this protein corona have been detected in other corona studies [43, 201-204], including our own [103, 200]. For example, albumin, prothrombin, antithrombin-III, complement factor H, plasma serine protease inhibitor, and alpha-2-HS-glycoprotein have been observed in the corona of carboxylate-modified polystyrene NPs (200 nm) [103, 200]. Albumin, antithrombin-III, complement factor H, and alpha-2-HS-glycoprotein were also observed in the corona of anionic TiO_2 NPs (300 nm aggregates) [200]. Complement factor H, the 4th most abundant protein in the corona (30 min, 10% FBS), is of specific interest. It has been observed in the corona of many types of NPs including silica (carboxylate-, amine-, un-modified; 30 nm,140 nm) [195], polystyrene (anionic; 112 nm) [195], and iron oxide (PEGylated, 13 nm) [205]. The presence of complement factor H in the corona has the potential to help avoid the alternative pathway response [42].

While the individual proteins in the corona have been observed in other NP studies, the enrichment or depletion of specific proteins, the "signature" of the protein

corona (Fig. 3-3), is unique to these NPs. This highlights a general challenge in the protein corona field. While there are common corona proteins, it is impossible to predict the specific compositional signature of a protein corona based on previous studies.

A summary view of protein enrichment (or depletion) of corona proteins relative to serum proteins shows that corona composition is relatively insensitive to incubation time or initial FBS concentration (Fig. 3-3). Sample grouping for each protein was assessed with an outlier test (ROUT, Q=5%), resulting in two outliers, hemoglobin subunit alpha (1 day, 10% FBS) and apolipoprotein E (1 day, 100% FBS). In addition to these two outliers, tetranectin and C4b-binding protein alpha chain show differences in enrichment at 10% and 100% FBS that may be of interest for future studies. C4b-binding protein alpha chain has been noted in previous corona studies [82, 206-208], including a high correlation with the uptake of gold NPs (10 nm) by macrophage cells [208].

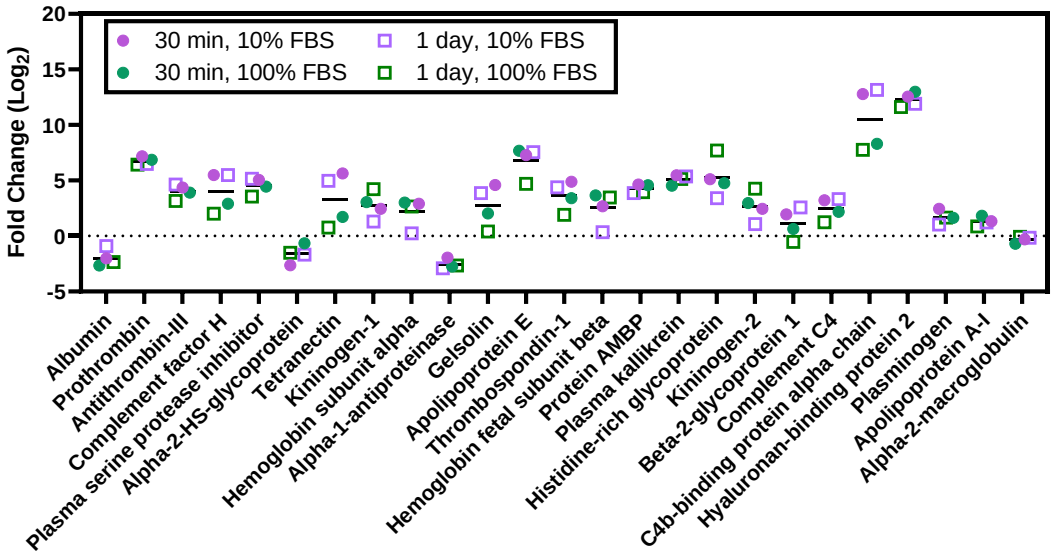

Figure 3-3: Enrichment of the 25 most abundant proteins present in the corona (with top 25 based on the 30 min, 10% FBS incubation sample) relative to their abundance in FBS. Fold change, log base 2, is plotted for each protein at 30 min (filled circles) or 1 day (open squares) incubations in 10% (purple) or 100% (green) FBS.

3.3.7 Increasing incubation time and concentration leads to slight increases in corona differentiation

To gain an understanding of how the composition of the protein corona changes with increasing FBS concentration and time, we ran proteomic analysis of three different FBS incubation concentrations: 5%, 10%, and 100% FBS. For each concentration we examined two incubation times: 30 minutes and 1 day. Venn diagrams of the 50 most abundant proteins present in the coronas are used to illustrate the differences in the protein coronas formed following a 30 min or 1 day incubation with a 10% and 100% FBS incubation concentration (Fig. 3-4). Data for the 5% FBS incubation and hierarchical clustered heat maps for all samples are included in ESM (Figs. 3-9 through 3-11).

For both 10% FBS and 100% FBS incubations, the similarities in the corona following 30 min and 1 day incubations are high with 41 common proteins (of 50) for 10% FBS incubations and 37 common proteins for a 100% FBS incubation. For the most part, the proteins that do change as a function of time are relatively low in abundance. The exceptions to this are serotransferrin, which becomes the 10th most abundant protein in the 10% FBS corona after a 1 day incubation, and cadherin-1, which becomes the 13th

most abundant protein in the 100% FBS corona after 1 day. Neither of these proteins were top 50 proteins following a 30 min incubation.

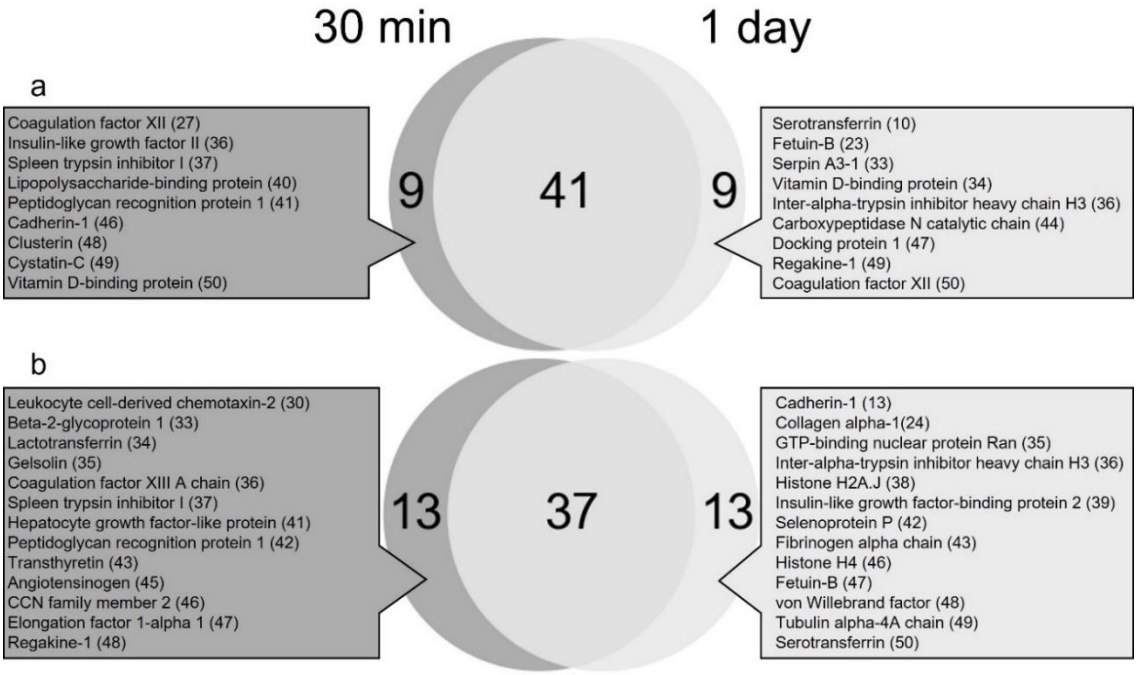

Figure 3-4: Venn diagrams illustrate the differences in the top 50 most abundant proteins following 30 min and 1 day incubations. The rank of each protein is shown in parentheses. (a) 10% FBS. (b) 100% FBS.

A similar approach was used to illustrate the differences in the protein coronas due to incubation concentration (Fig. 3-5). Protein coronas were formed following a 30 min or 1 day incubation with a 10% and 100% FBS incubation concentration. Tetranectin and C4b- binding protein alpha chain are observed in protein coronas formed with 10% FBS, but not 100% FBS. Coagulation factor XI is only observed in high abundance on protein coronas formed with 100% FBS.

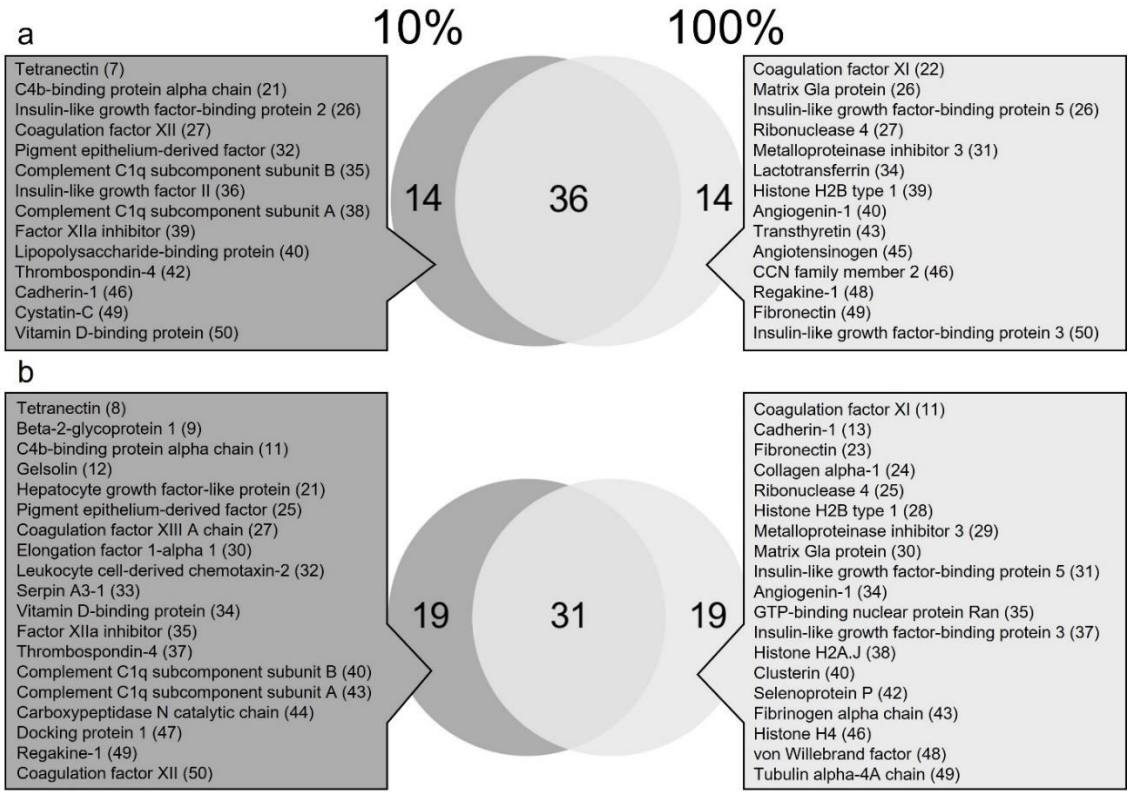

Figure 3-5: Venn diagrams showing the differences in the top 50 most abundant proteins for each incubation time at each incubation concentration. (a) 30 min incubation. (b) 1 day incubation.

An extensive previous proteomics study examined the corona formed on silica (35 nm, 140 nm) and polystyrene (120 nm) NPs, with a range of functional groups, including carboxylate-modified polystyrene, incubated with human plasma [195]. This research showed that the compositional signature of the protein corona was established quickly (30 s) with only the relative concentration of proteins changing over time (30 s - 480 min). Similarly, proteomic analysis of the corona formed from human plasma on polystyrene NPs (50 nm, 100 nm; amine, carboxylate, or no modification) showed no

time-dependence from 5 min- 5 hr [92]. We see similar results in that few new proteins with high (~top 10) abundance are altered as a function of time.

3.4 Conclusions

We describe the automated generation, isolation, and digestion of protein coronas formed on carboxylate-modified magnetic NPs (Table 3-1). We use these methods to characterize the concentration and composition of the protein corona as a function of incubation time and FBS concentration. Our experiments show that increased concentrations of FBS during the formation of the protein corona led to increased concentrations of corona proteins (Fig. 3-1a), as a function of both incubation time and concentration (Fig. 3-1b). The interdependence of time and concentration suggests possible reasons, although not a mechanism, for conflicting reports in the protein corona literature [36, 74, 75, 90, 195].

An overarching theme of our work has been the development of methods to reduce the costs associated with protein corona experiments, including lower cost proteomics [103], and, as described here, automating the digest step typically carried out by a core facility. We also carried out triplicate experiments to measure reproducibility across LC/MS-MS experiments. We find high reproducibility among triplicate samples (Fig. 10). We also compared proteins of interest identified by gel electrophoresis to our results from proteomics with the hope of further reducing costs (Fig. 3-12).

Unfortunately, bands of interest identified by gel electrophoresis did not correspond to results from proteomics.

Using proteomics data, we describe both the abundance (Table 3-2) and enrichment-depletion (Fig. 3-2) of the proteins adsorbed on the surface of the NPs. The composition of the protein corona shows small changes as a function of incubation time and FBS concentration (Figs. 3-4 and 3-5). The most abundant proteins are fairly insensitive to the conditions used to form the coronas.

Moving forward, this research points in two directions. The first is the use of enrichment-depletion data to select proteins for future molecular studies. The enrichment and depletion of proteins in the corona relative to the proteins present in the serum used to form the corona is a purely biophysical interaction, dependent on only the physical and chemical properties of the protein and NP. Understanding why hemoglobin subunit alpha, apolipoprotein E, tetranectin, and C4b-binding protein alpha show unique NP adsorption may help define the molecular interactions governing corona formation. Second, increasing automation and decreasing costs of protein corona characterization will help generate the large data sets necessary for machine learning approaches. We hope to stimulate future studies at both the single protein and big data levels.

3.5 Supplementary Information

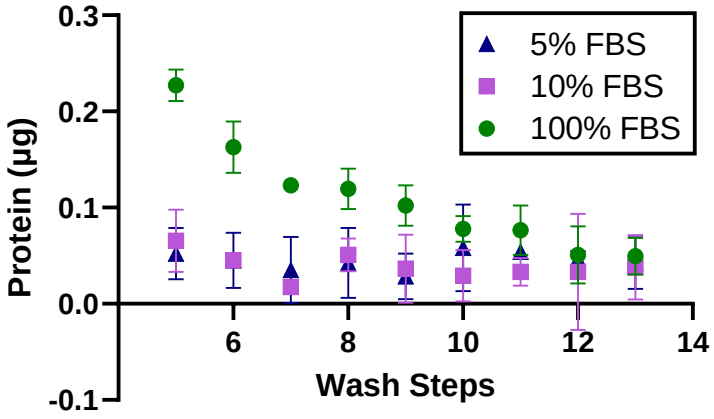

Figure 3-6: Concentration of protein present in the supernatant following "washing" of protein-NP samples. Wash #5 is the first sample shown. The first 4 wash steps show much higher protein concentrations and are not typically measured. NPs were incubated with FBS for 30 min and washed by the robot (n=2). Error bars show standard deviation.

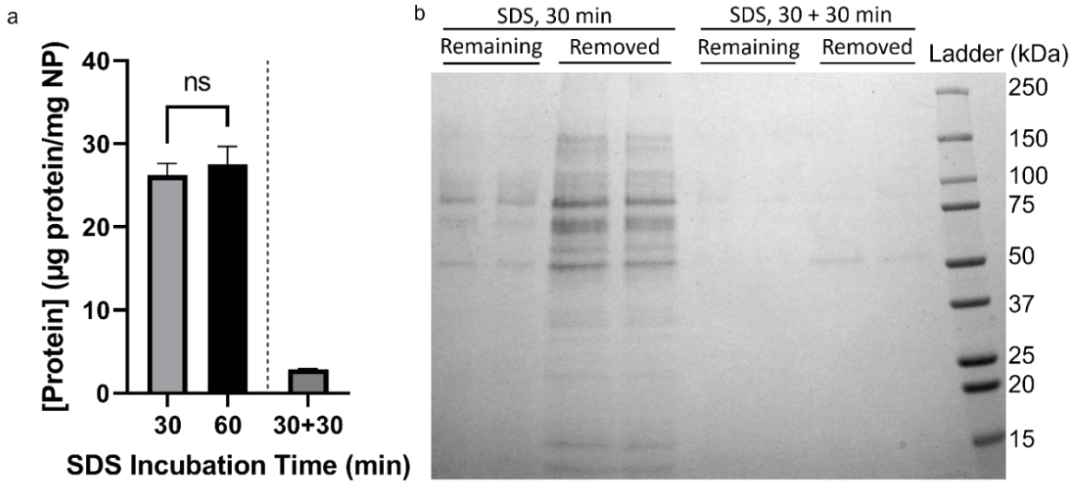

Figure 3-7: Residual protein is resistant to removal by SDS. The bulk of the protein corona is removed following a 30 min SDS incubation, as described in Materials and Methods. (a) There is no significant difference in protein removal following a 30 min (light gray) or 60 min (black) incubation with SDS, measured by the 660 nm assay for coronas formed following a 30 min, 100% FBS incubation. The concentration of protein removed after two sequential 30 min SDS incubation steps (dark gray) is at the edge of the working detection limit of the assay (50 µg/mL). (b) Gel electrophoresis was used to determine if SDS led to selective protein removal.

The proteins visible (Remaining) after SDS treatment (30 min) are the same as the most abundant proteins removed (Removed) by SDS. A second SDS removal step (30+30 min) shows removal of additional protein at ~50 kDa, likely albumin, and residual protein remaining at ~90 kDa. Given the non-significant difference in protein removal, a single SDS removal step was chosen to match with Proteomics Core Facility protocols.

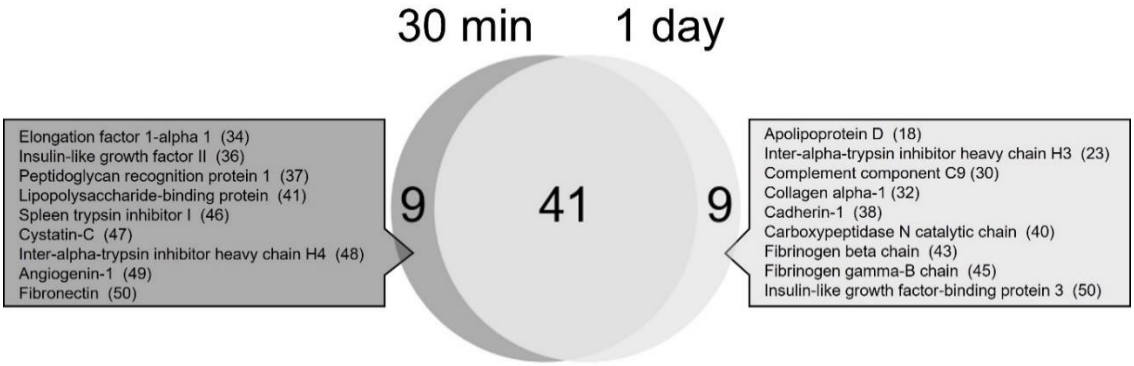

Figure 3-8: Venn diagram showing the differences in the top 50 most abundant proteins following 30 min and 1 day incubations with 5% FBS incubation. Protein ranks are included in parentheses.

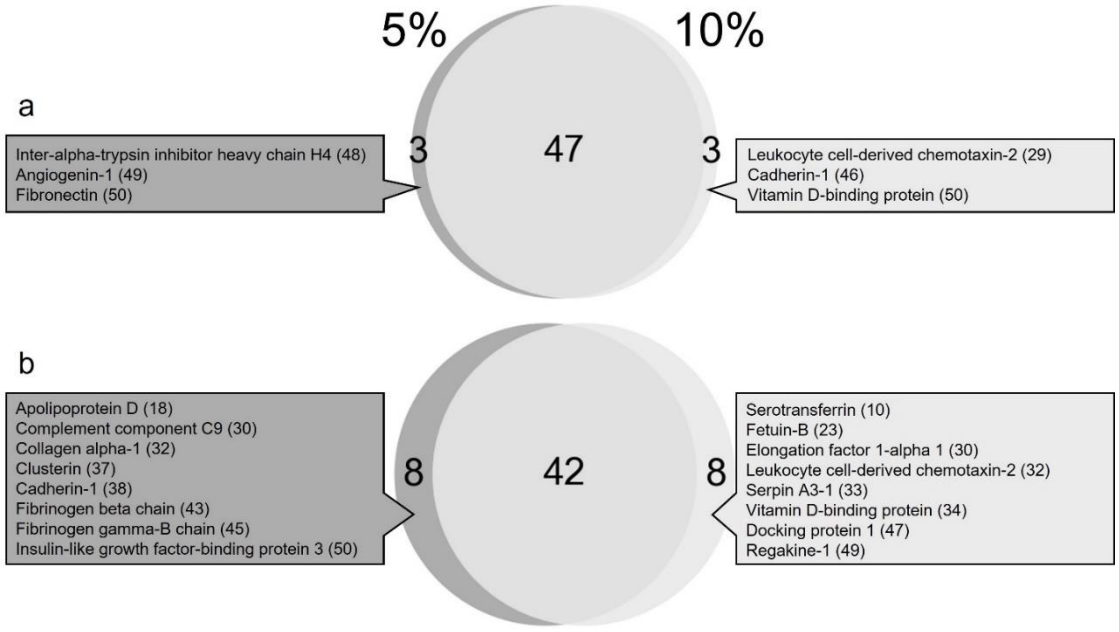

Figure 3-9: Venn diagrams showing the differences in the top 50 most abundant proteins for 5% and 10% FBS incubations. (a) 30 min incubation. (b) 1 day incubation. Protein ranks are included in parentheses.

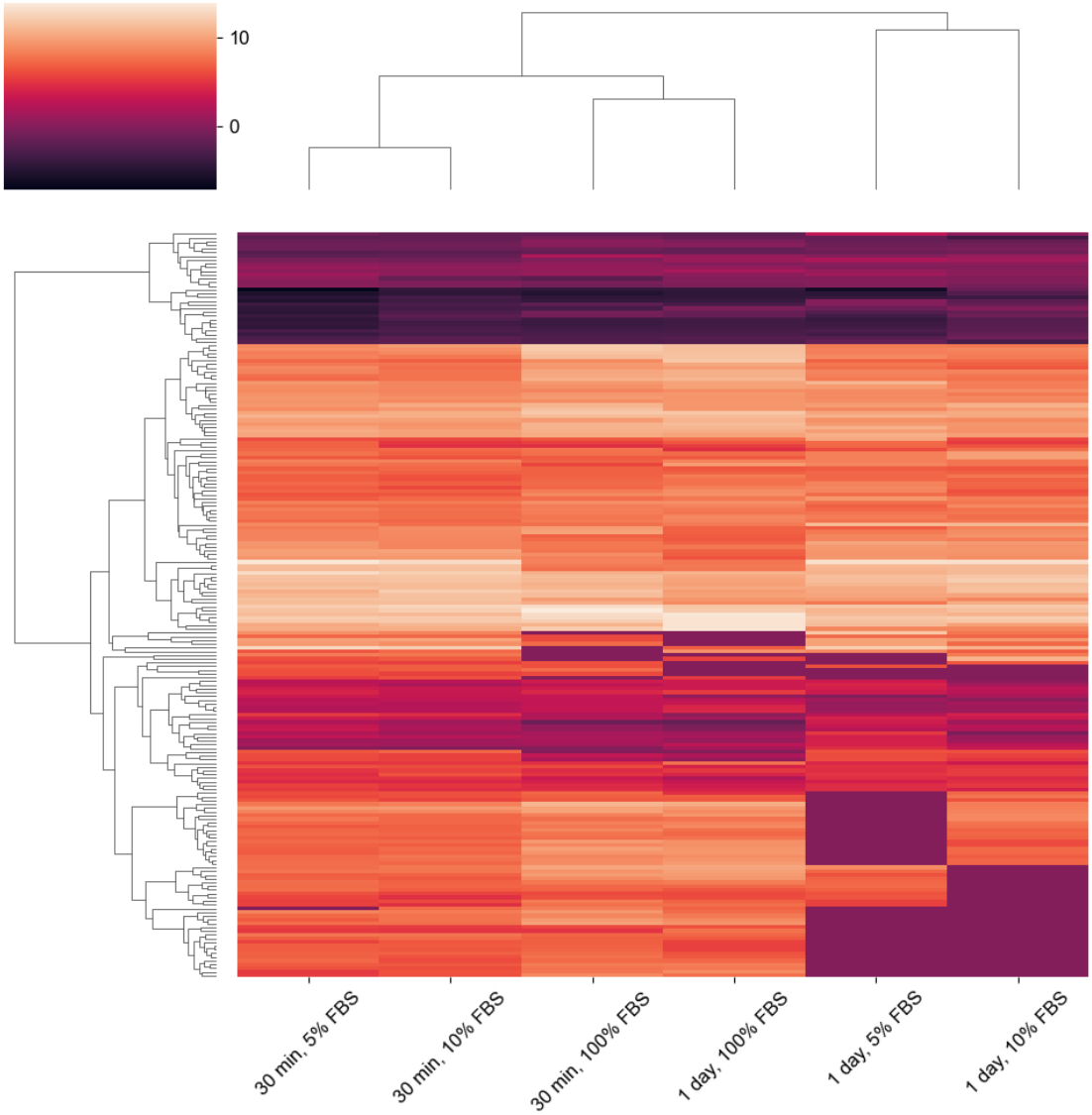

Figure 3-10: Hierarchical clustered heat map of proteomics data with each incubation condition on the x-axis and individual proteins on the y-axis. Fold change data is used for the clustering. 200 proteins are displayed across 6 six different conditions. Fig. 15 shows the top 25 proteins for better visualization.

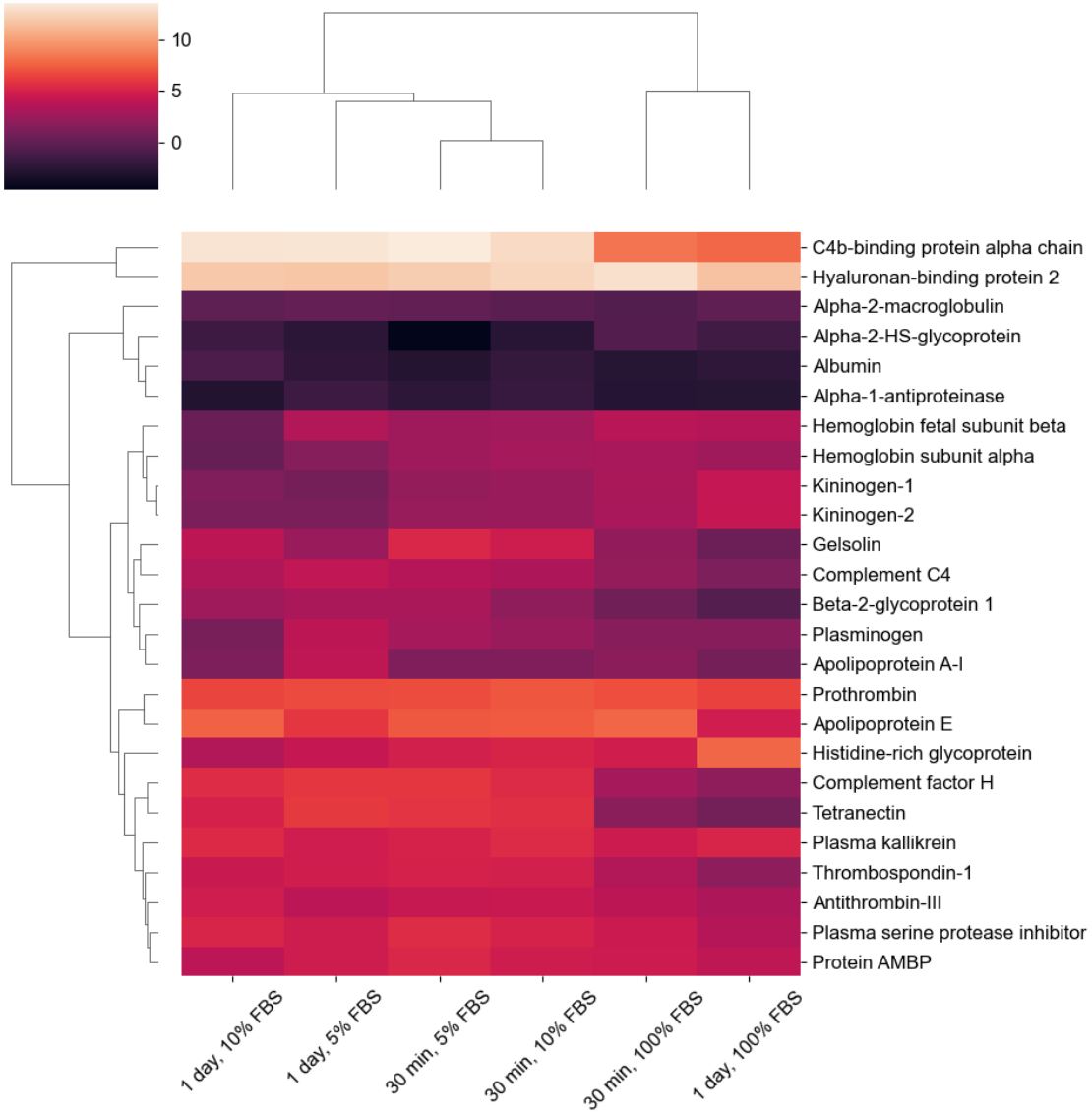

Figure 3-11: Hierarchical clustered heat map of proteomics data with each incubation concentration and incubation time on the x-axis and individual proteins on the y-axis. Fold change data is used for the clustering. The top 25 proteins, ordered by the 30 min, 10% FBS sample are displayed for each incubation condition.

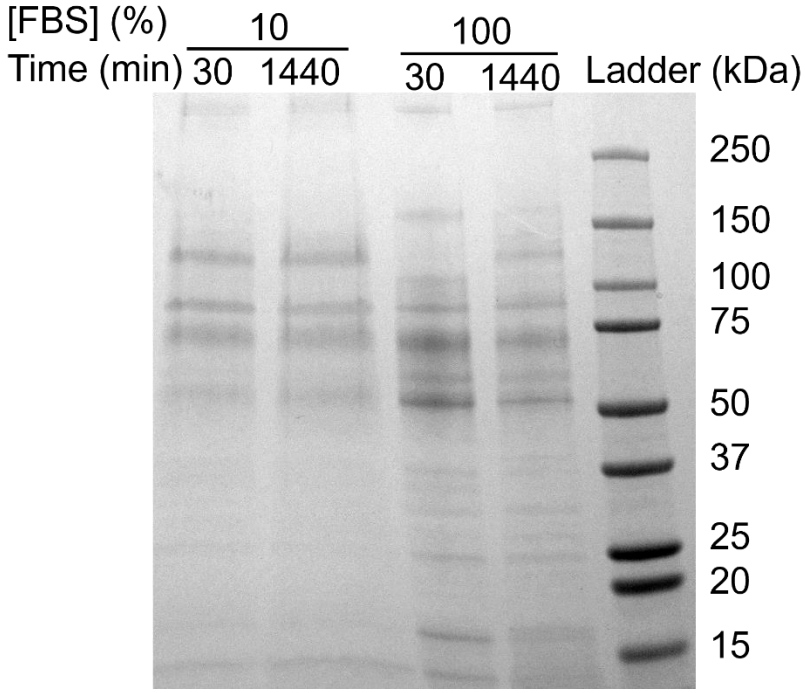

Figure 3-12: Gel electrophoresis of the protein corona formed on NPs as a function of incubation time (30 min and 1 day) and FBS concentration used to form the corona (10%, and 100% FBS). Gel electrophoresis and LC-MS/MS results were compared to see if gel electrophoresis could be used as a less expensive replacement for LC-MS/MS. Gel electrophoresis shows a change in two bands as a function of incubation time and incubation concentration. A band at ~100 kDa is only visible in the 30 min, 100% FBS sample. Based on proteomics this band is likely spondin-1 (MW 91 kDa), as its normalized abundance was ~9 times greater for the 30 min, 100% FBS sample than any other sample. A band observed at ~140 kDa is visible in all of the samples except for 30 min, 100% FBS. In comparison, based on proteomics, no protein is absent from only this sample. The band at ~160 kDa was excised and confirmed to be complement C3 (MW 187 kDa) by LC/MS-MS. In agreement with the gel, complement C3 was observed in higher abundance for 100% FBS than 10% FBS by proteomics, although it is present at all FBS concentrations. Unlike results from the gel, proteomics showed the highest complement C3 abundance at 1 day. Overall, we were unable to correlate gel electrophoresis with LC/MS-MS results. We expect this is a combination of differing levels of protein staining and protein modifications. The lack of correlation between gel electrophoresis and LC/MS-MS points towards the value of working to decrease the costs of proteomics experiments.

4. Interaction of TiO_2 nanoparticles with lung fluid proteins and the resulting macrophage inflammatory response[3]

Inhalation is a major nanoparticle exposure pathway. Following inhalation, nanoparticles first interact with the lung lining fluid, a complex mixture of proteins, lipids, and mucins. We measure the concentration and composition of lung fluid proteins adsorbed on the surface of titanium dioxide (TiO_2) nanoparticles. Using proteomics, we find that lung fluid results in a unique protein corona on the surface of the TiO_2 nanoparticles. We then measure the expression of three cytokines (interleukin 6 (IL-6), tumor necrosis factor-alpha (TNF-α), and macrophage inflammatory protein 2 (MIP-2)) associated with lung inflammation. We find that the corona formed from lung fluid leads to elevated expression of these cytokines in comparison to bare TiO_2 nanoparticles or coronas formed from serum or albumin. These experiments show that understanding the concentration and composition of the protein corona is essential for elucidating the pulmonary response associated with human exposure to nanoparticles.

4.1 Introduction

Nanoparticles (NPs) are increasingly used on a large scale in industrial materials and agriculture [44-50]. These applications pose the risk of environmental exposures

[3] This chapter comes from my first author work that is currently in review for publication. This work was co-authored by Michaela Albright and Nicholas Niemuth, along with professors Rob Tighe and Christine Payne.

either directly, as in during manufacturing and agricultural uses, or following the degradation of nanoparticle-containing materials [51-54]. A major concern is inhalation of NPs, especially titanium dioxide (TiO_2) NPs,[57, 209-212] which are used at very large scale. TiO_2 NPs are widely used in the paint, plastics, rubber, adhesives, coatings, and paper industries, and are also a common food coloring [57-59]. The Department of the Interior reported U.S. production of TiO_2 at 1.1 million metric tons, valued at $3.2 billion, in 2022 [213].

In any biological system, proteins adsorb on the surface of NPs forming a protein "corona" [23, 51, 94, 110, 214-217]. The specific proteins that adsorb on the NP surface determine the subsequent interactions of the NP with cells and organs including cellular internalization, immune response, biodistribution, and circulation time [23, 176, 216-219]. The majority of research on the protein corona, including our own,[60-62] has focused on blood serum proteins relevant to nanomedicines. In comparison, inhalation of NPs brings the NPs into contact with the more complex environment of lung lining fluid. The lung has a thin layer of fluid that lines the airspaces. This fluid is a mixture of proteins, lipids, and mucins. The proteins present in lung fluid include those designed for antimicrobial and oxidant defense. Therefore, this lung fluid has important functions in the innate defense mechanisms of the lung [63-65].

Our goal was to characterize the concentration and composition of the protein corona formed by bronchoalveolar lung fluid (BALF) and then determine the cellular response to the lung fluid proteins adsorbed on TiO_2 NPs. In comparison, previous work has examined a single lung fluid protein, lung surfactant protein A (SPA), isolated from BALF,[220] rather than considering the entire fluid mixture. The concentration of the protein corona was measured with a colorimetric assay and the protein composition was determined by gel electrophoresis and proteomics. We found that incubating TiO_2 NPs with BALF resulted in a protein corona dominated by albumin, the most abundant protein present in BALF. The production of cytokines (interleukin 6 (IL-6), tumor necrosis factor-alpha (TNF-α), and macrophage inflammatory protein 2 (MIP-2)) by macrophage cells has been associated with lung inflammation [221-223]. These cytokines were used as measures of the inflammatory response to the BALF-TiO2 NPs. Coronas formed from bovine serum albumin (BSA) and fetal bovine serum (FBS) were used for comparison. We observed that the BALF corona led to an increased production of pro-inflammatory cytokines in comparison to bare TiO_2 NPs and the other protein coronas. We hope our work will help address underlying mechanisms of the observed lung toxicity and inflammation associated with inhalation of TiO_2 NPs [224, 225].

4.2 Materials and Methods

4.2.1 Nanoparticles (NPs) and characterization

TiO_2 NPs (#718467, Sigma-Aldrich, Burlington, MA) were used for all experiments. Hydrodynamic diameter, polydispersity index, and zeta potential of the NPs (100 μg/mL in phosphate buffered saline (PBS) diluted 1:100 in ultrapure water) were measured using dynamic light scattering (DLS; Zetasizer, Malvern Instruments, Worcestershire, England). Measurements were carried out in triplicate with three distinct samples. Average and standard deviation are reported for all measurements. Electrophoretic mobility was converted to zeta potential using the Smoluchowski approximation.

4.2.2 Rodent bronchoalveolar lavage (BAL)

C57BL/6 male mice (8-10 weeks) were purchased from Jackson Laboratories (Bar Harbor, ME). All procedures were approved by the Duke University Institutional Animal Care and Use Committee (IACUC) and were performed under an IACUC approved animal protocol (A053-21-03). BAL was performed following a published protocol.[226] Prior to BAL, mice were deeply anesthetized with an intraperitoneal injection of ketamine (100 mg/kg), xylazine (100 mg/kg), and saline (0.9%), dosed by weight (350-500 μL). The chest and trachea were dissected to expose the lungs and trachea. Following a nick in the trachea, PE-60 tubing (#9565S30, Thomas Scientific, Swedesboro, NJ) was inserted into the trachea and attached to a 12-inch infusion set

(#SV-25BLK, Terumo, Tokyo, Japan), which was connected to a 10 mL syringe held on a ring stand. Lungs were passively filled to 20 cm H_2O with PBS to reach total lung capacity. The BALF was then passively drained. The BALF was placed on ice for further processing and for use in NP incubation experiments. BALF used for protein corona formation was pooled to reduce mouse-to-mouse variation.

4.2.3 Protein corona formation and quantification

A protein corona was formed by incubating TiO_2 NPs (1 mg/mL) in 10% solutions of FBS (#10437028, Thermo Fisher Scientific, Waltham, MA), BSA (#A2153, Sigma-Aldrich) or BALF diluted in PBS. Initial incubation protein concentrations for FBS, BSA, and BALF were 6, 1.6, and 0.017 mg/mL respectively. The incubations were performed at room temperature on an orbital shaker for 30 minutes. To remove unbound proteins, the NPs were "washed" with PBS. Each wash step consisted of a centrifugation cycle (18,000 rcf, 15 mins), removal of the supernatant, and then resuspension in an equal volume of PBS. The hard corona is defined as the protein that remains bound to the NPs with no unbound protein detected in the supernatant.

Protein concentration was measured with the bicinchoninic acid assay (BCA assay; #2260, Thermo Fisher Scientific) according to the manufacturer's instruction. In brief, washed protein corona samples were resuspended in 50 μL of PBS and sonicated

(5 min, room temperature). Protein concentration was then determined by following the BCA protocol with the addition of a centrifugation cycle (15 min, 18,000 rcf, 4 °C) to remove NPs from the suspension. The NP concentration after the washing process was measured by absorption at 440 nm and comparison with a calibration curve of known concentrations. Protein concentration is reported as protein relative to NP concentration (μg protein/mg NPs).

4.2.4 Gel electrophoresis

Gel electrophoresis was used to visualize individual proteins present in the corona of the NPs, as previously described [60]. The protein coronas were removed from the NPs by suspending the NPs in loading buffer (Laemmli, #BP-110R; Boston BioProducts, Ashland, MA), incubating for at least 5 min at 95°C, and then loaded onto a gel (tris-glycine sodium dodecyl sulfate (SDS) gel, #4561093, Bio-Rad, Hercules, CA) for SDS-polyacrylamide gel electrophoresis (PAGE; 230 V, 35 min). A 10 to 250 kDa molecular weight marker (Precision Plus Protein Dual Color Standards, #1610374, Bio-Rad) was included. Gels were rinsed by microwaving in deionized water (1 min heat, 1 min rocking at RT, replace water, x3), stained (SimplyBlue Safe Stain, #LC6060, Thermo Fisher) by microwaving until near boiling (1 min), and then rocked for 5 min. Gels were destained in deionized water (10 min, rocking) and NaCl solution (20% w/v, >5 minutes, rocking) and then imaged (PhotoDoc-It, Analytik Jena, Jena, Germany).

4.2.5 Proteomic analysis

Prior to proteomic analysis, samples were digested using a modified S-Trap micro column (Protifi, Farmingdale, NY) protocol, as previously described [60]. Proteins were removed from the NP surface by resuspending in 5% SDS buffer, sonicating to reduce aggregates followed by incubating at 95 °C for 15 minutes. Protein concentration was measured with the Pierce 660 nm Protein Assay Reagent (#2260, Thermo Fisher Scientific) with the addition of Ionic Detergent Compatibility Reagent (#22663, Thermo Fisher Scientific) according to the manufacturer's instructions. Samples were pooled to load a minimum of 5 μg of protein on each S-Trap. Two modifications were made to the S-trap protocol: dithiothreitol (DTT; #R0861, Thermo Fisher Scientific) and iodoacetamide (IAM; #I1149, Sigma-Aldrich) were used as the reducer (20 mM) and alkylator (40 mM), respectively. DTT and IAM are commonly used for proteomics and are recommended substitutions. Following the completion of the S-trap protocol, the resulting digested proteins were lyophilized and stored at -20 °C until proteomic analysis.

Proteomic analysis was carried out in the Proteomics and Metabolomics Core Facility, part of the Duke Center for Genomics and Computational Biology, as previously described [60, 61]. In brief, digested samples were analyzed using LC-MS/MS

with ≤ 5 μg of digested protein injected. NanoFlow LC was performed with an ultra-performance liquid chromatography (UPLC, 75 μm x 250 mm, nanoAcquity, Waters Corporation; 400 nL/min) column and a 60 minute total elution time. The column was run with an acetonitrile gradient (5-40%) with 0.1% formic acid. Peptide fragments were analyzed using in-line tandem mass spectrometry (Orbitrap Fusion Lumos, Thermo Fisher).

We analyzed the LC-MS/MS data using MaxQuant (v2.1.0, Max Plank Institute, Munich, Germany), an open-source software designed to analyze mass spectrometry data qualitatively and quantitatively [227, 228]. The raw LC-MS/MS spectra were searched, using their integrated Andromeda search engine, against the Swiss-Prot murine canonical protein database from UniProt, accessed on June 6th, 2022. A custom contaminants file was used while searching, which contained a relevant subset of the Common Repository of Adventitious Proteins (cRAP) database [229]. For protein and peptide quantification and identification, default MaxQuant parameters were used including a 0.01 false discovery rate, a minimum peptide length of 7 amino acids, a maximum peptide length of 25 amino acids, oxidation and acetyl groups as variable modifications, and carbamidomethyl as a fixed modification. The proteins were quantified by their summed intensity.

Proteomic data was analyzed and filtered in Perseus (v2.0.3.1, Max Plank Institute) [230]. Proteins were excluded if they were considered contaminants, quality control standards, or were not observed in at least 2 samples. After filtering in Perseus, 112 proteins were observed across the samples. To correct for any change in performance or differences in protein loading, each sample was normalized to itself by dividing by the mean of the interior 80% of the protein intensities. Each sample was scaled to have the same average. We report these values as percent normalized abundance. Fold change for each protein was calculated by taking the log base 2 of the normalized abundance of samples divided by BALF.

The complete lists of proteins are included in the Electronic Supplementary Information of the original publication. In addition, the mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the Proteomics Identification Database (PRIDE) partner repository with the dataset identifier PXD041036 and DOI 10.6019/PXD041036 [231].

4.2.6 Cell culture, TiO_2 NP incubations, and cytokine assays

RAW 264.7 mouse macrophage cells (TIB-71, ATCC, Manassas, VA) were cultured in Dulbecco's Modified Eagle Medium (DMEM, #12100046, Thermo Fisher Scientific) supplemented with 10% FBS (#F4135, Sigma-Aldrich) at 37 °C and 5% CO_2.

Cells were passaged by scraping (#08100240, Thermo Fisher Scientific) every 4-5 days. Cells were seeded at 250,000 cells/mL in 6-well plates (#353046, Corning, Corning, NY) for gene expression experiments. Cells were seeded and grown overnight in DMEM with 10% FBS, and this media was removed and replaced with DMEM without FBS immediately prior to addition of NPs. Cells were incubated with TiO_2 NPs (250 μg/mL) or a PBS vehicle control in serum-free media for 24 h.

Following NP or PBS exposure, cells were processed to extract RNA using the RNeasy Plus Micro Kit (#74034, Qiagen, Hilden Germany). After RNA extraction, RNA content was quantified by Nano Drop (Thermo Fisher Scientific). cDNA H Minus Synthesis Master Mix (#M1681, Thermo Fisher Scientific) was used to generate cDNA. Power Up SYBR Green Master Mix (#A25743, Thermo Fisher Scientific) was used for RT-PCR and reactions were run in QuantStudio 6 (Applied Biosystems, Waltham, MA). The following primer sequences were used: IL-6 Forward TTGGTCCTTAGCCACTCCTTC, IL-6 Reverse TAGTCCTTCCTACCCCAATTTCC; TNF-α Forward CTATGTCTCAGCCTCTTCTC, TNF-α Reverse CATTTGGGAACTTCTCATCC; Cxcl2(MIP-2) Forward GGGTTGACTTCAAGAACATC, CXCL2(MIP-2) Reverse CCTTGCCTTTGTTCAGTATC; 18s Forward TTGACGGAAGGGCACCACCAG, 18s Reverse GCACCACCACCCACGGAATCG. Data collected by QuantStudio 6 was analyzed in Prism (v. 9.5.1, GraphPad Software, San Diego, CA). Cycle threshold (CT)

values were normalized to housekeeping gene (18S). Data was expressed as a fold change compared to the control treated groups.

4.3 Results and Discussion

4.3.1 NP characterization

The TiO_2 NPs used in this study have a primary diameter of ~21 nm, but are observed as fused aggregates by electron microscopy and DLS [198, 232]. The hydrodynamic diameter and zeta potential were measured without sonication (Table 4-1), to avoid possible disruption of the protein corona.

4.3.2 BALF forms a protein corona on the surface of TiO_2 NPs

Protein coronas were formed by incubating TiO_2 NPs for 30 minutes at room temperature with three different protein sources: FBS (6 mg/mL, equivalent to a 10% v/v FBS solution), BSA (1.6 mg/mL), and BALF (0.017 mg/mL). FBS (10%) is a common nutrient source for cell lines, making it relevant to *in vitro* studies. Murine BALF obtained from mice by lavage was used to model the protein corona formed following inhalation. Albumin is the most abundant protein in both FBS and BALF. The concentration of BSA was chosen to be comparable to the amount of BSA that is found in FBS [60, 233]. Unbound proteins were removed from the protein-TiO_2 NP suspensions

by centrifugation (18,000 rcf, 15 minutes) and resuspension (3x), as described previously [198, 232, 234]. Removal of free protein was confirmed by BCA (Figure 4-5).

Hydrodynamic diameter and zeta potential of the protein-TiO_2 NP complexes were measured using DLS (Table 4-1). There was no significant change in the diameter of the NPs with the addition of the corona. Previous work examined the coronas formed from porcine BALF on 8 different metal oxide NPs and found that BALF did not lead to a disruption of the TiO_2 NP aggregates [220]. The zeta potential increased significantly from the bare NPs (-35 ± 3 mV) to the NPs with FBS corona (-23 ± 2 mV; $p < 0.05$), in agreement with previous studies by our lab in which a FBS corona on TiO_2 NPs was observed to have a zeta potential of (-24 ± 2) [232, 234]. In comparison, there was no significant change in zeta potential with the formation of a BSA (-32 ± 2 mV) or BALF (-38 ± 2 mV) corona. The differences in zeta potential are correlated with the protein corona concentration, as shown using BSA as a representative protein (Figure 26). In addition, previous work has shown the change in zeta potential is related to the initial charge of the NP and the specific proteins that form the corona [235].

Table 4-1: TiO_2 NP hydrodynamic diameter (d_h), polydispersity index (PDI), and zeta potential (ZP) in the absence and presence of a protein corona.

NP and corona	d_h (nm)	PDI	ZP (mV)	ΔZP
TiO_2	900 ± 410	0.40 ± 0.14	-43 ± 4	-

FBS-TiO_2	900 ± 120	0.41 ± 0.01	-23 ± 2	+20
BSA-TiO_2	800 ± 40	0.50 ± 0.06	-32 ± 2	+11
BALF-TiO_2	1200 ± 150	0.26 ± 0.08	-38 ± 3	+5

The concentration of protein present in the corona is measured by BCA analysis and reported as protein (μg)/NPs (mg), as described in the Materials and Methods (Fig. 1A). The differences in protein corona concentrations (FBS = 53 ± 1.3 μg/mg NP; BSA = 20 ± 1 μg/mg NP, BALF = 1.9 ± 0.3 μg/mg NP) is, at least in part, due to the concentration of the initial protein solution. Using BSA as a representative protein, we found that concentration of the protein corona scaled with the initial concentration of protein used to form the corona (Fig. 4-7), in agreement with previous work that showed a correlation between protein corona concentration and initial protein concentration [60, 236-238]. In addition, previous work with porcine BALF showed relatively low concentrations of BALF proteins on similar TiO_2 NPs [220].

4.3.3 Composition of the BALF corona

The composition of each protein corona was analyzed using gel electrophoresis (Fig. 4-1B) and proteomics (Table 4-1). Gel electrophoresis was in agreement with the protein corona concentration assay (Fig. 4-1A), with less protein present in the BSA and BALF coronas compared to FBS (Fig. 4-1B). The gel also shows that the most abundant protein in each of the coronas is albumin (66 kDa). Bands in the BALF corona at ~13 kDa

and ~8 kDa were tentatively assigned to uteroglobin (10 kDa) and the monomer of pulmonary surfactant-associated protein B (SPB; 8.7 kDa) [239].

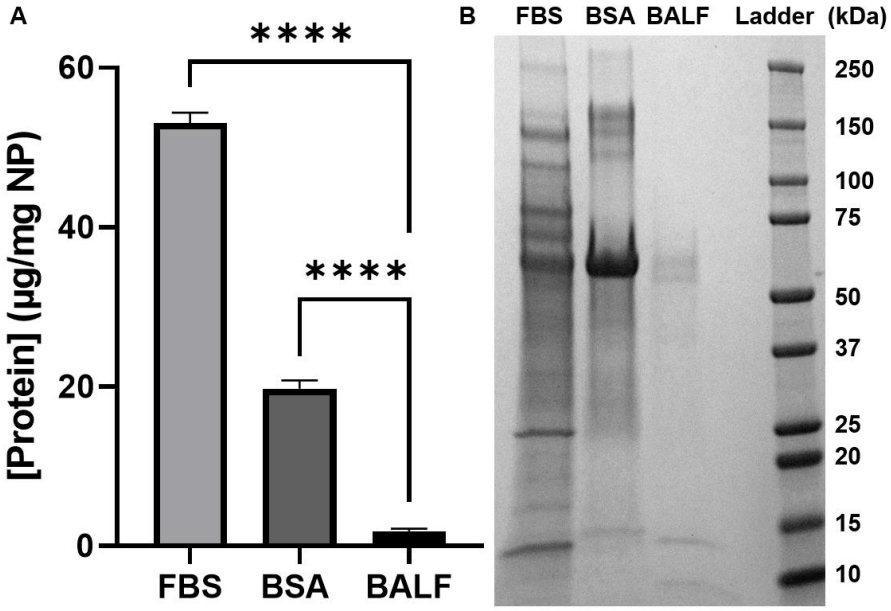

Figure 4-1: Concentration and composition of proteins adsorbed on TiO_2 NPs when incubated with FBS, BSA, and BALF. (A) Protein concentration per NP (µg/mg) when (FBS 6 mg/mL; BSA 1.6 mg/mL; BALF 0.017 mg/mL; n=3). The mean protein per NP (µg/mg) is reported with error bars showing standard deviation. Significance was calculated using an ANOVA with a post-hoc Tukey test. ****p<0.0001. (B) Gel electrophoresis of each protein corona.

Proteomic analysis of the BALF corona is shown as the relative amount of protein detected (normalized abundance; Table 4-2) and the amount of protein in the corona relative to the amount in BALF (enrichment, Fig. 4-2). In general, the normalized abundance shows that the top 15 proteins in a solution of BALF are also present in the corona formed from BALF. For example, albumin is the most abundant protein in the corona formed from BALF (78 ± 2.6%) and in BALF (88.4 ± 4.8%). In comparison, the

abundance of uteroglobin in the corona suggests selective adsorption on the TiO_2 NP surface. Uteroglobin is the 8th most abundant protein in the corona (0.6 ± 0.3%), while it is the 24th most abundant protein in BALF (0.04 ± 0.07%). Uteroglobin, also known as blastokinin and club-cell secretory protein (CCSP), is an immunomodulatory, anti-inflammatory, and anti-chemotaxis secretoglobin expressed by epithelial cells that interact with external environments.[240]

Table 4-2: Normalized abundance (%) of the top 15 most abundant corona proteins (n=3). For comparison, the rank order of proteins present in BALF is shown in parentheses. Mean and standard deviation are reported.

Protein	Protein Corona (%)	BALF (%)
Albumin	78 ± 2.6	88.4 ± 4.8 (1)
Serotransferrin	6 ± 3.2	3.4 ± 1.8 (2)
Hemoglobin subunit beta-1	4.2 ± 0.7	2.2 ± 1.7 (3)
Alpha-1-antitrypsin 1-1	2 ± 1.2	0.9 ± 0.5 (5)
Hemoglobin subunit alpha	2 ± 1.2	0.4 ± 0.6 (8)
Serine protease inhibitor A3K	1.0 ± 0.6	1.1 ± 0.6 (4)
Lysozyme C-2	0.7 ± 0.6	0.7 ± 0.4 (7)
Uteroglobin	0.6 ± 0.3	0.04 ± 0.07 (24)
Haptoglobin	0.6 ± 0.3	0.1 ± 0.1 (12)
Transthyretin	0.6 ± 0.2	0.08 ± 0.09 (14)
Alpha-1-antitrypsin 1-5	0.5 ± 0.8	0.4 ± 0.2 (10)
Hemopexin	0.5 ± 0.1	0.2 ± 0.1 (11)

Chitinase-like protein 3	0.47 ± 0.05	0.7 ± 0.6 (6)
Alpha-1-antitrypsin 1-4	0.4 ± 0.1	0.4 ± 0.2 (9)
Peroxiredoxin-6	0.2 ± 0.2	0.1 ± 0.1 (15)

Enrichment and depletion of a protein in the corona compared to the solution used to form the corona can be visualized in an enrichment plot (Fig. 4-2). In the enrichment plot, the fold change from serum to corona is displayed as $\log_2$ so that an enrichment of 0 is no change, while an enrichment of 1 is a 2-fold enrichment on the corona. Likewise, an enrichment of -1 is a 2-fold depletion in abundance on the corona. This type of plot can be used to determine which proteins are outliers suggesting they are enriched or depleted in the protein corona. For example, albumin is slightly depleted (-0.18 ± 0.05) in the corona relative to BALF. Of the top 25 most abundant proteins in the protein corona, only SPB is an outlier as determined by the robust regression and outlier removal (ROUT) method (Q = 0.1%). SPB is the most enriched protein within the top 25 proteins suggesting that it has high affinity for the TiO_2 NP surface relative to other proteins in BALF. SPB is a hydrophobic apolipoprotein that is essential for lung function [241]. The enrichment of SPB is interesting considering molecular dynamics studies that showed anionic NPs were selective for SPB, compared to the similar lung surfactant protein C (SPC) due to a combination of hydrophobics and electrostatics [242]. Although not an outlier, the next most highly enriched protein is uteroglobin (3.8 ± 0.8, corona abundance rank 8). Overall, this enrichment and depletion profile (Fig. 2) shows that

there is a selection of specific proteins in the NP corona. Several of these proteins (serotransferrin, uteroglobin, haptoglobin, hemopexin, and chitinase-like protein 3) have immunoregulatory properties suggesting that these proteins adsorbed on the surface of the TiO_2 NPs could regulate the inflammation associated with TiO_2 lung toxicity *in vivo* [243-246].

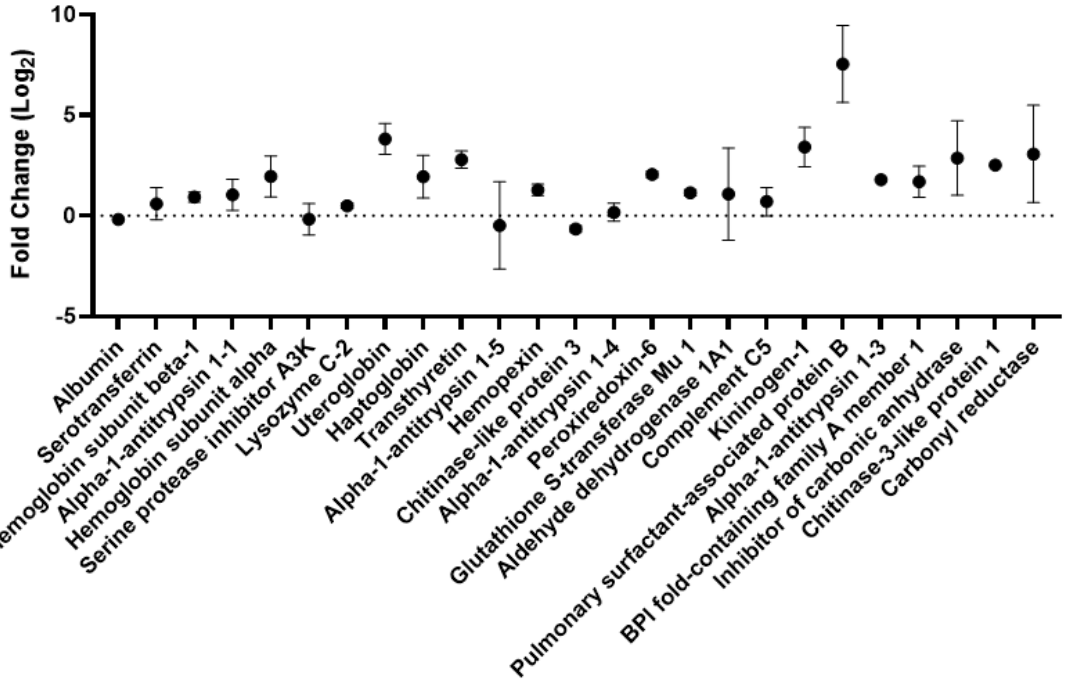

Figure 4-2: Enrichment (positive fold change) and depletion (negative fold change) of the 25 most abundant proteins present in the BALF protein corona present on TiO2 NPs relative to their abundance in BALF. Fold change, log base 2, with error bars showing standard deviation is plotted for each protein (n=3, n=2 for lysozyme C-2 and peroxiredoxin-6 with outliers removed). Proteins are listed in order of their relative abundance (Table 4-2).

4.3.4 BALF corona leads to increased expression of inflammatory cytokines in macrophage cells

To first determine if bare TiO2 NPs elicit an inflammatory response in macrophage cells, cells were incubated with TiO2 NPs for 24 hours. To prevent the formation of a protein corona in situ, serum-free media was used for the incubation. Following the incubation with TiO2 NPs, pro-inflammatory cytokine gene expression (IL-6, TNF-α, and MIP-2) was measured by real time PCR. We observed that bare TiO2 NPs significantly increased the expression of these pro-inflammatory cytokines (Fig. 3). This observation is in agreement with previous work showing TiO2 NPs increased pro-inflammatory cytokines associated with lung inflammation [210, 247, 248].

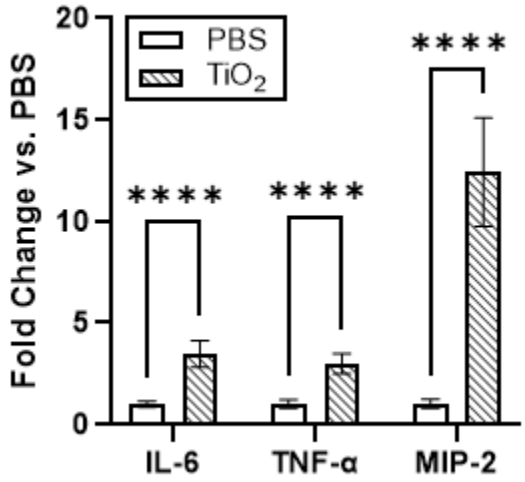

Figure 4-3: Expression of the pro-inflammatory genes IL-6, TNF-α, and MIP-2 increased in response to bare TiO2 NPs compared to untreated control cells (n=6). Expression change was found to be significant for each gene using unpaired t-tests. **** p<0.0001.

To determine if a BALF corona altered the cellular response to the TiO2 NPs, cells were incubated with BALF-TiO2 NPs and cytokine expression was compared to the response to bare TiO2 NPs (Fig. 4-3). FBS and BSA coronas were used for comparison as the predominant protein in BALF is albumin. The formation of the protein coronas is described in Materials and Methods. We observed that exposure of macrophage cells with BALF-TiO2 NPs increased the expression of pro-inflammatory cytokines (IL-6, TNF-α and MIP-2; Fig. 4-4). In comparison, coronas formed from FBS and BSA did not alter the expression of pro-inflammatory cytokines when compared to bare TiO2. These results suggest that a BALF corona is unique in increasing macrophage expression of pro-inflammatory cytokines and that proteins enriched in the BALF corona (Table 4-2 and Fig. 4-2) may drive the enhanced toxicity of these NPs.

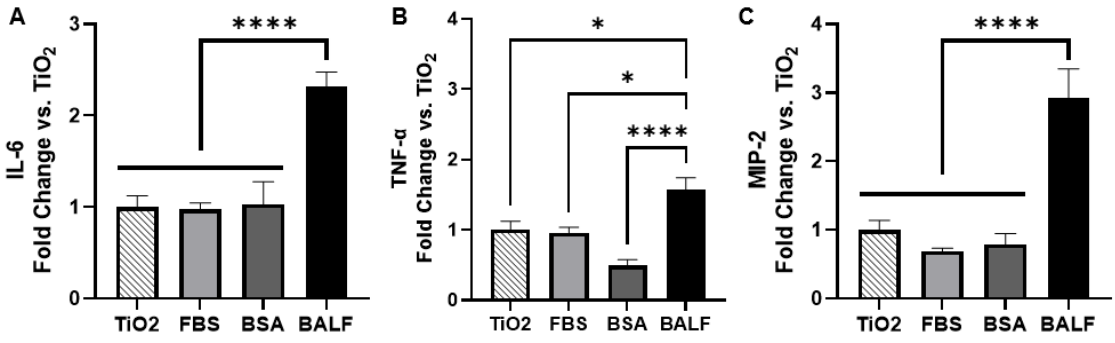

Figure 4-4: IL-6, TNF-α, and MIP-2 showed elevated expression levels in response to BALF-TiO2 NP compared to FBS- and BSA-TiO2 NPs (n=6). Cytokine expression is reported as fold change relative to TiO2. Significance was determined using a one-way ANOVA with a post hoc Tukey test. Comparisons to the PBS control were also performed (Fig 4-8). *p<0.05, ****p<0.0001.

4.4 Conclusion

Our studies provide the first description of the interaction of TiO_2 NPs with BALF and the subsequent response of macrophage cells. We propose that these interactions are critical initial steps in the pulmonary response following the inhalation of these NPs. We find that the proteins present in BALF form a corona on the surface of the TiO_2 NPs (Fig. 4-1). The composition of this corona (Table 4-1 and Fig. 2-2) results in a distinct protein-TiO_2 NP complex that is associated with an elevated cytokine response in macrophage cells. The incubation of macrophage cells with bare TiO_2 NPs leads to increased expression of pro-inflammatory genes, IL-6, TNF-α, and MIP-2 (Fig. 4-3). A BALF corona further enhances this pro-inflammatory response (Fig. 4-4). In comparison to FBS and BSA coronas, the BALF corona leads to a much greater response for all three cytokines. It is possible that BALF-specific proteins, other than albumin, are responsible for this increase in cytokine expression either through protein-dependent responses, an increased uptake of NPs, or a combination of both. For example, previous work has measured the uptake of polystyrene NPs (50 nm and 100 nm) by human alveolar lung cells (Type I and II) [249]. Pre-incubation of these polystyrene NPs with BALF led to increased cellular uptake by Type I cells. Type II cells did not internalize the polystyrene NPs. SPA and surfactant protein D (SPD) were identified on the surface of these polystyrene NPs by western blot. It is not known if SPB was probed. Similarly, work

with magnetite NPs (110 nm - 180 nm) showed that adsorption of SPA on the surface of the NPs led to increased uptake by alveolar macrophage cells [250]. Other surfactant proteins were not examined. In comparison, BSA led to decreased uptake.

In addition to characterizing the BALF corona formed on TiO_2 NPs and the resulting cytokine response, we hope these experiments will provide a starting point for future *in vivo* studies examining the mechanism of the toxicity associated with the inhalation of TiO_2 NPs [251, 252]. For example, if this enrichment of SPB on the surface of the TiO_2 NPs leads to a corresponding depletion of free SPB in the lung there could be a reduced surface tension in the lungs associated with respiratory distress [253]. It is also possible that protein coronas could sequester specific lung lining fluid proteins required for normal host-defense and maintenance of lung homeostasis or the interaction of proteins with the NP surface could activate proteins to make them more pro-inflammatory. These questions will need to be addressed in future studies to define specific effects of the protein corona on the functions of these proteins. One potential limitation of our studies is that our current work has not addressed the lipid corona that is expected to form following incubation of TiO_2 NPs with BALF [254-257]. Previous molecular dynamics simulations showed that adsorption of specific proteins (SPA, B, and C) on silver and polystyrene NPs (5 nm and 15 nm, anionic) was determined by the hydrophobicity of the NPs. The lipid composition was insensitive to the type of NP

[258]. We hope future work will probe the lipid composition on these much larger TiO_2 NPs.

4.5 Supplementary Information

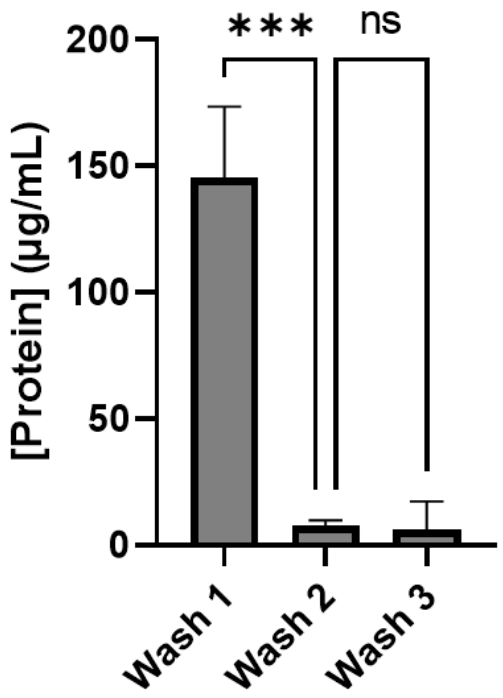

Figure 4-5: Concentration of protein present in the supernatant following "washing" of protein-TiO2 NP complexes incubated with BSA (8 mg/mL). TiO2 NPs were incubated with BSA for 30 min and washed with PBS (n=3). The first supernatant, wash 0, is not shown because it has much higher protein concentration and is not typically measured. BSA (8 mg/mL) is used as an example as it is the highest incubation concentration used for protein corona quantification. Error bars show standard deviation. Significance was determined using a one-way ANOVA with a post hoc Tukey test. *** p<0.001.

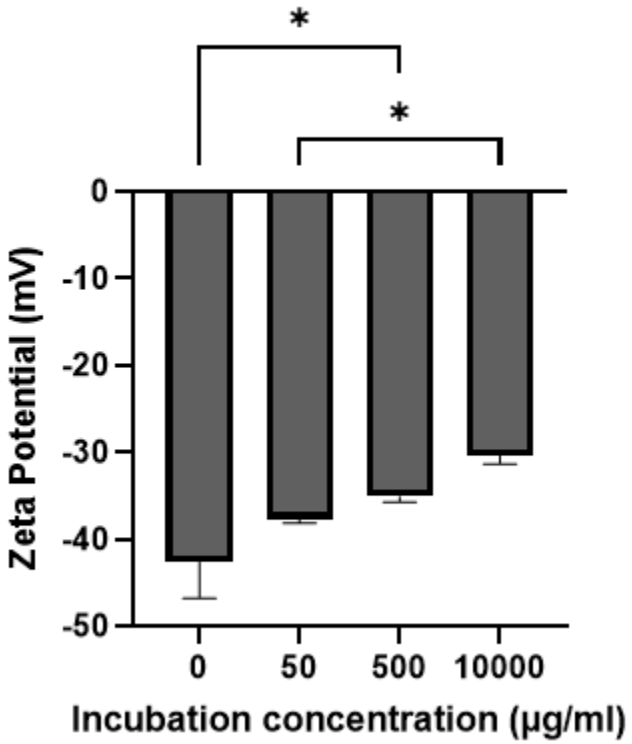

Figure 4-6: Increasing the amount of BSA present during the formation of the protein corona increases the zeta potential of the protein-TiO2 NP complexes. Experiments were carried out in triplicate. Error bars show standard deviation. Significance was determined using a one-way ANOVA with a post hoc Tukey test. *p<0.05.

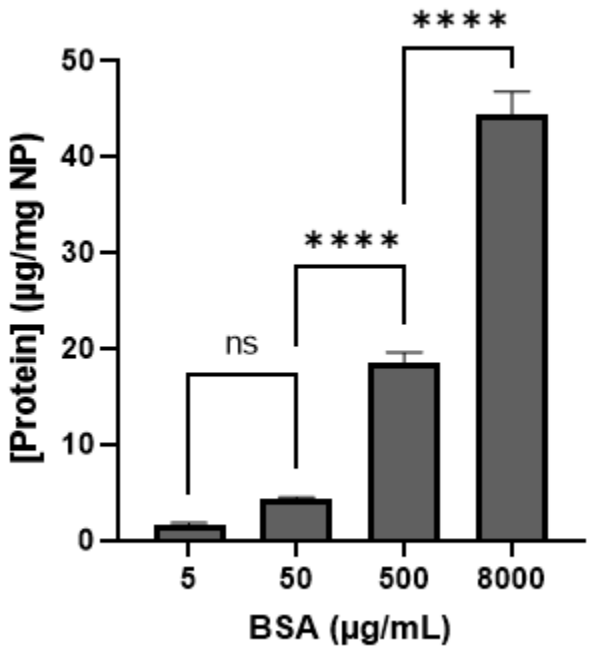

Figure 4-7: Increasing the amount of BSA (5 µg/mL – 8000 µg/mL) present during the formation of the protein corona increases the concentration of protein present in the corona formed on TiO2 NPs. Experiments were carried out in triplicate. Error bars show standard deviation. Significance was determined using a one-way ANOVA with a post hoc Tukey test. ****p<0.0001.

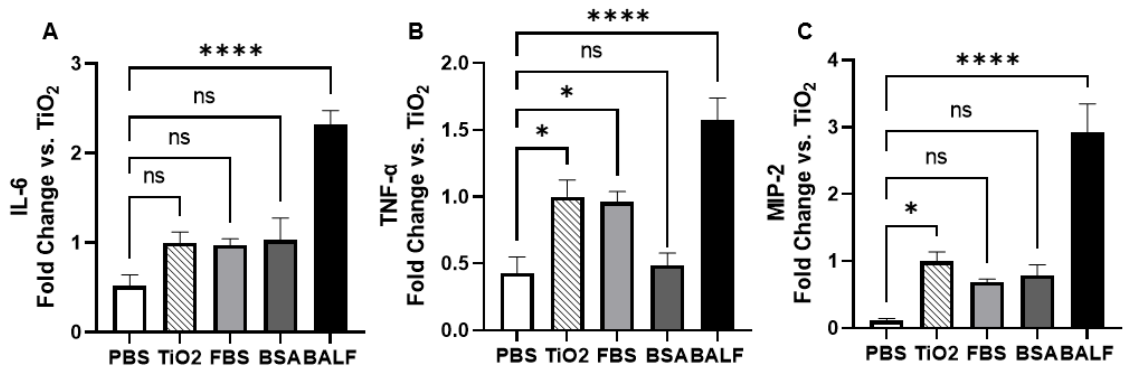

Figure 4-8: IL-6, TNF-α, and MIP-2 showed elevated expression levels in response to TiO2 NPs with and without proteins coronas compared to the use of PBS as vehicle control. Significance was determined using a one-way ANOVA with a post hoc Tukey test. This is the same dataset and one way ANOVA as shown in Fig. 4. *p<0.05, ****p<0.0001.

5. Conclusion

5.1 Summary

The overarching goal of our work is to understand what drives nanoparticle protein interactions, with the idea that leveraging this understanding will allow us to accurately predict the protein corona formation on any nanoparticle in any biofluid. The work presented here has furthered this goal by developing and validating semi-automated methodologies for high-throughput reproducible experimentation, building an understanding of how serum concentration and incubation time effect the protein corona, synthesizing a library of magnetic nanoparticles, and collating a dataset of nanoparticle protein coronas and protein properties for machine learning.

In developing and validating an automated methodology for generating, purifying, and characterizing coronas we have built the framework for building larger datasets. We also validated low-cost proteomics and the reproducibility of our methods with triplicate samples showing agreement between runs. We then used these methods to determine that increasing incubation concentration increases corona concentration but has only a slight effect on the corona composition. We also observed that incubation time does not significantly change the composition of the corona but does increase the corona concentration at higher incubation concentrations. These results suggest that higher affinity proteins are not slowly displacing higher abundant proteins, in other

words the protein corona formation is relatively static, on the time scale of a day, after it forms for our system.

We have used these automated methods and applied them to making a larger high throughput dataset for understanding and predicting protein corona formations. This work for characterizing and generating the dataset has been laid out in the appendix. The machine learning results on these datasets show the capability to predict protein coronas across proteomes and nanoparticle surfaces. While doing this, we have seen that it is easier to predict new nanoparticle surfaces rather than new protein coronas (new meaning not within the model training dataset).

To put the importance of the protein corona in perspective we determined that the BALF corona elicits an increased inflammation response relative to other coronas and the bare NPs. In chapter 4, we characterized the BALF corona for the first time as a model system for the protein corona that forms during inhalation exposures. It formed a unique corona with specific enrichment of lung proteins including pulmonary surfactant associated protein B and uteroglobin. These proteins have previously been enriched on other nanoparticles and require follow up experiments to determine if they play a significant role in the increased inflammation response that was observed. These experiments also laid the foundation for doing further in vivo experiments to understand the potential toxicity of NP exposures.

5.2 Future Work

With automated methods for generating, purifying, and characterizing protein coronas there are many potential directions for next steps. There is significant interest in having a generalizable model for predicting the formation of NP protein coronas. Moving forward, this model could be used to determine material properties of interest that could then be synthesized and tested using the automated workflow. One of the most important validation tests is to see how modulating the important features, as determined by the machine learning model, changes protein affinity for NP surfaces. For instance, if we find that the fraction of alanine amino acids in a protein is important for the model to predict protein adsorption, we can than find proteins or peptides with varying concentrations of alanine and see if there are any differences in affinity between an NP surface and the proteins. This sort of experiment could be done using methods like isothermal titration calorimetry, differential scanning fluorimetry or fluorescent microscopy. Once the model has been validated there a multitude of interesting areas to pursue but one of the most important directions from an applications standpoint is determining what material characteristics are necessary for enriching specific proteins either as a way to prevent immune response, inflammation response, or as a method for enriching specific biomarkers in disease models.

A further out next step is to build an understanding of how the other biomolecules in the serum and blood interact with nanoparticle surfaces. We understand

that there are a variety of lipids, carbohydrates, and salts that are also likely to interact with NP surfaces and may complicate our understanding of how the protein corona interacts with cells. There are many exciting questions to solve and interesting directions to go with the methods and models that have been generated here and I look forward to seeing the next breakthroughs in this field.

A. Appendix A – Predicting the Protein Corona

In addition to the published results, we have been working on predicting the formation of the protein corona on varied nanoparticle surfaces and across proteomes through the use of our high throughput workflow.

A.1 Materials and Methods

A.1.1 Preparation of Iron Oxide Nanoparticle Cores

The iron oxide NPs were synthesized using previously published protocols [12, 259]. In brief 40 mL of ethylene glycol (#324558, Sigma Aldrich, St.Louis, MO), 1.3 g $FeCl_3*6H_2O$ (#236489, Sigma Aldrich), 0.52 grams of trisodium citrate (#S4641, Sigma Aldrich), and 2.4 grams of sodium acetate (#S2889, Sigma Aldrich) were mixed together in an Erlenmeyer flask. At this point DI water can be added to vary the size of the NPs. Two sizes were chosen at 0 mL and 4 mL of water added as denoted by large and small in the text. The Erlenmeyer flask was covered and stirred for at least one hour or until a red-brown solution was obtained. This mixture was added to a Teflon lined stainless steel reaction flask and heated to 200°C for 10 hours. The reaction flask was allowed to cool to room temperature. After contents were appropriately cooled the resulting black NPs were washed at least 3 times with ethanol. A strong magnet was used to remove the NPs from suspension. The resulting nanoparticles were suspended in a minimal amount of ethanol and dried under a stream of nitrogen overnight.

A.1.2 Iron oxide nanoparticle functionalization

The resulting dried NPs were functionalized according to previously published protocols by Nguyen et al [12]. NPs were successively functionalized to generate gold shell iron oxide core nanoparticles. To obtain a positive surface charge, NPs (5 mg/ml) were mixed with polyethylenimine (PEI 0.1mM (aq), #408727, Sigma Aldrich) and shaken for at least 1 hour on a rotary shaker at room temperature. Following PEI functionalization, the NPs were functionalized with Gold nanoseeds by adding 1 mg of PEI-NPs to 50 mL of Gold nanoseed solution. To grow the Gold shell polyvinylpyrrolidone (PVP 9.85 mg/ml, #PVP40, Sigma Aldrich), NPs functionalized with Gold nanoseeds (25 μg/mL), hydroxylamine (75 μg/ml, #159417, Sigma Aldrich), and Gold (III) chloride trihydrate (75 μg/ml, #520918, Sigma Aldrich) were successively combined in DI water. The color of the solution took on a bluish-purple tint within minutes of adding the gold mixture. PVP was exchanged for PEI by shaking the PVP@Au NPs (1 mg/ml) in PEI 0.1 mM (aq) for at least 1 hour. Alternatively, the PVP was displaced with thiolated polyethylene glycol (PEG) by shaking the PVP@Au NPs in a thiol PEG solution (10 mM, A3021/MH-SH5000, JenKem USA, Plano, TX) for at least 1 hour. After each functionalization step the NPs were removed from suspension using a magnet and washed at least 4 times with DI water.

A.1.3 Gold nanoseed synthesis

Gold nanoseeds were synthesized following a previously published protocol [12, 260]. In brief an Erlenmeyer flask was charged with 44 mL of DI water, 3 mL of 100 mM sodium hydroxide (aq), and 1 mL of 50 mM tetrakis (hydroxymethyl) phosphonium chloride (aq). After mixing for 5 minutes with a magnetic stir bar, 1 mL of 25 mM gold (III) chloride trihydrate was added. After addition of the gold salt the solution turned a deep red signifying the formation of gold NPs.

A.1.4 PEG functionalization of polystyrene nanoparticles

Commercially available polystyrene (PS) NPs (200 nm, Carboxylate modified, ThermoFisher Scientific) were conjugated with PEG using EDC conjugation protocol. Stock PS NPs were prewashed, by diluting 10 fold with DI water and centrifuging. In a standard conjugation, 100 μl of 4 mg/mL washed PS NPs were added to 200 μL methoxy-PEG-Amine 2K (50mg/mL, #A3071, JenKemUSA) in 4-Morpholineethanesulfonic acid hemisoldium salt (MES 25 mM (aq), #M0164, Sigma Aldrich). The NP PEG mixture was vortexed and shaken for 5 minutes before adding 40 μL of N-(3-Dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride (EDC 45 mg/mL (aq), #E7750, Sigma Aldrich). The mixture was rotary shaken for at least 30 minutes before being washed 3 times with PBS via centrifugation, to remove excess PEG and EDC.

A.1.5 Nanoparticle characterization

NP diameter was measured with transmission electron microscopy (TEM) and dynamic light scattering (DLS). TEM was carried out using Tecnai G² TWIN TEM (FEI, Hillsboro, OR) at the Shared Materials Instrumentation Facility at Duke University or using the UNC characterization facilities TEM. All samples were prepared by drop casting on 400 mesh copper grids (#CF400-Cu, Electron Microscopy Sciences, Hatfield Township, PA) and drying at room temperature (RT) for 12-18 hrs. Nanoparticle diameters were measured using ImageJ [183]. Average and standard deviations are reported for all measurements.

Hydrodynamic diameter, polydispersity index, and zeta potential of the NPs (10-100 μg/mL in phosphate buffered saline (PBS) diluted 1:100 in ultrapure water) were measured using DLS (Zetasizer, Malvern Instruments, Worcestershire, England). Measurements were carried out with three distinct samples. Each measurement was performed for 12 - 30 runs. The average and standard deviation are reported for all measurements. Electrophoretic mobility was converted to zeta potential using the Smoluchowski approximation.

A.1.6 Liquid handling robot

A liquid handling robot (OT-2, Opentrons, Brooklyn, NY) with a magnetic baseplate was used to automate protein corona formation and isolation, as described previously [103, 261]. Protocol scripts were written in python using Opentrons API

v2.12. Pipette tips (300 µl, single and multi) and tip racks were purchased from Opentrons to verify compatibility and calibration. The locations of each reagent and sample were designated in the script and appropriately positioned prior to running the robot. Most experiments used a 96-well plate with three or six replicates, specified in the text. Within a row of eight wells, two wells were used for background subtraction.

A.1.7 Protein corona formation

A protein corona was formed by incubating NPs (1-5 mg/mL) in 1-100% solutions of FBS (#10437028, Thermo Fisher Scientific, Waltham, MA) diluted in PBS. The incubations were performed at RT on a microplate shaker for 30 minutes. Iron oxide samples were generated and purified in a liquid handling robot, in contrast with the polystyrene samples which were processed by hand. To remove unbound proteins, the NPs were "washed." Each wash step consisted of a magnetic pull down or centrifugation, removal of the supernatant, and then resuspension in an equal volume of PBS. The hard corona is defined as the protein that remains bound to the NPs with no unbound protein detected in the supernatant, as described below. The use of magnetic pulldown with the liquid handling robot provides a faster workflow and has been previously characterized [103, 261].

Protein concentration was measured with the Pierce 660 nm Protein Assay Reagent (referred to as a 660 nm assay; #2260, Thermo Fisher Scientific) with the

addition of Ionic Detergent Compatibility Reagent (#22663, Thermo Fisher Scientific) according to the manufacturer's instructions. The concentration of protein present in the hard corona was determined by removing the proteins from the NPs by incubating with sodium dodecyl sulfate (SDS) buffer (5% w/v, #L3771, Sigma-Aldrich, Burlington, MA) for 30 minutes at RT. Protein concentration was then determined by measuring absorbance at 660 nm using a plate reader (Spectramax iD3, Molecular Devices, San Jose, CA). A residual amount of protein is resistant to SDS removal independent of duration of SDS incubation as shown previously [261].

A.1.8 Protein digestion for proteomics

Prior to proteomic analysis, samples were digested using a modified S-Trap mini column (Protifi, Farmingdale, NY) protocol. Proteins were removed from the NP surface by incubating with SDS buffer for 30 minutes. Protein concentration was determined using the 660 nm assay. Samples were pooled to load a minimum of 25 μg of protein on each S-Trap. Two modifications were made to the S-trap protocol: dithiothreitol (DTT; #R0861, Thermo Fisher Scientific) and iodoacetamide (IAM; #I1149, Sigma-Aldrich) were used as the reducer (20 mM) and alkylator (40 mM), respectively. DTT and IAM are commonly used for proteomics and are recommended substitutions. Following the completion of the S-trap protocol the resulting digested proteins were lyophilized and stored at -20 °C until proteomic analysis.

A.1.9 Proteomic analysis

Proteomic analysis was carried out in the Proteomics and Metabolomics Core Facility, part of the Duke Center for Genomics and Computational Biology, as described previously [103, 261]. In brief, digested samples were analyzed using LC-MS/MS with ≤ 25 mg of digested protein injected. MicroFlow LC was performed with an ultra-performance liquid chromatography (UPLC, 1 mm x 100 mm, M-Class, Waters Corporation; 80 μL/min) column and a 17 minute total elution time. The column was run with an acetonitrile gradient (5-40%) with 0.1% formic acid. Peptide fragments were analyzed using in-line tandem mass spectrometry (Orbitrap Fusion Lumos, Thermo Fisher).

To further develop our low-cost workflow, we analyzed the LC-MS/MS data using MaxQuant (v2.1.0, Max Plank Institute, Munich, Germany), an open source software designed to qualitatively and quantitatively analyze mass spectrometry data [184, 185]. The raw LC-MS/MS spectra were searched, using their integrated Andromeda search engine, against the Swiss-Prot Bos Taurus canonical protein database from UniProt, accessed on April 27, 2022. A custom contaminants file was used while searching, which contained a relevant subset of the Common Repository of Adventitious Proteins (cRAP) database [186]. For protein and peptide quantification and identification, default MaxQuant parameters were used including a 0.01 false discovery rate, a minimum peptide length of 7 amino acids, a maximum peptide length of 25

amino acids, oxidation and acetyl groups as variable modifications, and carbamidomethyl as a fixed modification. The intensity method, in Maxquant, for quantification was used.

Resulting proteomic data was analyzed and filtered in Perseus (v2.0.3.1, Max Plank Institute). Proteins were excluded if they were considered contaminants, quality control standards, or only identified by site. Data normalization was performed in Python (v3.9, Python Software Foundation, Beaverton, OR). A quantitative internal standard was not used for these experiments. To correct for any change in performance or differences in protein loading, each sample was normalized to itself by dividing by the mean of the interior 80% of the protein intensities [190]. Each sample was scaled to have the same average. We report these values as percent normalized abundance. Fold change for each protein was calculated by taking the log base 2 of the normalized abundance of samples divided by serum.

A.1.10 Machine learning methods

To perform machine learning a database of protein features, nanoparticle characteristics, and experimental features were combined. Basic canonical protein information for homo sapiens (Tax ID), Bos Taurus (Tax ID), and Murine (Tax ID) were downloaded from the Swiss-Prot datasets [262]. The sequence information was used with NetSurfp3.0 to predict the exposed amino acids, secondary structures, accessible surface area, hydrophobicity, and polarity [263, 264]. NetSurfP3.0 uses a natural

language processing model to predict the protein structure and report results for each amino acid. These results were collated for each protein using modified protocols developed by Ouassil et al [101]. The Biopython Protein Analysis library [265] (v. 1.8.1) was used to calculate values based solely off of sequence data including percent amino acid composition, molecular weight, aromaticity, instability index, flexibility metrics, gravy score, and secondary structure. The BioPython Protein Analysis module functions for molecular weight, flexibility, instability index and gravy score are not able to account for proteins that have abbreviations for groups of amino acids in their sequence or that contain the amino acid selenocysteine. The sequence data was cleaned by replacing the unspecified amino acids with the most common amino acid, leucine [266] and replacing selenocysteine with the cysteine. The physicochemical properties of the proteins were merged with the proteomic protein abundance data, nanoparticle characteristics, and experimental features. NP ligand and core material were transformed into ordinal variables.

All of these features led to a dataset containing 92 features. To determine which features were important recursive feature elimination (RFE) and recursive feature elimination with cross validation (RFECV) from the scikit-learn library (v. 1.1.2) were employed. RFE was employed to reduce to 75 features and then RFECV was used to figure out the optimal number of features for the highest accuracy or R^2, for classification and regression respectively. For both dimensionality reduction methods, a step size of 1

was used with an 80/20 split and either random forest regressor or random forest classifier were used as the estimators.

A random forest regression model from scikit-learn was employed to predict the formation of the protein corona. The model was trained using squared error as the accuracy criterion and 100 trees. Random forest classification was employed to predict high affinity proteins. The random forest classification model was trained on fold change relative to the control protein abundance. For proteins that were not observed in the control, their abundance was estimated to be the minimum abundance observed in the control. Fold change threshold cutoffs are specified in the text. Model accuracy was determined using an 80/20 split using the train-test-split package from scikit-learn. Model testing was performed with 10-fold validation which is shown as mean and standard deviation.

A.2 Nanoparticle and protein corona characterization

Two sizes of magnetic nanoparticles were synthesized following previously published protocols [12]. The NPs were further functionalized with a variety of ligands to get 7 unique surface charges. For comparison to the automated formation of the NP and a core comparison polystyrene NP were also used. After synthesis, the NPs were characterized by TEM and DLS (Table A-1).

Table A-1 NP ligand and corresponding diameter (d_{TEM} and d_h), polydispersity index (PDI), and zeta potential (ZP).

NP Label	Ligand	d_{TEM} (nm)	d_h (nm)	PDI	ZP (mV)
PS Carb	Carboxylate	200 ± 10	220 ± 2	0.02 ± 0.01	-63 ± 9
PS PEG	Polyethylene glycol 2k	200 ± 23	270 ± 7	0.13 ± 0.06	-7 ± 3
Small Citrate	Citrate	80 ± 36	150 ± 3	0.11 ± 0.01	-44 ± 6
Large Citrate		183 ± 48	230 ± 2	0.19 ± 0.02	-49 ± 4
Small PEI	Polyethyleneimine	83 ± 36	230 ± 22	0.22 ± 0.08	29 ± 4
Large PEI		183 ± 48	280 ± 78	0.25 ± 0.09	39 ± 4
Small PVP@Au	Polyvinylpyrrolidone	100 ± 60	270 ± 17	0.31 ± 0.04	-12 ± 4
Large PVP@Au		244 ± 53	316 ± 85	0.22 ± 0.04	-11 ± 3
Small PEI@Au	Polyethyleneimine	100 ± 60	230 ± 17	0.19 ± 0.03	12 ± 3
Large PEI@Au		244 ± 53	291 ± 9	0.15 ± 0.04	12 ± 1
Large PEG@Au	Polyethylene glycol 5k	244 ± 53	610 ± 90	0.38 ± 0.04	-3 ± 3

Proteomic analysis is often performed in one of two ways including abundance and enrichment, which are respectively the amount of protein detected on the protein corona and the protein detected relative to the amount found in serum. We present abundance as normalized abundance as described in the materials and methods (Table A-2).

Table A-2 Normalized abundance (%) for the top fifteen most abundant corona proteins on average, shown for a sample of NPs incubated with 100% FBS. For comparison, the rank order of proteins is shown in parentheses.

Protein	PS Carb	Large Citrate	PS PEG	Large PEG@Au	Large PEI	FBS
Albumin	15.2	25.0	12.5	36.9	33.7	44.68 (1)
Alpha-2-HS-glycoprotein	9.2	13.2	4.4	16.4	14.0	17.12 (2)
Hemoglobin subunit alpha	13.6	1.4	38.1	2.4	1.6	0.64 (14)
Alpha-1-antiproteinase	3.1	6.8	5.4	7.0	7.0	11.43 (3)
Apolipoprotein A-I	5.5	1.5	16.5	1.6	2.2	1.29 (8)
Complement C3	0.6	4.6	0.3	0.9	1.2	0.94 (9)
Prothrombin	7.2	2.1	0.7	0.9	3.1	0.09 (29)
Alpha-2-macroglobulin	1.2	2.2	1.2	2.4	3.9	2.95 (5)
Inter-alpha-trypsin inhibitor heavy chain H3	0.1	0.1	1.0	0.2	3.8	0.2 (25)
Apolipoprotein E	2.1	3.5	4.6	8.5	0.2	0.02 (46)
Serotransferrin	0.2	4.0	0.5	3.7	5.2	8.5 (4)
Hemoglobin fetal subunit beta	8.2	1.1	4.0	1.5	1.0	0.7 (13)
Complement C4	0.1	0.4	0.1	0.4	1.6	0.07 (32)
Complement factor H	0.3	3.2	0.0	0.3	0.1	0.05 (35)
Tetranectin	0.0	4.1	0.0	0.8	0.0	0.02 (49)

Previously we have seen that the protein corona is relatively insensitive to time and incubation concentration [261]. This led us to focus on understanding the effects of core material, surface charge, ligand and size. The 12 unique NP surfaces, 7 unique surface charges, and two unique core sizes allow us to analyze the effects of surface charge, size, core material, and ligand. The difference caused by surface charge can be observed when looking at enrichment of large iron oxide nanoparticles with varied functionalization (Fig A-1). There are no clear generalizable trends between surface charge and enrichment. When looking at individual proteins surface charge trends can

be observed. For example, complement factor H and tetranectin show increased enrichment with decreasing surface charge. Inversely inter-alpha-trypsin heavy chain H3 and adiponectin show increased enrichment on increasing surface charge.

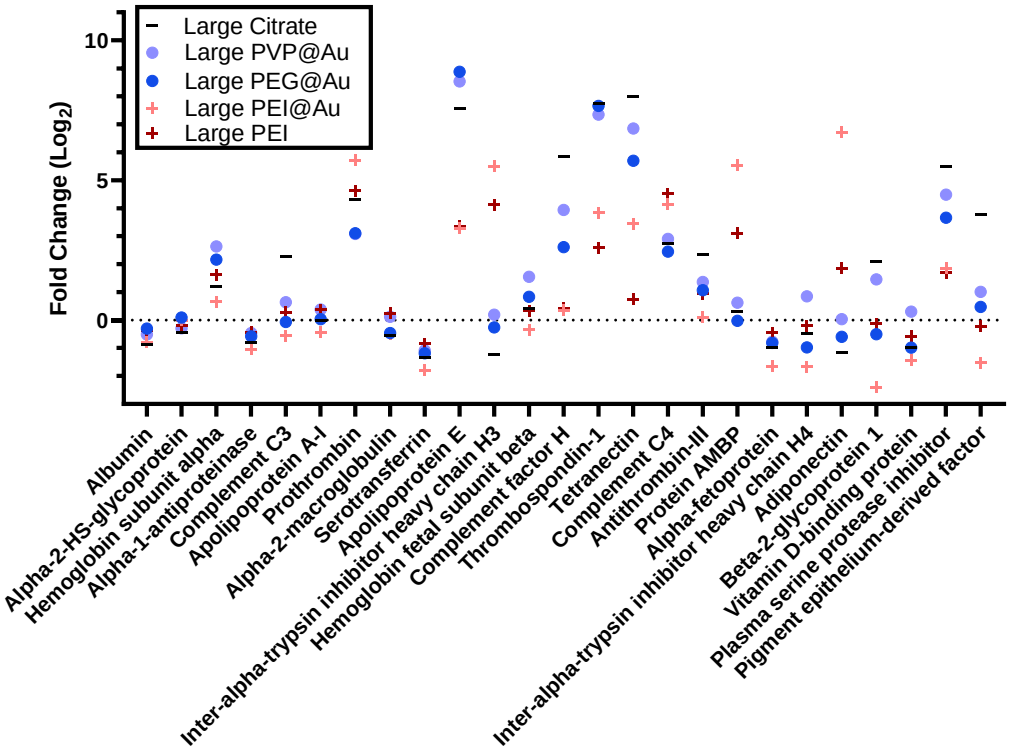

Figure A-1: Enrichment of the top 15 most abundant proteins present on average across all samples relative to their abundance in FBS. Fold change, log base 2, is plotted for each protein observed. Samples are arranged and symbolized by their surface charge; Large Citrate (black minus, -49 mV), Large PVP@Au (light blue circle, -11 mV), Large PEG@Au (dark blue circle,-3 mV), Large PEI@Au (light red plus, +12 mV), and Large PEI (dark red plus, 39 mV).

A.3 Machine learning

Using the above protein corona and NP dataset along with a protein property database built using UniProt [262], NetSurfP 3.0 [263, 264], and BioPython [265], a machine learning dataset was generated. NetSurfP is a natural language processing algorithm that predicts how a protein will fold and provides detailed secondary protein information. This information is output on a per amino acid basis. Using modified code written by Ouassil *et al* [101], we collated this data into aggregate protein quantities like the fraction of exposed polar amino acids, fraction beta sheet, and fraction exposed amino acids. These quantities were combined with data calculated by the BioPython package which calculates using the protein sequence paired with structure data from the Protein DataBank when available. Combining all of protein informatics resulted in 71 physicochemical, structural, and macro protein properties. Previous work in predicting the protein corona has either focused on protein properties [101] or the NP features [102]. Here we are trying to generalize the protein (71), NP (14), and experimental (7) properties of the system.

Random Forest models were utilized here as they are known for their interpretability and low likelihood of overfitting [100]. Predicting the protein corona can be broken down into two problems. Predicting high affinity proteins from a mechanistic perspective or predicting the full protein corona in any given proteome. We used Random Forest Classification to try to predict high affinity proteins and gain some

mechanistic insights, while using Random Forest Regression to try and predict what proteins are going to be in the corona and their specific protein abundance. We made a model for each of these problems independently, in the hopes that we could get better performance overall. With both models we performed recursive feature elimination with five-fold cross validation (Fig. A-2). The best performance for both models was found to be right around 15-20 features. We tested using a variety of scores, to find similar results, but here we show the use of R^2 and Accuracy for regression and classification respectively.

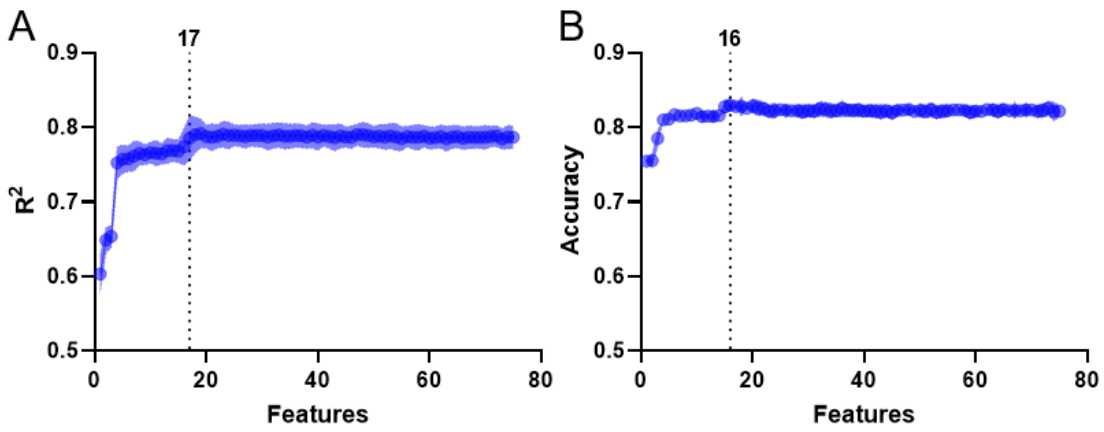

Figure A-2: Recursive Feature Elimination with Cross Validation shows the optimal number of features to use for (A) Random Forest Regression and (B) Random Forest Classification. Five-fold cross validation was performed.

From the feature selection we found the top 10 features for each model to have significant overlap (Table A-3). In the case of the regression model the most important feature was found to be the control abundance, which in this case is the abundance of

the protein in FBS. Both models showed that surface charge (zeta potential), and size of the NPs were important. Surprisingly we see in both models that the fraction of the protein that forms a sheet plays an important role which requires further investigation.

Table A-3 Top ten features with their feature importance for the Random Forest Regression and Random Forest Classification models

Feature	Regression Importance (%)	Feature	Classification Importance (%)
Abundance Controls	47.8	Zeta Potential	10.7
Zeta Potential	6.7	d_h functionalized	9.4
d_h functionalized	4.6	Surface Ligand	6.9
NetSurfP secondary structure sheet	3.5	Secondary structure fraction sheet	6.5
Fraction exposed polar	3.3	d_{tem}	6.0
Fraction phenylalanine	3.2	Isoelectric point	5.9
NetSurfP secondary structure coil	3.1	Fraction arginine	5.9
Fraction asparagine	3.1	Fraction alanine	5.7
Length	3.0	Molecular weight	5.6
NP Incubation Concentration (mg/mL)	2.9	Serum Incubation Concentration (mg/ml)	5.6

References

1. ISO/TC 229 - Nanotechnologies 2021 updated 2021-04-01.

2. Sudha PN, Sangeetha K, Vijayalakshmi K, Barhoum A. Chapter 12 - Nanomaterials history, classification, unique properties, production and market. In: Barhoum A, Makhlouf ASH, editors. Emerging Applications of Nanoparticles and Architecture Nanostructures. Micro and Nano Technologies: Elsevier; 2018. p. 341-84.

3. Bayda S, Adeel M, Tuccinardi T, Cordani M, Rizzolio F. The History of Nanoscience and Nanotechnology: From Chemical–Physical Applications to Nanomedicine. Molecules. 2019;25(1):112.

4. Leon L, Chung EJ, Rinaldi C. Chapter 1 - A brief history of nanotechnology and introduction to nanoparticles for biomedical applications. In: Chung EJ, Leon L, Rinaldi C, editors. Nanoparticles for Biomedical Applications. Micro and Nano Technologies: Elsevier; 2020. p. 1-4.

5. Agar AW. 4.1 The Story of European Commerical Electron Microscopes. Advances in Imaging and Electron Physics1996. p. 415-584.

6. Yeh YC, Creran B, Rotello VM. Gold nanoparticles: preparation, properties, and applications in bionanotechnology. Nanoscale. 2012;4(6):1871-80.

7. Wheeler KE, Chetwynd AJ, Fahy KM, Hong BS, Tochihuitl JA, Foster LA, et al. Environmental dimensions of the protein corona. Nat Nanotechnol. 2021;16(6):617-29.

8. Wang D, Saleh NB, Byro A, Zepp R, Sahle-Demessie E, Luxton TP, et al. Nano-enabled pesticides for sustainable agriculture and global food security. Nat Nanotechnol. 2022;17(4):347-60.

9. Anderson SD, Gwenin VV, Gwenin CD. Magnetic Functionalized Nanoparticles for Biomedical, Drug Delivery and Imaging Applications. Nanoscale Res Lett. 2019;14(1):188.

10. Ashraf SA, Siddiqui AJ, Elkhalifa AEO, Khan MI, Patel M, Alreshidi M, et al. Innovations in nanoscience for the sustainable development of food and agriculture with implications on health and environment. Sci Total Environ. 2021;768:144990.

11. Alkilany AM, Lohse SE, Murphy CJ. The Gold Standard: Gold Nanoparticle Libraries To Understand the Nano–Bio Interface. Accounts of Chemical Research. 2013;46(3):650-61.

12. Hoang KNL, Wheeler KE, Murphy CJ. Isolation Methods Influence the Protein Corona Composition on Gold-Coated Iron Oxide Nanoparticles. Anal Chem. 2022;94(11):4737-46.

13. Jayaram DT, Payne CK. Food-Grade TiO2 Particles Generate Intracellular Superoxide and Alter Epigenetic Modifiers in Human Lung Cells. Chem Res Toxicol. 2020;33(11):2872-9.

14. Fan Y, Marioli M, Zhang K. Analytical characterization of liposomes and other lipid nanoparticles for drug delivery. J Pharm Biomed Anal. 2021;192:113642.

15. Luan X, Sansanaphongpricha K, Myers I, Chen H, Yuan H, Sun D. Engineering exosomes as refined biological nanoplatforms for drug delivery. Acta Pharmacol Sin. 2017;38(6):754-63.

16. Mao G, Wu G, Chen M, Yan C, Tang J, Ma Y, et al. Synthesis of Dual-Emitting CdZnSe/Mn:ZnS Quantum Dots for Sensing the pH Change in Live Cells. Anal Chem. 2022;94(18):6665-71.

17. Pho T, Champion JA. Surface Engineering of Protein Nanoparticles Modulates Transport, Adsorption, and Uptake in Mucus. ACS Appl Mater Interfaces. 2022;14(46):51697-710.

18. Walkey CD, Chan WC. Understanding and controlling the interaction of nanomaterials with proteins in a physiological environment. Chem Soc Rev. 2012;41(7):2780-99.

19. Pelaz B, del Pino P, Maffre P, Hartmann R, Gallego M, Rivera-Fernández S, et al. Surface Functionalization of Nanoparticles with Polyethylene Glycol: Effects on Protein Adsorption and Cellular Uptake. ACS Nano. 2015;9(7):6996-7008.

20. Harris JM, Chess RB. Effect of pegylation on pharmaceuticals. Nat Rev Drug Discov. 2003;2(3):214-21.

21. Shi Z, Neoh KG, Kang ET, Poh CK, Wang W. Surface Functionalization of Titanium with Carboxymethyl Chitosan and Immobilized Bone Morphogenetic Protein-2 for Enhanced Osseointegration. Biomacromolecules. 2009;10(6):1603-11.

22. Wu W, Wu Z, Yu T, Jiang C, Kim WS. Recent progress on magnetic iron oxide nanoparticles: synthesis, surface functional strategies and biomedical applications. Sci Technol Adv Mater. 2015;16(2):023501.

23. Walkey CD, Olsen JB, Song F, Liu R, Guo H, Olsen DW, et al. Protein corona fingerprinting predicts the cellular interaction of gold and silver nanoparticles. ACS Nano. 2014;8(3):2439-55.

24. Castelli A, de Graaf J, Marras S, Brescia R, Goldoni L, Manna L, et al. Understanding and tailoring ligand interactions in the self-assembly of branched colloidal nanocrystals into planar superlattices. Nature Communications. 2018;9(1):1141.

25. Jia H, Zhang Y-F, Zhang C, Ouyang M, Du S. Ligand–Ligand-Interaction-Dominated Self-Assembly of Gold Nanoparticles at the Oil/Water Interface: An Atomic-Scale Simulation. The Journal of Physical Chemistry B. 2023;127(10):2258-66.

26. Zhou T, Li Q, Chen Y, Jiang X. Ligand-regulated self-assembly of luminescent Au nanoparticles towards diverse controllable superstructures. Chemical Communications. 2020;56(90):14023-6.

27. Fam SY, Chee CF, Yong CY, Ho KL, Mariatulqabtiah AR, Tan WS. Stealth Coating of Nanoparticles in Drug-Delivery Systems. Nanomaterials (Basel). 2020;10(4).

28. Mitragotri S, Burke PA, Langer R. Overcoming the challenges in administering biopharmaceuticals: formulation and delivery strategies. Nat Rev Drug Discov. 2014;13(9):655-72.

29. Allen TM, Cullis PR. Drug Delivery Systems: Entering the Mainstream. Science. 2004;303(5665):1818-22.

30. Aggarwal P, Hall JB, McLeland CB, Dobrovolskaia MA, McNeil SE. Nanoparticle interaction with plasma proteins as it relates to particle biodistribution, biocompatibility and therapeutic efficacy. Adv Drug Deliver Rev. 2009;61(6):428-37.

31. Mitchell MJ, Billingsley MM, Haley RM, Wechsler ME, Peppas NA, Langer R. Engineering precision nanoparticles for drug delivery. Nat Rev Drug Discov. 2021;20(2):101-24.

32. Mura S, Nicolas J, Couvreur P. Stimuli-responsive nanocarriers for drug delivery. Nature Materials. 2013;12(11):991-1003.

33. Rahim MA, Jan N, Khan S, Shah H, Madni A, Khan A, et al. Recent Advancements in Stimuli Responsive Drug Delivery Platforms for Active and Passive Cancer Targeting. Cancers. 2021;13(4):670.

34. Vinchhi P, Rawal SU, Patel MM. Chapter 13 - External stimuli-responsive drug delivery systems. In: Chappel E, editor. Drug Delivery Devices and Therapeutic Systems. Developments in Biomedical Engineering and Bioelectronics: Academic Press; 2021. p. 267-88.

35. Zhang J, Lin Y, Lin Z, Wei Q, Qian J, Ruan R, et al. Stimuli-Responsive Nanoparticles for Controlled Drug Delivery in Synergistic Cancer Immunotherapy. Advanced Science. 2022;9(5):2103444.

36. Casals E, Pfaller T Fau - Duschl A, Duschl A Fau - Oostingh GJ, Oostingh Gj Fau - Puntes V, Puntes V. Time evolution of the nanoparticle protein corona. ACS Nano. 2010;4(1936-086X (Electronic)):3623-32.

37. Grafe C, Weidner A, Luhe MV, Bergemann C, Schacher FH, Clement JH, et al. Intentional formation of a protein corona on nanoparticles: Serum concentration affects protein corona mass, surface charge, and nanoparticle-cell interaction. Int J Biochem Cell Biol. 2016;75:196-202.

38. Latreille PL, Le Goas M, Salimi S, Robert J, De Crescenzo G, Boffito DC, et al. Scratching the Surface of the Protein Corona: Challenging Measurements and Controversies. ACS Nano. 2022;16(2):1689-707.

39. Partikel K, Korte R, Mulac D, Humpf HU, Langer K. Serum type and concentration both affect the protein-corona composition of PLGA nanoparticles. Beilstein J Nanotechnol. 2019;10:1002-15.

40. Mahmoudi M, Landry MP, Moore A, Coreas R. The protein corona from nanomedicine to environmental science. Nature Reviews Materials. 2023.

41. Nel AE, Madler L, Velegol D, Xia T, Hoek EM, Somasundaran P, et al. Understanding biophysicochemical interactions at the nano-bio interface. Nat Mater. 2009;8(7):543-57.

42. Wang Z, Hood ED, Nong J, Ding J, Marcos-Contreras OA, Glassman PM, et al. Combating Complement's Deleterious Effects on Nanomedicine by Conjugating Complement Regulatory Proteins to Nanoparticles. Adv Mater. 2022;34(8):e2107070.

43. de Castro CE, Panico K, Stangherlin LM, Ribeiro CAS, da Silva MCC, Carneiro-Ramos MS, et al. The Protein Corona Conundrum: Exploring the Advantages and Drawbacks of its Presence around Amphiphilic Nanoparticles. Bioconjug Chem. 2020;31(11):2638-47.

44. Mohajerani A, Burnett L, Smith JV, Kurmus H, Milas J, Arulrajah A, et al. Nanoparticles in Construction Materials and Other Applications, and Implications of Nanoparticle Use. Materials. 2019;12(19):3052.

45. Giraldo JP, Kruss S. Nanosensors for monitoring plant health. Nature Nanotechnology. 2023;18(2):107-8.

46. Newkirk GM, de Allende P, Jinkerson RE, Giraldo JP. Nanotechnology Approaches for Chloroplast Biotechnology Advancements. Frontiers in Plant Science. 2021;12:691295.

47. Xu L, Xu M, Wang R, Yin Y, Lynch I, Liu S. The Crucial Role of Environmental Coronas in Determining the Biological Effects of Engineered Nanomaterials. Small. 2020;16(36):e2003691.

48. Hofmann T, Lowry GV, Ghoshal S, Tufenkji N, Brambilla D, Dutcher JR, et al. Technology readiness and overcoming barriers to sustainably implement nanotechnology-enabled plant agriculture. Nature Food. 2020;1(7):416-25.

49. Hanus MJ, Harris AT. Nanotechnology innovations for the construction industry. Progress in Materials Science. 2013;58(7):1056-102.

50. Lee J, Mahendra S, Alvarez PJJ. Nanomaterials in the Construction Industry: A Review of Their Applications and Environmental Health and Safety Considerations. ACS Nano. 2010;4(7):3580-90.

51. Wheeler KE, Chetwynd AJ, Fahy KM, Hong BS, Tochihuitl JA, Foster LA, et al. Environmental dimensions of the protein corona. Nature Nanotechnology. 2021;16(6):617-29.

52. Zahra Z, Habib Z, Chung S, Badshah MA. Exposure Route of TiO2 NPs from Industrial Applications to Wastewater Treatment and Their Impacts on the Agro-Environment. Nanomaterials. 2020;10(8):1469.

53. Runa S, Hussey M, Payne CK. Nanoparticle–Cell Interactions: Relevance for Public Health. Journal of Physical Chemistry B. 2018;122(3):1009-16.

54. Mahmudun Nabi M, Wang J, Erfani M, Goharian E, Baalousha M. Urban runoff drives titanium dioxide engineered particle concentrations in urban watersheds: field measurements. Environmental Science: Nano. 2023.

55. Donaldson K, Seaton A. A short history of the toxicology of inhaled particles. Particle and Fibre Toxicology. 2012;9(1):13.

56. Rayens NT, Rayens EA, Tighe RM. Co-occurrence of pneumoconiosis with COPD, pneumonia and lung cancer. Occup Med (Lond). 2022;72(8):527-33.

57. Current intelligence bulletin 63: occupational exposure to titanium dioxide.: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health; 2011 2011-04-01.

58. Weir A, Westerhoff P, Fabricius L, Hristovski K, von Goetz N. Titanium Dioxide Nanoparticles in Food and Personal Care Products. Environmental Science and Technology. 2012;46(4):2242-50.

59. CFR. Part 73 Listing of Color Additives Exempt from Certification. Washington DC: Code of Federal Regulations; 2023.

60. Poulsen KM, Payne CK. Concentration and composition of the protein corona as a function of incubation time and serum concentration: an automated approach to the protein corona. Analytical and Bioanalytical Chemistry. 2022;414(24):7265-75.

61. Poulsen KM, Pho T, Champion JA, Payne CK. Automation and low-cost proteomics for characterization of the protein corona: experimental methods for big data. Analytical and Bioanalytical Chemistry. 2020;412(24):6543-51.

62. Fleischer CC, Payne CK. Nanoparticle-cell interactions: molecular structure of the protein corona and cellular outcomes. Accounts of Chemical Research. 2014;47(8):2651-9.

63. Rogan MP, Geraghty P, Greene CM, O'Neill SJ, Taggart CC, McElvaney NG. Antimicrobial proteins and polypeptides in pulmonary innate defence. Respir Res. 2006;7(1):29.

64. Whitsett JA, Alenghat T. Respiratory epithelial cells orchestrate pulmonary innate immunity. Nat Immunol. 2015;16(1):27-35.

65. Numata M, Voelker DR. Anti-inflammatory and anti-viral actions of anionic pulmonary surfactant phospholipids. Biochim Biophys Acta Mol Cell Biol Lipids. 2022;1867(6):159139.

66. Kumar A, Bicer EM, Morgan AB, Pfeffer PE, Monopoli M, Dawson KA, et al. Enrichment of immunoregulatory proteins in the biomolecular corona of nanoparticles within human respiratory tract lining fluid. Nanomedicine. 2016;12(4):1033-43.

67. Bonner JC, Silva RM, Taylor AJ, Brown JM, Hilderbrand SC, Castranova V, et al. Interlaboratory evaluation of rodent pulmonary responses to engineered nanomaterials: the NIEHS Nano GO Consortium. Environ Health Perspect. 2013;121(6):676-82.

68. Xia T, Hamilton RF, Bonner JC, Crandall ED, Elder A, Fazlollahi F, et al. Interlaboratory evaluation of in vitro cytotoxicity and inflammatory responses to engineered nanomaterials: the NIEHS Nano GO Consortium. Environ Health Perspect. 2013;121(6):683-90.

69. Dowell JA, Wright LJ, Armstrong EA, Denu JM. Benchmarking Quantitative Performance in Label-Free Proteomics. ACS Omega. 2021;6(4):2494-504.

70. Millán-Oropeza A, Blein-Nicolas M, Monnet V, Zivy M, Henry C. Comparison of Different Label-Free Techniques for the Semi-Absolute Quantification of Protein Abundance. Proteomes. 2022;10(1):2.

71. Nakayasu ES, Gritsenko M, Piehowski PD, Gao Y, Orton DJ, Schepmoes AA, et al. Tutorial: best practices and considerations for mass-spectrometry-based protein biomarker discovery and validation. Nature Protocols. 2021;16(8):3737-60.

72. Pusch W, Flocco MT, Leung S-M, Thiele H, Kostrzewa M. Mass spectrometry-based clinical proteomics. Pharmacogenomics. 2003;4(4):463-76.

73. Thomas SN, Zhang H. Targeted proteomic assays for the verification of global proteomics insights. Expert review of proteomics. 2016;13(10):897-9.

74. Vroman L, Adams A, Fischer G, Munoz P. Interaction of high molecular weight kininogen, factor XII, and fibrinogen in plasma at interfaces. Blood. 1980;55(1):156-9.

75. Pisani C, Gaillard JC, Odorico M, Nyalosaso JL, Charnay C, Guari Y, et al. The timeline of corona formation around silica nanocarriers highlights the role of the protein interactome. Nanoscale. 2017;9(5):1840-51.

76. Fleischer CC, Payne CK. Nanoparticle surface charge mediates the cellular receptors used by protein-nanoparticle complexes. J Phys Chem B. 2012;116(30):8901-7.

77. Rabel M, Warncke P, Thurmer M, Gruttner C, Bergemann C, Kurland HD, et al. The differences of the impact of a lipid and protein corona on the colloidal stability, toxicity, and degradation behavior of iron oxide nanoparticles. Nanoscale. 2021;13(20):9415-35.

78. Hedayati M, Kipper MJ, Krapf D. Anomalous protein kinetics on low-fouling surfaces. Phys Chem Chem Phys. 2020;22(9):5264-71.

79. Hutanu D. Recent Applications of Polyethylene Glycols (PEGs) and PEG Derivatives. Modern Chemistry & Applications. 2014;02(02).

80. Jang GJ, Jeong JY, Kang J, Cho W, Han SY. Size Dependence Unveiling the Adsorption Interaction of High-Density Lipoprotein Particles with PEGylated Gold Nanoparticles in Biomolecular Corona Formation. Langmuir. 2021;37(32):9755-63.

81. Knowles DB, Shkel IA, Phan NM, Sternke M, Lingeman E, Cheng X, et al. Chemical Interactions of Polyethylene Glycols (PEGs) and Glycerol with Protein Functional Groups: Applications to Effects of PEG and Glycerol on Protein Processes. Biochemistry. 2015;54(22):3528-42.

82. Naidu PSR, Norret M, Smith NM, Dunlop SA, Taylor NL, Fitzgerald M, et al. The Protein Corona of PEGylated PGMA-Based Nanoparticles is Preferentially Enriched with Specific Serum Proteins of Varied Biological Function. Langmuir. 2017;33(45):12926-33.

83. Partikel K, Korte R, Stein NC, Mulac D, Herrmann FC, Humpf HU, et al. Effect of nanoparticle size and PEGylation on the protein corona of PLGA nanoparticles. Eur J Pharm Biopharm. 2019;141:70-80.

84. Sherman MR, Williams LD, Sobczyk MA, Michaels SJ, Saifer MG. Role of the methoxy group in immune responses to mPEG-protein conjugates. Bioconjug Chem. 2012;23(3):485-99.

85. Wu J, Zhao C, Lin W, Hu R, Wang Q, Chen H, et al. Binding characteristics between polyethylene glycol (PEG) and proteins in aqueous solution. Journal of Materials Chemistry B. 2014;2(20).

86. Yang Q, Jones SW, Parker CL, Zamboni WC, Bear JE, Lai SK. Evading immune cell uptake and clearance requires PEG grafting at densities substantially exceeding the minimum for brush conformation. Mol Pharm. 2014;11(4):1250-8.

87. Yang Q, Lai SK. Anti-PEG immunity: emergence, characteristics, and unaddressed questions. Wiley Interdiscip Rev Nanomed Nanobiotechnol. 2015;7(5):655-77.

88. Zhang P, Sun F, Liu S, Jiang S. Anti-PEG antibodies in the clinic: Current issues and beyond PEGylation. J Control Release. 2016;244(Pt B):184-93.

89. Lubich C, Allacher P, de la Rosa M, Bauer A, Prenninger T, Horling FM, et al. The Mystery of Antibodies Against Polyethylene Glycol (PEG) - What do we Know? Pharm Res. 2016;33(9):2239-49.

90. Lima T, Bernfur K, Vilanova M, Cedervall T. Understanding the Lipid and Protein Corona Formation on Different Sized Polymeric Nanoparticles. Sci Rep. 2020;10(1):1129.

91. Melby ES, Lohse SE, Park JE, Vartanian AM, Putans RA, Abbott HB, et al. Cascading Effects of Nanoparticle Coatings: Surface Functionalization Dictates the Assemblage of Complexed Proteins and Subsequent Interaction with Model Cell Membranes. ACS Nano. 2017;11(6):5489-99.

92. Zhang H, Burnum KE, Luna ML, Petritis BO, Kim JS, Qian WJ, et al. Quantitative proteomics analysis of adsorbed plasma proteins classifies nanoparticles with different surface properties and size. Proteomics. 2011;11(23):4569-77.

93. Jayaram DT, Pustulka SM, Mannino RG, Lam WA, Payne CK. Protein Corona in Response to Flow: Effect on Protein Concentration and Structure. Biophys J. 2018;115(2):209-16.

94. Tomak A, Cesmeli S, Hanoglu BD, Winkler D, Oksel Karakus C. Nanoparticle-protein corona complex: understanding multiple interactions between environmental factors, corona formation, and biological activity. Nanotoxicology. 2021;15(10):1331-57.

95. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. the Journal of machine Learning research. 2011;12:2825-30.

96. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. 2016.

97. Kuhn M. Building Predictive Models in R Using the caret Package. Journal of Statistical Software. 2008;028(i05).

98. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. 2019.

99. Duan Y, Coreas R, Liu Y, Bitounis D, Zhang Z, Parviz D, et al. Prediction of protein corona on nanomaterials by machine learning using novel descriptors. NanoImpact. 2020;17.

100. Findlay MR, Freitas DN, Mobed-Miremadi M, Wheeler KE. Machine learning provides predictive analysis into silver nanoparticle protein corona formation from physicochemical properties. Environ Sci Nano. 2018;5(1):64-71.

101. Ouassil N, Pinals Rebecca L, Del Bonis-O'Donnell Jackson T, Wang Jeffrey W, Landry Markita P. Supervised learning model predicts protein adsorption to carbon nanotubes. Sci Adv.8(1):eabm0898.

102. Ban Z, Yuan P, Yu F, Peng T, Zhou Q, Hu X. Machine learning predicts the functional composition of the protein corona and the cellular recognition of nanoparticles. Proc Natl Acad Sci U S A. 2020;117(19):10492-9.

103. Poulsen KM, Pho T, Champion JA, Payne CK. Automation and low-cost proteomics for characterization of the protein corona: experimental methods for big data. Anal Bioanal Chem. 2020;412(24):6543-51.

104. Lynch I, Cedervall T, Lundqvist M, Cabaleiro-Lago C, Linse S, Dawson KA. The nanoparticle - protein complex as a biological entity; a complex fluids and surface science challenge for the 21st century. Adv Colloid Interfac. 2007;134-135:167-74.

105. Walczyk D, Bombelli FB, Monopoli MP, Lynch I, Dawson KA. What the cell "sees" in bionanoscience. J Am Chem Soc. 2010;132:5761-8.

106. Park S, Hamad-Schifferli K. Nanoscale interfaces to biology. Curr Opin Chem Biol. 2010;14(5):616-22.

107. Walkey CD, Chan WCW. Understanding and controlling the interaction of nanomaterials with proteins in a physiological environment. Chem Soc Rev. 2012;41(7):2780-99.

108. Del Pino P, Pelaz B, Zhang Q, Maffre P, Nienhaus GU, Parak WJ. Protein corona formation around nanoparticles - from the past to the future. Mater Horiz. 2014;1(3):301-13.

109. Fleischer CC, Payne CK. Nanoparticle-cell interactions: Molecular structure of the protein corona and cellular outcomes. Acc Chem Res. 2014;47:2651-9.

110. Ke PC, Lin S, Parak WJ, Davis TP, Caruso F. A Decade of the Protein Corona. ACS Nano. 2017;11(12):11773-6.

111. Mahmoudi M, Bertrand N, Zope H, Farokhzad OC. Emerging understanding of the protein corona at the nano-bio interfaces. Nano Today. 2016;11(6):817-32.

112. Yang ST, Liu Y, Wang YW, Cao AN. Biosafety and Bioapplication of Nanomaterials by Designing Protein-Nanoparticle Interactions. Small. 2013;9(9-10):1635-53.

113. Payne CK. A protein corona primer for physical chemists. J Chem Phys. 2019;151(13):130901.

114. Owens III DE, Peppas NA. Opsonization, biodistribution, and pharmacokinetics of polymeric nanoparticles. Int J Pharm. 2006;307(1):93-102.

115. Khlebtsov N, Dykman L. Biodistribution and toxicity of engineered gold nanoparticles: a review of in vitro and in vivo studies. Chem Soc Rev. 2011;40(3):1647-71.

116. Duan XP, Li YP. Physicochemical Characteristics of Nanoparticles Affect Circulation, Biodistribution, Cellular Internalization, and Trafficking. Small. 2013;9(9-10):1521-32.

117. Chinen AB, Guan CM, Ko CH, Mirkin CA. The Impact of Protein Corona Formation on the Macrophage Cellular Uptake and Biodistribution of Spherical Nucleic Acids. Small. 2017;13(16):#1603847.

118. Wilhelm S, Tavares AJ, Dai Q, Ohta S, Audet J, Dvorak HF, et al. Analysis of nanoparticle delivery to tumours. Nat Rev Mater. 2016;1:16014.

119. Hellstrand E, Lynch I, Andersson A, Drakenberg T, Dahlbäck B, Dawson KA, et al. Complete high-density lipoproteins in nanoparticle corona. FEBS J. 2009;276(12):3372-81.

120. Olenick LL, Troiano JM, Vartanian A, Melby ES, Mensch AC, Zhang L, et al. Lipid Corona Formation from Nanoparticle Interactions with Bilayers. Chem. 2018;4(11):2709-23.

121. Pink M, Verma N, Kersch C, Schmitz-Spanke S. Identification and characterization of small organic compounds within the corona formed around engineered nanoparticles. Environ Sci Nano. 2018;5(6):1420-7.

122. Khot LR, Sankaran S, Maja JM, Ehsani R, Schuster EW. Applications of nanomaterials in agricultural production and crop protection: A review. Crop Protection. 2012;35:64-70.

123. Liu R, Lal R. Potentials of engineered nanoparticles as fertilizers for increasing agronomic productions. Sci Total Environ. 2015;514:131-9.

124. Kah M, Hofmann T. Nanopesticide research: Current trends and future priorities. Environ Int. 2014;63:224-35.

125. Giraldo JP, Wu HH, Newkirk GM, Kruss S. Nanobiotechnology approaches for engineering smart plant sensors. Nat Nanotechnol. 2019;14(6):541-53.

126. Hochella MF, Mogk DW, Ranville J, Allen IC, Luther GW, Marr LC, et al. Natural, incidental, and engineered nanomaterials and their impacts on the Earth system. Science. 2019;363:eaau8299.

127. Mauter MS, Zucker I, Perreault F, Werber JR, Kim JH, Elimelech M. The role of nanotechnology in tackling global water challenges. Nat Sustain. 2018;1(4):166-75.

128. Kaegi R, Voegelin A, Sinnet B, Zuleeg S, Siegrist H, Burkhardt M. Transformation of AgCl nanoparticles in a sewer system - A field study. Sci Total Environ. 2015;535:20-7.

129. Lopez H, Lobaskin V. Coarse-grained model of adsorption of blood plasma proteins onto nanoparticles. J Chem Phys. 2015;143(24):#243138.

130. Voicescu M, Ionescu S, Angelescu DG. Spectroscopic and coarse-grained simulation studies of the BSA and HSA protein adsorption on silver nanoparticles. J Nanopart Res. 2012;14(10):1174.

131. Ding F, Radic S, Chen R, Chen P, Geitner NK, Brown JM, et al. Direct observation of a single nanoparticle–ubiquitin corona formation. Nanoscale. 2013;5(19):9162-9.

132. Wei S, Ahlstrom LS, Brooks CL. Exploring protein–nanoparticle interactions with coarse-grained protein folding models. Small. 2017;13(18):#1603748.

133. Ding H-m, Ma Y-q. Computer simulation of the role of protein corona in cellular delivery of nanoparticles. Biomaterials. 2014;35(30):8703-10.

134. Li R, Chen R, Chen P, Wen Y, Ke PC, Cho SS. Computational and experimental characterizations of silver nanoparticle–apolipoprotein biocorona. J Phys Chem B. 2013;117(43):13451-6.

135. Li R, Stevens CA, Cho SS. Molecular Dynamics Simulations of Biocorona Formation. In: Suzuki J, Nakano T, Moore MJ, editors. Modeling, Methodologies and Tools for Molecular and Nano-scale Communications: Modeling, Methodologies and Tools. Cham: Springer International Publishing; 2017. p. 241-56.

136. Manning MD, Kwansa AL, Oweida T, Peerless JS, Singh A, Yingling YG. Progress in ligand design for monolayer-protected nanoparticles for nanobio interfaces. Biointerphases. 2018;13(6):06D502.

137. Deyev S, Proshkina G, Ryabova A, Tavanti F, Menziani MC, Eidelshtein G, et al. Synthesis, Characterization, and Selective Delivery of DARPin–Gold Nanoparticle Conjugates to Cancer Cells. Bioconj Chem. 2017;28(10):2569-74.

138. Tavanti F, Pedone A, Menziani MC. Competitive binding of proteins to gold nanoparticles disclosed by molecular dynamics simulations. J Phys Chem C. 2015;119(38):22172-80.

139. Cui Q, Hernandez R, Mason SE, Frauenheim T, Pedersen JA, Geiger F. Sustainable nanotechnology: opportunities and challenges for theoretical/computational studies. J Phys Chem B. 2016;120(30):7297-306.

140. Liang D, Hong J, Fang D, Bennett JW, Mason SE, Hamers RJ, et al. Analysis of the conformational properties of amine ligands at the gold/water interface with QM, MM and QM/MM simulations. Phys Chem Chem Phys. 2018;20:3349-62.

141. Van Lehn RC, Alexander-Katz A. Structure of mixed-monolayer-protected nanoparticles in aqueous salt solution from atomistic molecular dynamics simulations. J Phys Chem C. 2013;117(39):20104-15.

142. Brancolini G, Maschio MC, Cantarutti C, Corazza A, Fogolari F, Bellotti V, et al. Citrate stabilized gold nanoparticles interfere with amyloid fibril formation: D76N and ΔN6 β2-microglobulin variants. Nanoscale. 2018;10(10):4793-806.

143. Fleischer CC, Payne CK. Nanoparticle surface charge mediates the cellular receptors used by protein-nanoparticle complexes. J Phys Chem B. 2012;116:8901-7.

144. Fleischer CC, Kumar U, Payne CK. Cellular binding of anionic nanoparticles is inhibited by serum proteins independent of nanoparticle composition. Biomat Sci. 2013;1:975-82.

145. Fleischer CC, Payne CK. Secondary structure of corona proteins determines the cell surface receptors used by nanoparticles. J Phys Chem B. 2014;118:14017-26.

146. Hill A, Payne CK. Impact of serum proteins on MRI contrast agents: Cellular binding and T2 relaxation. RSC Adv. 2014;4(60):31735-44.

147. Walkey CD, Olsen JB, Song F, Liu R, Guo H, Olsen DWH, et al. Protein corona fingerprinting predicts the cellular interaction of gold and silver nanoparticles. ACS Nano. 2014;8(3):2439-55.

148. Findlay MR, Freitas DN, Mobed-Miremadi M, Wheeler KE. Machine learning provides predictive analysis into silver nanoparticle protein corona formation from physicochemical properties. Environ Sci Nano. 2018;5(1):64-71.

149. Lazarovits J, Sindhwani S, Tavares AJ, Zhang Y, Song F, Audet J, et al. Supervised learning and mass spectrometry predicts the in vivo fate of nanomaterials. ACS Nano. 2019.

150. Chan EM, Xu C, Mao AW, Han G, Owen JS, Cohen BE, et al. Reproducible, high-throughput synthesis of colloidal nanocrystals for optimization in multidimensional parameter space. Nano Lett. 2010;10(5):1874-85.

151. Pendleton IM, Cattabriga G, Li Z, Najeeb MA, Friedler SA, Norquist AJ, et al. Experiment Specification, Capture and Laboratory Automation Technology (ESCALATE): a software pipeline for automated chemical experimentation and data management. MRS Communications. 2019;9(3):846-59.

152. Runa S, Khanal D, Kemp ML, Payne CK. TiO2 nanoparticles alter the expression of peroxiredoxin antioxidant genes. J Phys Chem C. 2016;120:20736-42.

153. Jayaram DT, Runa S, Kemp ML, Payne CK. Nanoparticle-induced oxidation of corona proteins initiates an oxidative stress response in cells. Nanoscale. 2017;9:7595-601.

154. Jayaram DT, Pustulka SM, Mannino RG, Lam WA. Protein corona in response to flow: Effect on protein concentration and structure. Biophysical J. 2018;115:209-16.

155. Jayaram DT, Kumar A, Kippner LE, Ho P-Y, Kemp ML, Fan Y, et al. TiO 2 nanoparticles generate superoxide and alter gene expression in human lung cells. RSC Adv. 2019;9(43):25039-47.

156. Etorki AM, Gao M, Sadeghi R, Maldonado-mejia LF, Kokini JL. Effects of Desolvating Agent Types , Ratios , and Temperature on Size and Nanostructure of Nanoparticles from α -Lactalbumin and Ovalbumin. J Food Sci. 2016;81:E2511-E20.

157. Chang TZ, Stadmiller SS, Staskevicius E, Champion JA. Effects of ovalbumin protein nanoparticle vaccine size and coating on dendritic cell processing. Biomat Sci. 2017;5:223-33.

158. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9(7):671.

159. Rimkus G, Bremer-Streck S, Grüttner C, Kaiser WA, Hilger I. Can we accurately quantify nanoparticle associated proteins when constructing high-affinity MRI molecular imaging probes? Contrast media & molecular imaging. 2011;6(3):119-25.

160. Gruttner C, Muller K, Teller J. A rapid assay to measure the shielding of iron oxide cores by the particle shell. IEEE Trans Magn. 2012;49(1):177-81.

161. Muller J, Bauer KN, Prozeller D, Simon J, Mailander V, Wurm FM, et al. Coating nanoparticles with tunable surfactants facilitates control over the protein corona. Biomaterials. 2017;115:1-8.

162. Chang TZ, Diambou I, Rok J, Baozhong K, Champion JA. Host- and Pathogen-Derived Adjuvant Coatings on Protein Nanoparticle Vaccines. Bioeng Transl Med. 2017;2:120-30.

163. Deng L, Mohan T, Chang TZ, Gonzalez GX, Wang Y, Kwon Y, et al. Double-layerd protein nanoparticles induce broad protection against divergent influenza A viruses. Nature Commun. 2018;9:1-12.

164. Sakulkhu U, Maurizi L, Mahmoudi M, Motazacker M, Vries M, Gramoun A, et al. Ex situ evaluation of the composition of protein corona of intravenously injected superparamagnetic nanoparticles in rats. Nanoscale. 2014;6(19):11439-50.

165. Chen F, Wang G, Griffin JI, Brenneman B, Banda NK, Holers VM, et al. Complement proteins bind to nanoparticle protein corona and undergo dynamic exchange in vivo. Nat Nanotechnol. 2017;12(4):387.

166. Vizcaíno JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Rios D, et al. ProteomeXchange provides globally coordinated proteomics data submission and dissemination. Nat Biotechnol. 2014;32(3):223-6.

167. Fleischer CC, Payne CK. Nanoparticle-cell interactions: molecular structure of the protein corona and cellular outcomes. Acc Chem Res. 2014;47(8):2651-9.

168. Monopoli MP, Aberg C, Salvati A, Dawson KA. Biomolecular coronas provide the biological identity of nanosized materials. Nat Nanotechnol. 2012;7(12):779-86.

169. Nienhaus K, Nienhaus GU. Towards a molecular-level understanding of the protein corona around nanoparticles – Recent advances and persisting challenges. Curr Opin Biomed Eng. 2019;10:11-22.

170. Kobos L, Shannahan J. Biocorona-induced modifications in engineered nanomaterial-cellular interactions impacting biomedical applications. Wiley Interdiscip Rev Nanomed Nanobiotechnol. 2020;12(3):e1608.

171. Abarca-Cabrera L, Fraga-Garcia P, Berensmeier S. Bio-nano interactions: binding proteins, polysaccharides, lipids and nucleic acids onto magnetic nanoparticles. Biomater Res. 2021;25(1):12.

172. Cai R, Chen C. The Crown and the Scepter: Roles of the Protein Corona in Nanomedicine. Adv Mater. 2019;31(45):e1805740.

173. Payne CK. A protein corona primer for physical chemists. J Chem Phys. 2019;151(13):130901.

174. Docter D, Strieth S, Westmeier D, Hayden O, Gao M, Knauer SK, et al. No king without a crown – impact of the nanomaterial-protein corona on nanobiomedicine. Nanomedicine. 2015;10(3):503-19.

175. Frtus A, Smolkova B, Uzhytchak M, Lunova M, Jirsa M, Kubinova S, et al. Analyzing the mechanisms of iron oxide nanoparticles interactions with cells: A road from failure to success in clinical applications. J Control Release. 2020;328:59-77.

176. Hamad-Schifferli K. Exploiting the novel properties of protein coronas: emerging applications in nanomedicine. Nanomedicine. 2015;10(10):1663-74.

177. Deng D, Li Y, Xue J, Wang J, Ai G, Li X, et al. Gold nanoparticle-based beacon to detect STAT5b mRNA expression in living cells: a case optimized by bioinformatics screen. Int J Nanomedicine. 2015;10:3231-44.

178. Verma A, Stellacci F. Effect of surface properties on nanoparticle-cell interactions. Small. 2010;6(1):12-21.

179. Carrillo-Carrion C, Carril M, Parak WJ. Techniques for the experimental investigation of the protein corona. Curr Opin Biotechnol. 2017;46:106-13.

180. Pisani C, Gaillard JC, Dorandeu C, Charnay C, Guari Y, Chopineau J, et al. Experimental separation steps influence the protein content of corona around mesoporous silica nanoparticles. Nanoscale. 2017;9(18):5769-72.

181. Zarei M, Aalaie J. Profiling of nanoparticle-protein interactions by electrophoresis techniques. Anal Bioanal Chem. 2019;411(1):79-96.

182. Blume JE, Manning WC, Troiano G, Hornburg D, Figa M, Hesterberg L, et al. Rapid, deep and precise profiling of the plasma proteome with multi-nanoparticle protein corona. Nat Commun. 2020;11(1):3662.

183. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9(7):671-5.

184. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008;26(12):1367-72.

185. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat Protoc. 2016;11(12):2301-19.

186. cRAP protein sequences: The Global Proteome Machine; 2012 [Available from: https://www.thegpm.org/crap/index.html.

187. Bruckner M, Simon J, Jiang S, Landfester K, Mailander V. Preparation of the protein corona: How washing shapes the proteome and influences cellular uptake of nanocarriers. Acta Biomater. 2020;114:333-42.

188. Docter D, Distler U, Storck W, Kuharev J, Wunsch D, Hahlbrock A, et al. Quantitative profiling of the protein coronas that form around nanoparticles. Nat Protoc. 2014;9(9):2030-44.

189. Simon J, Kuhn G, Fichter M, Gehring S, Landfester K, Mailander V. Unraveling the In Vivo Protein Corona. Cells. 2021;10(1):132.

190. Kastan Jonathan P, Dobrikova Elena Y, Bryant Jeffrey D, Gromeier M. CReP mediates selective translation initiation at the endoplasmic reticulum. Sci Adv.6(23):eaba0745.

191. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. 2019;47(1362-4962 (Electronic)):D442-D50.

192. Doorley GW, Payne CK. Nanoparticles act as protein carriers during cellular internalization. Chem Commun (Camb). 2012;48(24):2961-3.

193. Hill A, Payne CK. Impact of Serum Proteins on MRI Contrast Agents: Cellular Binding and T2 relaxation. RSC Adv. 2014;4:31735-44.

194. Doorley GW, Payne CK. Cellular binding of nanoparticles in the presence of serum proteins. Chem Commun (Camb). 2011;47(1):466-8.

195. Tenzer S, Docter D, Kuharev J, Musyanovych A, Fetz V, Hecht R, et al. Rapid formation of plasma protein corona critically affects nanoparticle pathophysiology. Nat Nanotechnol. 2013;8(10):772-81.

196. Monopoli MP, Walczyk D, Campbell A, Elia G, Lynch I, Bombelli FB, et al. Physical-chemical aspects of protein corona: relevance to in vitro and in vivo biological impacts of nanoparticles. J Am Chem Soc. 2011;133(8):2525-34.

197. Runa S, Lakadamyali M, Kemp ML, Payne CK. TiO2 Nanoparticle-Induced Oxidation of the Plasma Membrane: Importance of the Protein Corona. J Phys Chem B. 2017;121(37):8619-25.

198. Jayaram DT, Runa S, Kemp ML, Payne CK. Nanoparticle-induced oxidation of corona proteins initiates an oxidative stress response in cells. Nanoscale. 2017;9(22):7595-601.

199. Dudzik D, Barbas-Bernardos C, Garcia A, Barbas C. Quality assurance procedures for mass spectrometry untargeted metabolomics. a review. J Pharm Biomed Anal. 2018;147:149-73.

200. Runa S, Khanal D, Kemp ML, Payne CK. TiO2 Nanoparticles Alter the Expression of Peroxiredoxin Antioxidant Genes. J Phys Chem C. 2016;120(37):20736-42.

201. Yu Q, Zhao L, Guo C, Yan B, Su G. Regulating Protein Corona Formation and Dynamic Protein Exchange by Controlling Nanoparticle Hydrophobicity. Front Bioeng Biotechnol. 2020;8:210.

202. Strojan K, Leonardi A, Bregar VB, Krizaj I, Svete J, Pavlin M. Dispersion of Nanoparticles in Different Media Importantly Determines the Composition of Their Protein Corona. PLoS One. 2017;12(1):e0169552.

203. Venerando R, Miotto G, Magro M, Dallan M, Baratella D, Bonaiuto E, et al. Magnetic Nanoparticles with Covalently Bound Self-Assembled Protein Corona for Advanced Biomedical Applications. J Phys Chem C. 2013;117(39):20320-31.

204. Sakulkhu U, Maurizi L, Mahmoudi M, Motazacker M, Vries M, Gramoun A, et al. Ex situ evaluation of the composition of protein corona of intravenously injected superparamagnetic nanoparticles in rats. Nanoscale. 2014;6(19):11439-50.

205. Zhang T, Li G, Miao Y, Lu J, Gong N, Zhang Y, et al. Magnetothermal regulation of in vivo protein corona formation on magnetic nanoparticles for improved cancer nanotherapy. Biomaterials. 2021;276:121021.

206. Caracciolo G, Pozzi D, Capriotti AL, Cavaliere C, Piovesana S, La Barbera G, et al. The liposome-protein corona in mice and humans and its implications for in vivo delivery. J Mater Chem B. 2014;2(42):7419-28.

207. Bigdeli A, Palchetti S, Pozzi D, Hormozi-Nezhad MR, Baldelli Bombelli F, Caracciolo G, et al. Exploring Cellular Interactions of Liposomes Using Protein Corona Fingerprints and Physicochemical Properties. ACS Nano. 2016;10(3):3723-37.

208. Saha K, Rahimi M, Yazdani M, Kim ST, Moyano DF, Hou S, et al. Regulation of Macrophage Recognition through the Interplay of Nanoparticle Surface Functionality and Protein Corona. ACS Nano. 2016;10(4):4421-30.

209. Sha B, Gao W, Cui X, Wang L, Xu F. The potential health challenges of TiO2 nanomaterials. Journal of Applied Toxicology. 2015;35(10):1086-101.

210. Shi H, Magaye R, Castranova V, Zhao J. Titanium dioxide nanoparticles: a review of current toxicological data. Part Fibre Toxicol. 2013;10(1):15.

211. Pelclova D, Zdimal V, Fenclova Z, Vlckova S, Turci F, Corazzari I, et al. Markers of oxidative damage of nucleic acids and proteins among workers exposed to TiO2 (nano) particles. Occupational and Environmental Medicine. 2016;73(2):110-8.

212. Hussain S, Vanoirbeek JAJ, Hoet PHM. Interactions of nanomaterials with the immune system. Wiley Interdisciplinary Reviews Nanomedicine and Nanobiotechnology. 2012;4(2):169-83.

213. Survey USG. Mineral commodity summaries 2022. Mineral Commodity Summaries. 2022(2022).

214. Payne CK. A protein corona primer for physical chemists. Journal of Chemical Physics. 2019;151(13):130901.

215. Lynch I, Cedervall T, Lundqvist M, Cabaleiro-Lago C, Linse S, Dawson KA. The nanoparticle-protein complex as a biological entity; a complex fluids and surface science challenge for the 21st century. Adv Colloid Interface Sci. 2007;134-135:167-74.

216. Walczyk D, Bombelli FB, Monopoli MP, Lynch I, Dawson KA. What the Cell "Sees" in Bionanoscience. Journal of the American Chemical Society. 2010;132(16):5761-8.

217. Monopoli MP, Aberg C, Salvati A, Dawson KA. Biomolecular coronas provide the biological identity of nanosized materials. Nature Nanotechnology. 2012;7(12):779-86.

218. Kobos L, Shannahan J. Biocorona-induced modifications in engineered nanomaterial-cellular interactions impacting biomedical applications. Wiley Interdisciplinary Reviews Nanomedicine and Nanobiotechnology. 2020;12(3):e1608.

219. Wang Z, Hood ED, Nong J, Ding J, Marcos-Contreras OA, Glassman PM, et al. Combating Complement's Deleterious Effects on Nanomedicine by Conjugating Complement Regulatory Proteins to Nanoparticles. Advanced Materials. 2022;34(8):e2107070.

220. Schulze C, Schaefer UF, Ruge CA, Wohlleben W, Lehr C-M. Interaction of metal oxide nanoparticles with lung surfactant protein A. European Journal of Pharmaceutics and Biopharmaceutics. 2011;77(3):376-83.

221. Arango Duque G, Descoteaux A. Macrophage cytokines: involvement in immunity and infectious diseases. Front Immunol. 2014;5:491.

222. Hussain S, Johnson CG, Sciurba J, Meng X, Stober VP, Liu C, et al. TLR5 participates in the TLR4 receptor complex and promotes MyD88-dependent signaling in environmental lung injury. Elife. 2020;9.

223. Tighe RM, Birukova A, Yaeger MJ, Reece SW, Gowdy KM. Euthanasia- and Lavage-mediated Effects on Bronchoalveolar Measures of Lung Injury and Inflammation. Am J Respir Cell Mol Biol. 2018;59(2):257-66.

224. Bonner James C, Silva Rona M, Taylor Alexia J, Brown Jared M, Hilderbrand Susana C, Castranova V, et al. Interlaboratory Evaluation of Rodent Pulmonary Responses to Engineered Nanomaterials: The NIEHS Nano GO Consortium. Environmental Health Perspectives. 2013;121(6):676-82.

225. Xia T, Hamilton Raymond F, Bonner James C, Crandall Edward D, Elder A, Fazlollahi F, et al. Interlaboratory Evaluation of in Vitro Cytotoxicity and Inflammatory Responses to Engineered Nanomaterials: The NIEHS Nano GO Consortium. Environmental Health Perspectives. 2013;121(6):683-90.

226. Tighe RM, Birukova A, Yaeger MJ, Reece SW, Gowdy KM. Euthanasia- and Lavage-mediated Effects on Bronchoalveolar Measures of Lung Injury and Inflammation. American Journal of Respiratory Cell and Molecular Biology. 2018;59(2):257-66.

227. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature Biotechnology. 2008;26(12):1367-72.

228. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nature Protocols. 2016;11(12):2301-19.

229. cRAP protein sequences. https://www.thegpm.org/crap/index.html; 2012 2012.

230. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nature Methods. 2016;13(9):731-40.

231. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. 2019;47(D1):D442-d50.

232. Jayaram DT, Payne CK. Intracellular Generation of Superoxide by TiO2 Nanoparticles Decreases Histone Deacetylase 9 (HDAC9), an Epigenetic Modifier. Bioconjugate Chemistry. 2020;31(5):1354-61.

233. Martel J, Young D, Young A, Wu C-Y, Chen C-D, Yu J-S, et al. Comprehensive proteomic analysis of mineral nanoparticles derived from human body fluids and analyzed by liquid chromatography–tandem mass spectrometry. Analytical Biochemistry. 2011;418(1):111-25.

234. Runa S, Khanal D, Kemp ML, Payne CK. TiO2 Nanoparticles Alter the Expression of Peroxiredoxin Antioxidant Genes. Journal of Physical Chemistry C. 2016;120(37):20736-42.

235. Fleischer CC, Payne CK. Nanoparticle surface charge mediates the cellular receptors used by protein-nanoparticle complexes. Journal of Physical Chemistry B. 2012;116(30):8901-7.

236. Monopoli MP, Walczyk D, Campbell A, Elia G, Lynch I, Bombelli FB, et al. Physical-chemical aspects of protein corona: relevance to in vitro and in vivo biological impacts of nanoparticles. Journal of the American Chemical Society. 2011;133(8):2525-34.

237. Partikel K, Korte R, Mulac D, Humpf HU, Langer K. Serum type and concentration both affect the protein-corona composition of PLGA nanoparticles. Beilstein Journal Nanotechnolology. 2019;10:1002-15.

238. Grafe C, Weidner A, Luhe MV, Bergemann C, Schacher FH, Clement JH, et al. Intentional formation of a protein corona on nanoparticles: Serum concentration affects protein corona mass, surface charge, and nanoparticle-cell interaction. International Journal of Biochemistry & Cell Biology. 2016;75:196-202.

239. V. Courtney Broaddus RJM, Joel D. Ernst, Talmadge E. King Jr, Stephen C. Lazarus, John F. Murray, Jay A. Nadel, Arthur Slutsky. Murray and Nadel´s textbook of respiratory medicine. 2 / Ed.-in-chief: V. Courtney Broaddus. 6. ed ed. Philadelphia, Pa.: Elsevier Saunders; 2016 2016. 912 p.

240. Mukherjee AB, Zhang Z, Chilton BS. Uteroglobin: a steroid-inducible immunomodulatory protein that founded the Secretoglobin superfamily. Endocrine Reviews. 2007;28(7):707-25.

241. Ryan MA, Qi X, Serrano AG, Ikegami M, Perez-Gil J, Johansson J, et al. Mapping and Analysis of the Lytic and Fusogenic Domains of Surfactant Protein B. Biochemistry. 2005;44(3):861-72.

242. Hu G, Jiao B, Shi X, Valle RP, Fan Q, Zuo YY. Physicochemical properties of nanoparticles regulate translocation across pulmonary surfactant monolayer and formation of lipoprotein corona. ACS nano. 2013;7(12):10525-33.

243. Mandal AK, Zhang Z, Ray R, Choi MS, Chowdhury B, Pattabiraman N, et al. Uteroglobin represses allergen-induced inflammatory response by blocking PGD2 receptor-mediated functions. J Exp Med. 2004;199(10):1317-30.

244. Janicova A, Becker N, Xu B, Wutzler S, Vollrath JT, Hildebrand F, et al. Endogenous Uteroglobin as Intrinsic Anti-inflammatory Signal Modulates Monocyte and Macrophage Subsets Distribution Upon Sepsis Induced Lung Injury. Front Immunol. 2019;10:2276.

245. Zhao T, Su Z, Li Y, Zhang X, You Q. Chitinase-3 like-protein-1 function and its role in diseases. Signal Transduct Target Ther. 2020;5(1):201.

246. Lin T, Sammy F, Yang H, Thundivalappil S, Hellman J, Tracey KJ, et al. Identification of hemopexin as an anti-inflammatory factor that inhibits synergy of hemoglobin with HMGB1 in sterile and infectious inflammation. J Immunol. 2012;189(4):2017-22.

247. L. KJ, Moon C, Lee HS, Lee HW, Park EM, Kim HS, et al. Comparison of the biological activity between ultrafine and fine titanium dioxide particles in RAW 264.7 cells associated with oxidative stress. Journal of toxicology and environmental health Part A. 2008;71(8).

248. Miller TJ, Knapton A, Adeyemo OO, Noory LS, Weaver JL, Hanig JP, et al. Toxicology of Titanium Dioxide (TiO2) Nanoparticles: In vitro and in vivo evaluation of macrophage uptake of TiO2. The FASEB Journal. 2007;21(6):A812-A.

249. Thorley AJ, Ruenraroengsak P, Potter TE, Tetley TD. Critical Determinants of Uptake and Translocation of Nanoparticles by the Human Pulmonary Alveolar Epithelium. ACS Nano. 2014;8(11):11778-89.

250. Ruge CA, Kirch J, Cañadas O, Schneider M, Perez-Gil J, Schaefer UF, et al. Uptake of nanoparticles by alveolar macrophages is triggered by surfactant protein A. Nanomedicine. 2011;7(6):690-3.

251. Gakis GP, Aviziotis IG, Charitidis CA. Metal and metal oxide nanoparticle toxicity: moving towards a more holistic structure–activity approach. Environmental Science: Nano. 2023.

252. Ma Y, Shi J, Zhang Y, Chen Z, Jia G. Titanium Dioxide Nanoparticles Altered the lncRNA Expression Profile in Human Lung Cells. International Journal of Environmental Research and Public Health. 2023;20(2).

253. Tredano M, van Elburg RM, Kaspers AG, Zimmermann LJ, Houdayer C, Aymard P, et al. Compound SFTPB 1549C-->GAA (121ins2) and 457delC heterozygosity in severe congenital lung disease and surfactant protein B (SP-B) deficiency. Human Mutation. 1999;14(6):502-9.

254. Olenick LL, Troiano JM, Vartanian A, Melby ES, Mensch AC, Zhang L, et al. Lipid Corona Formation from Nanoparticle Interactions with Bilayers. Chem. 2018;4(11):2709-23.

255. Zhang X, Pandiakumar AK, Hamers RJ, Murphy CJ. Quantification of Lipid Corona Formation on Colloidal Nanoparticles from Lipid Vesicles. Analytical Chemistry. 2018;90(24):14387-94.

256. G. LB, Capriotti AL, Caracciolo G, Cavaliere C, Cerrato A, Montone CM, et al. A comprehensive analysis of liposomal biomolecular corona upon human plasma incubation: The evolution towards the lipid corona. Talanta. 2020;209.

257. T. L, Bernfur K, Vilanova M, Cedervall. Understanding the Lipid and Protein Corona Formation on Different Sized Polymeric Nanoparticles. Scientific reports. 2020;10(1).

258. Hu Q, Bai X, Hu G, Zuo YY. Unveiling the Molecular Structure of Pulmonary Surfactant Corona on Nanoparticles. ACS Nano. 2017;11(7):6832-42.

259. Deng H, Li X, Peng Q, Wang X, Chen J, Li Y. Monodisperse magnetic single-crystal ferrite microspheres. Angew Chem Int Ed Engl. 2005;44(18):2782-5.

260. Duff DG, Baiker A, Edwards PP. A new hydrosol of gold clusters. 1. Formation and particle size variation. Langmuir. 1993;9(9):2301-9.

261. Poulsen KM, Payne CK. Concentration and composition of the protein corona as a function of incubation time and serum concentration: an automated approach to the protein corona. Anal Bioanal Chem. 2022;414(24):7265-75.

262. UniProt C. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res. 2023;51(D1):D523-D31.

263. Hoie MH, Kiehl EN, Petersen B, Nielsen M, Winther O, Nielsen H, et al. NetSurfP-3.0: accurate and fast prediction of protein structural features by protein language models and deep learning. Nucleic Acids Res. 2022;50(W1):W510-W5.

264. Klausen MS, Jespersen MC, Nielsen H, Jensen KK, Jurtz VI, Sonderby CK, et al. NetSurfP-2.0: Improved prediction of protein structural features by integrated deep learning. Proteins. 2019;87(6):520-7.

265. Cock PJ, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics. 2009;25(11):1422-3.

266. Krick T, Verstraete N, Alonso LG, Shub DA, Ferreiro DU, Shub M, et al. Amino Acid Metabolism Conflicts with Protein Diversity. Molecular Biology and Evolution. 2014;31(11):2905-12.

Biography

Karsten Poulsen earned a Bachelor of Engineering in Materials Science along with a Bachelor of Science in Chemistry from the University of Minnesota Twin Cities in 2017. Afterwards he worked as a pilot plant technician at Cargill and then as a quality engineer at Kluge Design Incorporated. After his initial stint in industry, he joined the Payne Lab at Duke University. During his PhD he could often be found working on his house in Durham with his spouse Maren Poulsen and dog Nalu. Starting his fourth year he became a dad when Nora Poulsen joined the family. In his work at Duke, Karsten has worked on automating protein corona formation, purification, and characterization (chapter 2 and chapter 3). Along with this work he collaborated with the Tighe lab to understand how a lung fluid protein corona can affect the cellular response (chapter 4). Most recently his time has been spent working on predicting the formation of protein coronas using machine learning (Appendix A).