UNIVERSITY OF CALIFORNIA
RIVERSIDE

Fluorescence Labeling and Limited Proteolysis for Demystifying Protein Corona

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

Doctor of Philosophy

in

Chemistry

by

Yaokai Duan

September 2018

Dissertation Committee:
Dr. Wenwan Zhong, Chairperson
Dr. Yinsheng Wang
Dr. Quan Cheng

The Dissertation of Yaokai Duan is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

Personal acknowledgement

Firstly and most importantly, I would like to give my acknowledgement to my advisor Prof. Wenwan Zhong, for all her inspiration, insights, encouragement, mentorship and support. As a newbie for analytical chemistry five years ago, I could not survive graduate school without her patience, trust, and advices.

I would also like to acknowledge and thank the other members in my committee: Prof. Yinsheng Wang and Prof. Quan Cheng, for their time, suggestions, and assistances.

I would like to acknowledge and thank Prof. Huiwang Ai, Prof. Richard Hooley, and Prof. Jikui Song, for serving on my oral exam committee, and providing many valuable comments and suggestions.

I would like to thank Department of Chemistry, faculty and staff at UCR, for their support. I also wish to thank my labmates, classmates, and friends, for their help and accompany. Finally, I would like to acknowledge and thank my family, especially my parents, for their dedicated love, understanding, encouragement, and support.

Copyright acknowledgement

The text and figures in chapter 2, in part or full, are reprinted with permission from the: Ashby, J.; Duan, Y.; Ligans, E.; Tamsi, M.; Zhong, W. High-Throughput Profiling of Nanoparticle–Protein Interactions by Fluorescamine Labeling. Analytical chemistry 2015, 87, 2213-2219. The co-author listed in that publication, Prof. Wenwan Zhong, provided guidance and editing for the work.

The text and figures in chapter 3, in part or full, are reprinted with permission from the: Duan, Y.; Liu, Y.; Shen, W.; Zhong, W. Fluorescamine Labeling for Assessment of Protein Conformational Change and Binding Affinity in Protein–Nanoparticle Interaction. Analytical chemistry 2017, 89, 12160-12167. The co-author listed in that publication, Prof. Wenwan Zhong, provided guidance and editing for the work.

ABSTRACT OF THE DISSERTATION

Fluorescence Labeling and Limited Proteolysis for Demystifying Protein Corona

by

Yaokai Duan

Doctor of Philosophy, Graduate Program in Chemistry
University of California, Riverside, September 2018
Dr. Wenwan Zhong, Chairperson

The emerging production and usage of engineered nanomaterials (ENMs) make them exposed more and more to biological systems, including human beings. However, the adsorbed layer of biomolecules, i.e. biocorona, outside of the ENMs present in any biofluid complicates the study of the biological impacts induced by exposure to the ENMs, calling for more understanding to be obtained about the formation and functions of the biocorona. Within the biocorona, proteins occupy a major portion because of their high abundance in most of the biofluids, like plasma, attracting tremendous research efforts in the past decade to study protein corona. It has been reported that, protein corona could disguise the native properties of the ENMs and present as a new "biological identity" of the ENMs to influence their cell recognition and other *in vivo* activities. With great endeavors devoted to understand the formation and biological impacts of the protein corona, there is still knowledge gap on how the intrinsic properties of ENMs affect corona formation, and little is known on the impacts on the proteins in the corona, including the orientation of the proteins on ENM surface and any conformational changes they

experience. Due to the non-specific forces behind protein-ENMs interactions, desired and controlled arrangement of proteins on ENMs surface has always been a significant but difficult task. Gaining detailed knowledge on the changes occurred to the proteins in ENM corona then could provide insights to help achieve this goal, calling for rapid and routine methods for corona composition analysis and nano-bio interface exploration of protein corona at the molecular level. My research thus focuses on development of new methods to gain more understanding on protein corona formation. Firstly, a fluorescamine labeling method is applied to screen interactions between individual proteins and ENMs, which could reveal both the presence of protein-ENM interaction and any conformational change in proteins induced by ENMs. Then, we identify the correlation between the labeling data and protein corona composition, which suggests that a structure activity quantification model using those descriptors could be built for prediction of corona formation in biofluids. Secondly, the method of limited proteolysis coupled with LC-MS/MS is developed which is capable of identifying the binding sites of the protein by a macromolecule like synthetic receptor and nanoparticle. This method is applied to study the proteins showing positive results in fluorescamine labeling and the impacts on the proteins including their orientation and conformation are revealed. The methods reported here aim to obtain rapid knowledge about protein binding and corona formation. Still, a comprehensive view of the molecular details at the ENM-protein interface require the employment of complimentary analytical techniques, which will be explored in our future works. Overall, the knowledge gained from such studies can guide rational design of ENMs with higher safety and sustainability.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: Introduction

The rational design of engineered nanomaterials (ENMs) has become one practical methodology and thriving research area, to meet the strict requirements for biological applications and environmental health and safety.[1,2] The success of these endeavors has facilitated the development of ENMs with improved safety to serve as nanomedicines and drug delivery vesicles, since the first of their kind was approved by FDA 23 years ago.[3-6] However, challenges about their smooth transition to the biomedical fields and debates about their biosafety have emerged along with the rapid increase of their implementations, both demanding more understanding of how the ENMs behave in biological systems. One unavoidable aspect in the study of the nano-bio interface is to understand how the biocorona is formed, which represents the layer of biological molecules coating the ENMs present in every biological matrix. It is believed that the biocorona establishes a new identity of ENMs masking the intrinsic properties of ENMs and being "seen" by the cells.[7,8] Compared to the physicochemical properties of ENMs, which can be controlled by advanced nanotechnology and measured precisely by sophisticated spectroscopic and imaging techniques, the biocorona is more difficult to manipulate and evaluate, due to the complexity of the biological surroundings ENMs will encounter.[4,8-11] There are diverse types of biomolecules in the biofluid, including proteins, nucleic acids, lipids and metabolites, each competing with thousands of others for the limited surface of the ENM and yielding a highly dynamic and complex biocorona.[7,12] Development of effective methods to unveil the composition and arrangement of biomolecules in the biocorona is

greatly desired to move forward the biological applications of ENMs, but not yet fulfilled despite lots of progress made recently.[12,13]

1.1: Understand protein corona of ENMs

Because of the importance of their biological functions and high abundance in biofluids, protein is considered the major component of the biocorona, and thus enormous research efforts have been devoted to study the protein corona since the concept was proposed a decade ago.[14-16]

Evolution of protein corona. Originally the focus of protein corona study was on the binding equilibrium and kinetics between the protein and ENM.[11,14] Later, physical separation relying on centrifugation and columns to effectively isolate the proteins adsorbed on ENMs from the free ones in solution, have been coupled with LC-MS/MS to facilitate routine identification of the protein composition in corona. Then, more understanding on how the corona is evolved has been obtained. It has been reported in numerous studies that, protein corona is highly dynamic, i.e. its composition is highly time and matrix dependent,[17,18] and "soft" or "hard" corona has been named based on the time evolution of the corona and the binding strength of the corona proteins, or the Vroman effect.[18] Proteins with fast binding rates will mainly contribute to the "soft" corona that is formed at short incubation duration (e.g. 0.5 min). Because the binding rate is determined by both the rate constant and the protein concentration, the "soft" corona is typically comprised of the highly abundant proteins in the matrix. With increasing incubation duration, proteins with lower abundance but higher affinities will replace the ones with lower affinities in the soft corona, establishing the hard corona. Corona formation is highly

responsive to the surrounding: the composition of both the "soft" and "hard" corona will change and new binding equilibrium will be established when the ENMs are transported among different biological matrices.[17] Moreover, physical forces experienced by the ENMs, like that present in the circulation system, can also influence the corona formation, which might be responsible for differences observed *in vivo* and *in vitro*.[19] The dynamic nature of protein corona and the high complexity of the biological environment present great challenges in the study of corona formation and its impacts on the fate of ENMs.

Biological outcomes of protein corona. Protein corona is thought to contribute to the biological behaviors of ENMs, because it is the forefront during the interaction of ENMs with cells and tissues. One significant outcome of protein corona is to elongate the circulating time of ENMs used for drug delivery, owing to the enhanced colloidal stability and biocompatibility. Aggregation is a common phenomenon occurred to the ENMs in the high salt physiological environment, if there is no hydrophilic ligand protecting their surface. Aggregation will enlarge the sizes of ENMs, making them easier to be removed by macrophages.[20] But protein corona could form a hydrophilic and charged surface on the ENMs, preventing them from aggregation and improving their distribution in biosystems. Moreover, protein corona could disguise the exogenous ENMs by the endogenous proteins, which reduces the uptake of ENMs by immune cells. Both could greatly help nanomedicines or ENM-based drug carriers to be stable in the circulation systems until reaching their target cells.

However, denaturation and aggregation may happen for the proteins in the corona.[21] For example, conformational change and decrease of stability were reported for the tubulin

adsorbed on the titanium oxide NPs, which prevented the polymerization of tubulin and disrupted the function of microtubule.[22,23] It was also reported that serum albumin could be denatured by the gold NPs;[24] and fibrinogen could be unfolded on the gold NPs coated by poly(acrylic) acid.[25,26] Such denaturation and conformational changes of proteins on ENMs could induce numerous adverse impacts. For instance, the denaturation of serum albumin could expose the cryptic epitopes that can be recognized by immune cells, increasing immunogenicity.[16,27] Fibrinogen unfolding could trigger NF-B signaling pathway directed by Mac-1, one integrin receptor for macrophage, which may induce inflammation.[28] In some circumstances, protein unfolding or fibrillation induced by ENMs could be beneficial. For instance, fibrillation inhibition could be useful for the cure of Alzheimer's and Parkinson's diseases. While some reports attribute such effects to the surface properties of the ENMs, how exactly the intrinsic properties of the ENMs induce the conformational change of proteins is not entirely clear yet, and it demands further study.

Protein corona has also been found to be able to reduce the cytotoxicity that causes cell death via necrosis, apoptosis, and proptosis. This mitigation effect has been observed for diverse ENMs. In particular, the ENMs with metal core, such as CuO, ZnO, and silver, usually have high cell toxicity, due to their interactions with the negatively charged lipids membrane, the release of metal ions, the induction of oxidative stress, and the peroxidation of lipids.[29] Protein corona has been shown able to minimize all these effects. Preincubating ENMs with serum to form protein corona could reduce the production of reactive oxygen species (ROS) by almost half, when exposing ENMs to cells.[30] Surface coating by individual proteins like human serum albumin (HSA), bovine serum albumin (BSA), and

transferrin, or by the biological matrix like fetal bovine serum (FBS), mouse or human serum, could increase the colloidal stability of ENMs and decrease the ion dissolution.[31] The antibacterial activity of silver NPs that is due to the release of silver ions, can be completely inhibited by coating the NPs with serum or plasma proteins.[32] The damage on lipid membrane induced by the strong positive charges on the surface of ENMs can also be mitigated by the formation of protein corona that changes the surface zeta potential.[33] Protein corona can also reduce the hazardous effects of two-dimensional (2D) ENMs, e.g. graphene and graphene oxide (GO). Graphene and GO both exhibit dose dependent cell toxicity, which could be alleviated by the protein corona formed in FBS or cell culture media.[34] However, protein corona cannot alleviate the cytotoxicity of all ENMs. The silver NPs coated with citrates and glucose still showed high ROS production and cytotoxicity, even after the formation of protein corona on the NP surface.[32] The different effect of protein corona might be caused by the its different composition, which is closely dependent on the intrinsic properties of ENMs.

The impacts of protein corona are not limited on the mitigation of ENMs' cytotoxicity, which only occurred at the high dose of ENMs. Instead, there are more interests on protein corona to modulate the cellular interactions and activities of ENMs at lower concentrations. One example is the active targeting of ENMs, since one of the most important, and also the very first clinical application of ENMs has been the drug delivery tool.[3,35] Although the passive targeting to cancer cells could be achieved through enhanced permeability and retention (EPR) effect,[36] protein corona would improve the biocompatibility and stability of ENMs by covering of intrinsic proteins, which will

increase the pharmacokinetics and target accumulations of nanomedicines.[37,38] Moreover, the capability of ENMs to penetrate biological barriers, such as the blood brain barrier (BBB), can be enhanced by protein corona.[2,39] Often, this passive targeting potential can be further enhanced by the active targeting schemes[40-42] induced by specific molecular interactions, like ligand-receptor, antibody-antigen, aptamer-target interaction.[43-45] For example, the gold NPs coated with transferrin can actively target solid tumors that overexpress transferrin receptor on the cell membranes.[46] However, protein corona could have unexpected impacts on active targeting. It has been reported that the transferrin-modified silica NPs lost their targeting ability after incubation in serum, due to the formation of serum protein corona. The protein corona can contain multiple layers of proteins,[47] which could block the desired interactions between the ligands on ENM surface and the receptors on cell surface.[48] Again, depending on the types of ENMs, different effects of protein corona on active targeting have been reported, e.g. no influence from the protein corona on the targeting capability of the antibody functionalized NPs towards colon cancer cells has been observed.[49]

Another example of protein corona' impact on ENMs is the modulation of immune response. Since protein corona could influence the cell recognition of ENMs, it will help ENMs to escape from the immune cells, including monocytes and macrophage cells. It has been reported that, the adsorption and internalization of the porous polymer NPs by monocytic cells can be dramatically inhibited by the serum albumin corona, but there was no influence of this corona on the uptake of NPs by macrophages. Instead, more uptakes by macrophages were observed during short time incubation (i.e. 1hr).[4,50] At the same time,

the unfolding of serum albumin in corona could induce the release of the inflammatory cytokines and the activation of phagocytosis. A similar phenomenon was also observed for the silica NP, showing high macrophage internalization mediated.[51] Besides serum albumin, complement components that has been proved to be one of the most abundant components in the serum protein corona, can also help to trigger the opsonization uptake of ENMs, by activating the complement process.[52,53] Based on this phenomenon, the surface properties of ENMs can be modified and therefore the composition of protein corona could be modulated, to control the activation of the complement pathway.[54] The conformational change of protein in corona also present cryptic epitopes to the surface as a danger signal to alarm immune system, which will induce adaptive immunity response and scavenger of ENMs.[16,55,56] Modulation of immune response of ENMs has become a very active research area to facilitate the recognition of cancer cells by immune systems, especially after the success of immune therapy on cancer treatment.[57,58]

1.2: Predict compositions of protein corona

Due to the critical role of protein corona on determining biological outcomes of ENMs, it is very important to get the detailed information about the composition of protein corona. However, the high complexity of biological systems present great challenges to detailed characterization of protein corona for each ENM which is far behind the rapid growing development of diverse types of ENMs. The composition of protein corona usually can be obtained through LC-MS/MS based proteomics study, after separating the adsorbed proteins with ENMs from the free proteins, mostly via centrifugation. Despite that tremendous works have been done recently in this area, the whole procedure is not

trivial at all. Thus, the knowledge on how the properties of ENMs and proteins can govern the adsorption of proteins on ENMs becomes essential for the prediction of corona formation in diverse biological matrices, which can accelerate the biological applications of ENMs.

Properties of ENMs on protein corona composition. It is intuitive to conclude the preference of protein corona composition based on physicochemical characteristics of ENMs, including size, shape, charge, hydrophobicity, and aspect ratio. By keeping most of parameters constant and only changing one at a time, the correlation between the individual parameter and the protein it preferred could be concluded based on the changes in the protein corona composition. Based on it, there have been many conclusions drawn in literatures. For two polystyrene (PS) NPs with 50- or 100-nm diameter, more apolipoprotein B-100 was identified in the serum corona of the larger NPs, which indicated the preference to low density lipoprotein (LDL) or very low density lipoprotein (VLDL) by the bigger PS NPs.[59] The difference on size also resulted in the variation in surface curvature, which influences the binding stability of proteins and requires different degrees of protein conformational change to accommodate the surface. As a result, more unfolding of serum albumin was observed on the smaller gold NPs.[60]

Besides size, surface charge is also important for protein corona composition. Apolipoproteins, in general, were less adsorbed on the positively charged PS NPs, while more immunoglobulin were shown in their corona.[27] Since most of the serum proteins are negatively charged, ENMs with positive surface charge can have higher binding potentials.[61] Meanwhile, more dramatic protein conformational change has been seen on

the positive charged ENMs;[9] and more protein binding could occur on the ENMs with higher surface hydrophobicity.[62] The composition of protein corona was different for the hydrophobic and hydrophilic ENMs.[63]

Moreover, the shape or aspect ratio of the ENMs can also influence the protein corona, as higher protein density was observed on the surface of the gold nanorods, comparing to the gold nanospheres with similar surface coatings.[64] Besides, more proteins in the corona of the gold nanorods were found with the loss of the secondary structures, as well as the protein functions.[64]

Although the progress shown above are valuable to illustrate the most important factors of ENMs for protein corona composition, the results are not quantitative, and some of them are even controversial. As a result, it is unreliable to use these correlations to predict protein corona compositions for other ENMs.

Structure-activity quantification (SAR) modeling and molecular simulation. As a complementary method, computational technology has been applied in many areas to accelerate research. Aspired by its success in the drug development, molecular dynamic (MD) simulation has been used to simulate the formation and time evolution of protein corona.[65,66] Thanks to the well-developed nanotechnology and characterization tools, many details of ENMs could be modeled. However, the complexity of proteins remains as a big issue. It is very challenging and impossible to include all possible corona proteins in one simulation, since the computational power is always the bottleneck.

SAR provides another rapid and robust paradigm to model the protein corona. One foundation for SAR is the molecular parameter, or the descriptor, for either the molecule

of interested or the target. After building the quantitative model to correlate the molecule and the target, the effect of a new molecule on the target could be predicted by fitting its descriptor into the model. In the scenario of protein corona composition, the physicochemical properties of ENMs can be used as their descriptors, and the composition of protein corona will be the target of prediction. Surprisingly, there is very limited practice using SAR model to predict protein corona. Only one paper of its kind developed one model to predicte whether one protein was in the corona of silver nanoparticles. The quantitative value of protein corona composition was never been predicted. In contrast, SAR has many success on predicting biological outcomes of ENMs, using either the physicochemical properties or the protein corona composition of ENMs as descriptors.[67-71] More work can be done in this area.

Although the physicochemical properties of ENMs are easy to measure, the relationships between them and protein corona composition are difficult to interpret. As shown in the previous section, the effects of these parameters were sometimes controversial. One reason is that the interaction between ENMs and proteins combines multiple forces crossing the 3-dimensional space, and it is not enough to take only one single factor in the consideration. To overcome this limit, individual proteins have been used as the unit probes to characterize the interactions with ENMs.[72] The 3-D structure and complexity of protein make the characterized binding behavior is the combination of all forces in vicinity. One drawback of this method is its throughput. The speed to measure the affinity of individual protein-ENM pairs is far slower than characterization of ENMs.

The development of a high throughput screening method to rapidly quantify the interactions between proteins and ENMs is necessary.

1.3: Probe the molecular details of protein corona

The biological importance of protein corona is dependent not only on its diverse composition, but also on the molecular orientation of proteins on the surface of ENMs. The adsorption of proteins on ENMs is mostly based on nonspecific interactions, which would allow proteins to show multiple orientations. However, the epitopes of proteins need to be exposed to the solvent to play their functions. Moreover, considering that the protein corona could compose of multiple layers, the position of the protein in the corona can also contribute to the biological identity of ENMs. Besides, the unfolding and aggregation of protein is quite common in the corona of the ENMs, and the newly exposed cryptic epitopes could induce new interactions with cells. Thus, methods for probing the molecular topology of the protein corona will help illustrate biological significance of protein corona, and it will improve the rational design of ENMs for biological applications.

Protein conformational change and aggregation in corona. The most widely used method to analyze protein folding is circular dichroism (CD).[73] The ratio of secondary structures and conformational change of proteins can be calculated from the CD spectra. However, CD is only good for the single protein; and ENMs with high absorption at UV region could interfere with the detection. Thermal stability test is another common way to show small conformational change of proteins, but its application for ENMs is not popular. One application used CD as the detection tool to monitor conformational changes of protein under different temperatures or urea concentrations, showing that the thermal stability of

serum albumin could be compromised by ENMs.[62,74] Another successful case used isothermal scanning calorimetry (ISC) to study the binding of fibrinogen to titanium oxide NPs, but the high concentration of the protein (10 mg/ml) is requited for ISC to show discriminable signal, which might not be realistic for other perishable proteins.[23] A third case used fluorescence dye, e.g. thioflavin T, to study the aggregation or fibrillation of fibrinogen induced by ENMs.[75-77]. Besides the nonspecific labeling dyes, intrinsic fluorescence from aromatic residues can also be monitored, to characterize the binding and unfolding of proteins in the corona of ENMs.[78] One drawback of these fluorescence based methods is the complicated optical influence from ENMs, including fluorescence quenching, enhancement, and the inner filter effect. Thus, careful experimental design and data interpretation are needed for those fluorescence methods. A fourth case used electrical chemistry to study the unfolding of protein induced by ENMs. Binding of the unfolded protein on ENMs can induce a shift of formal potential that can be measured and quantified, which was demonstrated on cytochrome c with the titanium oxide NPs.[79]

Epitope mapping of protein corona. Since the functional motif of protein only accounts for a small part of the whole structure, "Yes" or "No" for describing the conformational change of protein in corona is not enough, and more molecular details are needed to characterize the conditions of these functional motifs, or epitopes. Protein epitopes are usually exposed to the solvent, which is critical for their recognition and interaction with other biological molecules. Since ENMs could bind and block parts of the protein surface, it is very important to evaluate the arrangement and orientation of proteins on ENMs. Antibody conjugated gold NPs has been used to map out the orientation of

transferrin on the surface of PS NPs, and the result showed only a very small portion of transferrin would orientate correctly for the recognition of the transferrin receptor.[13,51] This method is largely limited by the availability of antibodies. As an alternative, cells, e.g. macrophage, could be used to replace antibodies in this method, since cells would express various kinds of receptor. Its success application to verify the unfolding of serum proteins on silica NPs has been reported, while the recognition process was still remined as a black box and difficult to interpret.[51] Most importantly, the information obtained in these testes was still limited on the molecular details. Among all technologies, X-ray crystallography and NMR could provide the most detailed and accurate information for the protein structure, but they are not practical for proteins on ENMs. Mass spectrometer (MS), on the contrary, has shown successful applications to explore the binding sites of protein on ENMs, as well the protein conformational change induced by ENMs, with the help of cross-linking, isotope exchange, or chemical labeling.[80-82] However, these reagents could not coverage the whole protein sequence due to their limited reactivity, and the protein conformation would also be interrupted during the labeling. To overcome these limitations, limited proteolysis has been used as an alternative tool to map out the structure of protein, because of the high reactivity, specificity, and compatibility of proteases. The binding of proteins to ENMs could prevent proteases from accessing the binding sites. Within short time, only the part of protein exposed to the solvent could be digested. After the generated peptides being identified in LC-MS/MS, the digestion efficiency crossing the whole protein sequence could be mapped out. Limited proteolysis coupled with LC-MS/MS has shown

its success on the study of ligand-protein binding and protein-protein interactions, but there is no application on protein-ENMs interactions yet.[83,84]

1.4: Dissertation scope

The focus of this dissertation is to develop routine and robust analytical methods applicable for the study of protein corona. The protein composition and protein molecular details will be the two aspects of protein corona study covered in this thesis. Two main problems will be addressed: 1) the deficiency of rapid experimental characterizations for protein-ENMs interactions, and 2) the lack of tools to disclose molecular details of proteins in corona.

Chapter 2 starts with the proof-of-concept of a high-throughput assay for screening interactions between proteins and ENMs. Using the fluorogenic dye, fluorescamine, the lysine residues on the protein surface will be labeled and quantified. The interaction between the protein and ENMs would have two opposite effects on the number of surface lysine residues: 1) the blockage of the binding interface by ENMs could reduce the number, while 2) the unfolding of protein induced by ENMs would expose more lysine residues to the surface and increase the number. Correspondingly, either diminished or enhanced fluorescence will be observed after fluorescamine labeling. Standard proteins, including HSA and transferrin, are incubated with silica NPs and PS NPs with three different sizes. The unfolding of proteins upon binding to PS NPs is verified by CD spectra. Moreover, the fluorescence profile of each NP shows the characteristic pattern, which can be used to differentiate different ENMs.

In chapter 3, we apply the fluorescamine labeling method to screen a larger set of proteins with various properties. The fluorescence profiles obtained could not only differentiate ENMs, but also cluster proteins into two groups. One of the groups includes proteins with flexible structures and cation binding abilities resulting in higher fluorescence changes. This indicates that flexible and cation bound proteins are easier to unfold upon binding to negatively charged ENMs, e.g. carboxylated PS NPs. Besides, this fluorescamine labeling method could be used to measure the binding affinities of different proteins toward PS NPs, by fitting the fluorescence responses to the Hill's equation.

Chapter 4 focuses on the prediction model for protein corona composition based on SAR. The fluorescence profiles obtained in the fluorescamine labeling screening are used as novel descriptors for ENMs. 19 ENMs that could be found in real life are used for the test. Due to the week and transient interactions for most protein-ENMs pairs, heating pressure is adopted to enlarge the fluorescence change in the fluorescamine labeling, which is benefited from the protein thermal stability shift induced by ENMs. The impact of ENMs on protein thermal stabilities can be either stabilization or destabilization, depending the balance of protein unfolding and binding. Moreover, high correlations between fluorescamine labeling profiles and the protein corona compositions were observed for some proteins in the serum protein corona of ENMs. Encouraged by these correlations, two prediction models using fluorescamine labeling profiles as descriptors for ENMs are successfully built. These two models, based on either classification or regression, could provide accurate predictions of which protein could be enriched in the serum protein corona. At the last, the prediction for the cytotoxicity of ENMs on CEM cells are also

explored using the fluorescence profiles as descriptors of ENMs, and rough classification could be achieved.

Chapter 5 shifts to the method development of limited proteolysis to study the binding of synthetic receptors to proteins. Although the ultimate goal of this method will be to explore the molecular details of protein in the corona of ENMs, synthetic receptors are chosen to verify the practicability of the method firstly, due to the similar interactions between these two systems and multiple crystal structures of protein/synthetic receptors complex available. Two different types of synthetic receptors, e.g. cucurbit[7]uril (CB7) and 4-sulfonatocalix[4]arene (sclx4), are used for this study. The amino acids targeted by CB7 and sclx4 are different: CB7 prefers aromatic amino acids, but sclx4 binds to positively charged ones. By choosing specific protease that can cut corresponding amino acids, the binding sites of CB7 on insulin and sclx4 on lysozyme can be identified by the limited proteolysis method. Moreover, the nonspecific protease, such as proteinase K, can also be used for synthetic receptors with unknown binding preferences, which is verified by the successful identification of binding sites of sclx4 on lysozyme. One possible difficult of this method could be the formation of protein oligomers induced by synthetic receptors, which would also block the protein surface and change the digestion result. To minimize the interference of oligomer formation, on surface limited proteolysis is performed, in which proteins will be firstly immobilized on the hydrophilic surface by nonspecific adsorption. Thus, less oligomerization will not happen in the following process.

After verifying its practicability, limited proteolysis is applied to explore the molecular details of proteins in the corona of ENMs, in chapter 6. The proteins shown

higher fluorescence changes in previous screening, e.g. transferrin and catalase, are used for the study. Consistent with their enhanced fluorescence signals indicating unfolding, the increased digestion efficiency of both transferrin and catalase induced by PS NPs are observed. The increase is caused by the unfolding of protein, but not by the change of trypsin activity, since no obvious enhancement is observed for HAS with PS NPs. Peptides generated in the limited proteolysis will be separated and identified by LC-MS/MS. Afterwards, the digestion efficiency of each cutting site could be quantified by the relative abundance (RA) of each peptide. The change of RA indicates either the blockage (shown as decreased RA) or the unfolding of protein (shown as increased RA), which could help to map the orientation and conformational change of proteins in the corona of ENMs. Multiple binding sites are found on the same protein, which could induce the aggregation of ENMs by the bridging effect. Moreover, another ENMs-protein pair: Tf/apoTf with iron oxide NP (IONP) is also tested, because it shows decreased fluorescence signal in the fluorescamine labeling, different from the enhanced signal of Tf and CAT with PS NPs. As expected, lower digestion efficiency is observed for Tf/apo-Tf incubated with IONP. The limited digestion at a higher temperature is used to enlarge the difference induced by IONP. After identifying the digestion efficiency of each cutting site, it is interesting that the binding sites of Tf/apoTf on IONP are different from these of Tf on PS NPs. This limited proteolysis method can be used as a routine tool to evaluate the orientation and conformational change for proteins in the corona of ENMs, to facilitate the rational design of ENMs functionalized by proteins.

At the last, chapter 7 will conclude the results undertaken in this thesis and discuss future plans. Despite their good performance, the methods developed in this work have plenty of room to improve. New descriptors could be incorporated to improve the prediction ability of the SAR model, because the biological matrix contains various kinds of molecules other than proteins. The interactions between ENMs and nucleic acids, lipids, and metabolites can be screened and used as the new descriptors of ENMs. As for the limited digestion method, it will be benefited from the fixation of proteins in the corona of ENMs before the proteolysis, since there will be less exchanges between the original protein and generated peptides during the digestion.

References

(1) Bobo, D.; Robinson, K. J.; Islam, J.; Thurecht, K. J.; Corrie, S. R. Nanoparticle-based medicines: a review of FDA-approved materials and clinical trials to date. *Pharmaceutical research* **2016**, *33*, 2373-2387.

(2) Blanco, E.; Shen, H.; Ferrari, M. Principles of nanoparticle design for overcoming biological barriers to drug delivery. *Nature biotechnology* **2015**, *33*, 941.

(3) Shi, J.; Kantoff, P. W.; Wooster, R.; Farokhzad, O. C. Cancer nanomedicine: progress, challenges and opportunities. *Nature Reviews Cancer* **2017**, *17*, 20.

(4) Lee, Y. K.; Choi, E.-J.; Webster, T. J.; Kim, S.-H.; Khang, D. Effect of the protein corona on nanoparticles for modulating cytotoxicity and immunotoxicity. *International journal of nanomedicine* **2015**, *10*, 97.

(5) Slowing, I. I.; Vivero-Escoto, J. L.; Wu, C.-W.; Lin, V. S.-Y. Mesoporous silica nanoparticles as controlled release drug delivery and gene transfection carriers. *Advanced drug delivery reviews* **2008**, *60*, 1278-1288.

(6) Ai, J.; Biazar, E.; Jafarpour, M.; Montazeri, M.; Majdi, A.; Aminifard, S.; Zafari, M.; Akbari, H. R.; Rad, H. G. Nanotoxicology and nanoparticle safety in biomedical designs. *International journal of nanomedicine* **2011**, *6*, 1117.

(7) Lynch, I.; Salvati, A.; Dawson, K. A. Protein-nanoparticle interactions: what does the cell see? *Nature nanotechnology* **2009**, *4*, 546.

(8) Lesniak, A.; Fenaroli, F.; Monopoli, M. P.; Åberg, C.; Dawson, K. A.; Salvati, A. Effects of the presence or absence of a protein corona on silica nanoparticle uptake and impact on cells. *ACS nano* **2012**, *6*, 5845-5857.

(9) Van Hong Nguyen, B.-J. L. Protein corona: a new approach for nanomedicine design. *International journal of nanomedicine* **2017**, *12*, 3137.

(10) Liu, R.; Jiang, W.; Walkey, C. D.; Chan, W. C.; Cohen, Y. Prediction of nanoparticles-cell association based on corona proteins and physicochemical properties. *Nanoscale* **2015**, *7*, 9664-9675.

(11) Monopoli, M. P.; Åberg, C.; Salvati, A.; Dawson, K. A. Biomolecular coronas provide the biological identity of nanosized materials. *Nature nanotechnology* **2012**, *7*, 779.

(12) Caracciolo, G.; Farokhzad, O. C.; Mahmoudi, M. Biological identity of nanoparticles in vivo: clinical implications of the protein corona. *Trends in biotechnology* **2017**, *35*, 257-264.

(13) Kelly, P. M.; Åberg, C.; Polo, E.; O'connell, A.; Cookman, J.; Fallon, J.; Krpetić, Ž.; Dawson, K. A. Mapping protein binding sites on the biomolecular corona of nanoparticles. *Nature nanotechnology* **2015**, *10*, 472.

(14) Cedervall, T.; Lynch, I.; Lindman, S.; Berggård, T.; Thulin, E.; Nilsson, H.; Dawson, K. A.; Linse, S. Understanding the nanoparticle–protein corona using methods to quantify exchange rates and affinities of proteins for nanoparticles. *Proceedings of the National Academy of Sciences* **2007**, *104*, 2050-2055.

(15) Docter, D.; Westmeier, D.; Markiewicz, M.; Stolte, S.; Knauer, S.; Stauber, R. The nanoparticle biomolecule corona: lessons learned–challenge accepted? *Chemical Society Reviews* **2015**, *44*, 6094-6121.

(16) Lynch, I.; Dawson, K. A.; Linse, S. Detecting cryptic epitopes created by nanoparticles. *Sci. Stke* **2006**, *2006*, pe14-pe14.

(17) Lundqvist, M.; Stigler, J.; Cedervall, T.; Berggård, T.; Flanagan, M. B.; Lynch, I.; Elia, G.; Dawson, K. The evolution of the protein corona around nanoparticles: a test study. *ACS nano* **2011**, *5*, 7503-7509.

(18) Tenzer, S.; Docter, D.; Kuharev, J.; Musyanovych, A.; Fetz, V.; Hecht, R.; Schlenk, F.; Fischer, D.; Kiouptsi, K.; Reinhardt, C. Rapid formation of plasma protein

corona critically affects nanoparticle pathophysiology. *Nature nanotechnology* **2013**, *8*, 772.

(19) Palchetti, S.; Pozzi, D.; Capriotti, A. L.; La Barbera, G.; Chiozzi, R. Z.; Digiacomo, L.; Peruzzi, G.; Caracciolo, G.; Laganà, A. Influence of dynamic flow environment on nanoparticle-protein corona: From protein patterns to uptake in cancer cells. *Colloids and Surfaces B: Biointerfaces* **2017**, *153*, 263-271.

(20) Dobrovolskaia, M. A.; Aggarwal, P.; Hall, J. B.; McNeil, S. E. Preclinical studies to understand nanoparticle interaction with the immune system and its potential effects on nanoparticle biodistribution. *Molecular pharmaceutics* **2008**, *5*, 487-495.

(21) Fei, L.; Perrett, S. Effect of nanoparticles on protein folding and fibrillogenesis. *International journal of molecular sciences* **2009**, *10*, 646-655.

(22) Gheshlaghi, Z. N.; Riazi, G. H.; Ahmadian, S.; Ghafari, M.; Mahinpour, R. Toxicity and interaction of titanium dioxide nanoparticles with microtubule protein. *Acta biochimica et biophysica Sinica* **2008**, *40*, 777-782.

(23) Zhao, A.; Zhou, S.; Wang, Y.; Chen, J.; Ye, C.; Huang, N. Molecular interaction of fibrinogen with thermally modified titanium dioxide nanoparticles. *RSC Advances* **2014**, *4*, 40428-40434.

(24) Wangoo, N.; Suri, C. R.; Shekhawat, G. Interaction of gold nanoparticles with protein: a spectroscopic study to monitor protein conformational changes. *Applied Physics Letters* **2008**, *92*, 133104.

(25) Saptarshi, S. R.; Duschl, A.; Lopata, A. L. Interaction of nanoparticles with proteins: relation to bio-reactivity of the nanoparticle. *Journal of nanobiotechnology* **2013**, *11*, 26.

(26) Yoo, S. I.; Yang, M.; Brender, J. R.; Subramanian, V.; Sun, K.; Joo, N. E.; Jeong, S. H.; Ramamoorthy, A.; Kotov, N. A. Inhibition of amyloid peptide fibrillation by inorganic nanoparticles: functional similarities with proteins. *Angewandte Chemie International Edition* **2011**, *50*, 5110-5115.

(27) Nel, A. E.; Mädler, L.; Velegol, D.; Xia, T.; Hoek, E. M.; Somasundaran, P.; Klaessig, F.; Castranova, V.; Thompson, M. Understanding biophysicochemical interactions at the nano–bio interface. *Nature materials* **2009**, *8*, 543.

(28) Deng, Z. J.; Liang, M.; Monteiro, M.; Toth, I.; Minchin, R. F. Nanoparticle-induced unfolding of fibrinogen promotes Mac-1 receptor activation and inflammation. *Nature nanotechnology* **2011**, *6*, 39.

(29) Mirshafiee, V.; Sun, B.; Chang, C. H.; Liao, Y.-P.; Jiang, W.; Jiang, J.; Liu, X.; Wang, X.; Xia, T.; Nel, A. E. Toxicological Profiling of Metal Oxide Nanoparticles in Liver Context Reveals Pyroptosis in Kupffer Cells and Macrophages versus Apoptosis in Hepatocytes. *ACS nano* **2018**, *12*, 3836-3852.

(30) Casals, E.; Pfaller, T.; Duschl, A.; Oostingh, G. J.; Puntes, V. F. Hardening of the nanoparticle–protein corona in metal (Au, Ag) and oxide (Fe3O4, CoO, and CeO2) nanoparticles. *Small* **2011**, *7*, 3479-3486.

(31) Moore, T. L.; Rodriguez-Lorenzo, L.; Hirsch, V.; Balog, S.; Urban, D.; Jud, C.; Rothen-Rutishauser, B.; Lattuada, M.; Petri-Fink, A. Nanoparticle colloidal stability in cell culture media and impact on cellular interactions. *Chemical Society Reviews* **2015**, *44*, 6287-6305.

(32) Durán, N.; Silveira, C. P.; Durán, M.; Martinez, D. S. T. Silver nanoparticle protein corona and toxicity: a mini-review. *Journal of nanobiotechnology* **2015**, *13*, 55.

(33) Saha, K.; Moyano, D. F.; Rotello, V. M. Protein coronas suppress the hemolytic activity of hydrophilic and hydrophobic nanoparticles. *Materials horizons* **2014**, *1*, 102-105.

(34) Hu, W.; Peng, C.; Lv, M.; Li, X.; Zhang, Y.; Chen, N.; Fan, C.; Huang, Q. Protein corona-mediated mitigation of cytotoxicity of graphene oxide. *ACS nano* **2011**, *5*, 3693-3700.

(35) Kirui, D. K.; Celia, C.; Molinaro, R.; Bansal, S. S.; Cosco, D.; Fresta, M.; Shen, H.; Ferrari, M. Mild hyperthermia enhances transport of liposomal gemcitabine and improves in vivo therapeutic response. *Advanced healthcare materials* **2015**, *4*, 1092-1103.

(36) Maeda, H.; Nakamura, H.; Fang, J. The EPR effect for macromolecular drug delivery to solid tumors: Improvement of tumor uptake, lowering of systemic toxicity, and distinct tumor imaging in vivo. *Advanced drug delivery reviews* **2013**, *65*, 71-79.

(37) Parhi, P.; Mohanty, C.; Sahoo, S. K. Nanotechnology-based combinational drug delivery: an emerging approach for cancer therapy. *Drug discovery today* **2012**, *17*, 1044-1052.

(38) Dobrovolskaia, M. A.; McNeil, S. E. Strategy for selecting nanotechnology carriers to overcome immunological and hematological toxicities challenging clinical translation of nucleic acid-based therapeutics. *Expert opinion on drug delivery* **2015**, *12*, 1163-1175.

(39) Parodi, A.; Corbo, C.; Cevenini, A.; Molinaro, R.; Palomba, R.; Pandolfi, L.; Agostini, M.; Salvatore, F.; Tasciotti, E. Enabling cytoplasmic delivery and organelle targeting by surface modification of nanocarriers. *Nanomedicine* **2015**, *10*, 1923-1940.

(40) Byrne, J. D.; Betancourt, T.; Brannon-Peppas, L. Active targeting schemes for nanoparticle systems in cancer therapeutics. *Advanced drug delivery reviews* **2008**, *60*, 1615-1626.

(41) Petros, R. A.; DeSimone, J. M. Strategies in the design of nanoparticles for therapeutic applications. *Nature reviews Drug discovery* **2010**, *9*, 615.

(42) Bertrand, N.; Wu, J.; Xu, X.; Kamaly, N.; Farokhzad, O. C. Cancer nanotechnology: the impact of passive and active targeting in the era of modern cancer biology. *Advanced drug delivery reviews* **2014**, *66*, 2-25.

(43) Torchilin, V. P. Multifunctional, stimuli-sensitive nanoparticulate systems for drug delivery. *Nature reviews Drug discovery* **2014**, *13*, 813.

(44) Koren, E.; Apte, A.; Jani, A.; Torchilin, V. P. Multifunctional PEGylated 2C5-immunoliposomes containing pH-sensitive bonds and TAT peptide for enhanced tumor cell internalization and cytotoxicity. *Journal of controlled release* **2012**, *160*, 264-273.

(45) Jiang, F.; Liu, B.; Lu, J.; Li, F.; Li, D.; Liang, C.; Dang, L.; Liu, J.; He, B.; Badshah, S. A. Progress and challenges in developing aptamer-functionalized targeted drug delivery systems. *International journal of molecular sciences* **2015**, *16*, 23784-23822.

(46) Choi, C. H. J.; Alabi, C. A.; Webster, P.; Davis, M. E. Mechanism of active targeting in solid tumors with transferrin-containing gold nanoparticles. *Proceedings of the National Academy of Sciences* **2010**, *107*, 1235-1240.

(47) Salvati, A.; Pitek, A. S.; Monopoli, M. P.; Prapainop, K.; Bombelli, F. B.; Hristov, D. R.; Kelly, P. M.; Åberg, C.; Mahon, E.; Dawson, K. A. Transferrin-functionalized nanoparticles lose their targeting capabilities when a biomolecule corona adsorbs on the surface. *Nature nanotechnology* **2013**, *8*, 137.

(48) Davis, M. E. The first targeted delivery of siRNA in humans via a self-assembling, cyclodextrin polymer-based nanoparticle: from concept to clinic. *Molecular pharmaceutics* **2009**, *6*, 659-668.

(49) Dai, Q.; Yan, Y.; Ang, C.-S.; Kempe, K.; Kamphuis, M. M.; Dodds, S. J.; Caruso, F. Monoclonal antibody-functionalized multilayered particles: targeting cancer cells in the presence of protein coronas. *ACS nano* **2015**, *9*, 2876-2885.

(50) Yan, Y.; Gause, K. T.; Kamphuis, M. M.; Ang, C.-S.; O'Brien-Simpson, N. M.; Lenzo, J. C.; Reynolds, E. C.; Nice, E. C.; Caruso, F. Differential roles of the protein corona in the cellular uptake of nanoporous polymer particles by monocyte and macrophage cell lines. *ACS nano* **2013**, *7*, 10960-10970.

(51) Lara, S.; Perez-Potti, A.; Herda, L. M.; Adumeau, L.; Dawson, K. A.; Yan, Y. Differential Recognition of Nanoparticle Protein Corona and Modified Low-Density Lipoprotein by Macrophage Receptor with Collagenous Structure. *ACS nano* **2018**, *12*, 4930-4937.

(52) Reddy, S. T.; Van Der Vlies, A. J.; Simeoni, E.; Angeli, V.; Randolph, G. J.; O'Neil, C. P.; Lee, L. K.; Swartz, M. A.; Hubbell, J. A. Exploiting lymphatic transport and complement activation in nanoparticle vaccines. *Nature biotechnology* **2007**, *25*, 1159.

(53) Salvador-Morales, C.; Flahaut, E.; Sim, E.; Sloan, J.; Green, M. L.; Sim, R. B. Complement activation and protein adsorption by carbon nanotubes. *Molecular immunology* **2006**, *43*, 193-201.

(54) Yu, K.; Lai, B. F.; Foley, J. H.; Krisinger, M. J.; Conway, E. M.; Kizhakkedathu, J. N. Modulation of complement activation and amplification on nanoparticle surfaces by glycopolymer conformation and chemistry. *ACS nano* **2014**, *8*, 7687-7703.

(55) Hubbell, J. A.; Thomas, S. N.; Swartz, M. A. Materials engineering for immunomodulation. *Nature* **2009**, *462*, 449.

(56) Boraschi, D.; Duschl, A. *Nanoparticles and the immune system: safety and effects*; Academic Press, 2013.

(57) Zheng, D.-W.; Chen, J.-L.; Zhu, J.-Y.; Rong, L.; Li, B.; Lei, Q.; Fan, J.-X.; Zou, M.-Z.; Li, C.; Cheng, S.-X. Highly integrated nano-platform for breaking the barrier between chemotherapy and immunotherapy. *Nano letters* **2016**, *16*, 4341-4347.

(58) Rosenberg, S. A.; Restifo, N. P. Adoptive cell transfer as personalized immunotherapy for human cancer. *Science* **2015**, *348*, 62-68.

(59) Lundqvist, M.; Stigler, J.; Elia, G.; Lynch, I.; Cedervall, T.; Dawson, K. A. Nanoparticle size and surface properties determine the protein corona with possible

implications for biological impacts. *Proceedings of the National Academy of Sciences* **2008**.

(60) Goy-López, S.; Juárez, J.; Alatorre-Meda, M.; Casals, E.; Puntes, V. F.; Taboada, P.; Mosquera, V. Physicochemical characteristics of protein–NP bioconjugates: the role of particle curvature and solution conditions on human serum albumin conformation and fibrillogenesis inhibition. *Langmuir* **2012**, *28*, 9113-9126.

(61) Hühn, D.; Kantner, K.; Geidel, C.; Brandholt, S.; De Cock, I.; Soenen, S. J.; Rivera_Gil, P.; Montenegro, J.-M.; Braeckmans, K.; Mullen, K. Polymer-coated nanoparticles interacting with proteins and cells: focusing on the sign of the net charge. *ACS nano* **2013**, *7*, 3253-3263.

(62) Lindman, S.; Lynch, I.; Thulin, E.; Nilsson, H.; Dawson, K. A.; Linse, S. Systematic investigation of the thermodynamics of HSA adsorption to N-iso-propylacrylamide/N-tert-butylacrylamide copolymer nanoparticles. Effects of particle size and hydrophobicity. *Nano letters* **2007**, *7*, 914-920.

(63) Gessner, A.; Waicz, R.; Lieske, A.; Paulke, B.-R.; Mäder, K.; Müller, R. Nanoparticles with decreasing surface hydrophobicities: influence on plasma protein adsorption. *International journal of pharmaceutics* **2000**, *196*, 245-249.

(64) Gagner, J. E.; Lopez, M. D.; Dordick, J. S.; Siegel, R. W. Effect of gold nanoparticle morphology on adsorbed protein structure and function. *Biomaterials* **2011**, *32*, 7241-7252.

(65) Ding, H.-m.; Ma, Y.-q. Computer simulation of the role of protein corona in cellular delivery of nanoparticles. *Biomaterials* **2014**, *35*, 8703-8710.

(66) Fan, H.; Zhao, D.; Li, Y.; Zhou, J. Lysozyme orientation and conformation on MoS2 surface: Insights from molecular simulations. *Biointerphases* **2017**, *12*, 02D416.

(67) Pan, Y.; Li, T.; Cheng, J.; Telesca, D.; Zink, J. I.; Jiang, J. Nano-QSAR modeling for predicting the cytotoxicity of metal oxide nanoparticles using novel descriptors. *RSC Advances* **2016**, *6*, 25766-25775.

(68) Papa, E.; Doucet, J.; Sangion, A.; Doucet-Panaye, A. Investigation of the influence of protein corona composition on gold nanoparticle bioactivity using machine learning approaches. *SAR and QSAR in Environmental Research* **2016**, *27*, 521-538.

(69) Gajewicz, A.; Schaeublin, N.; Rasulev, B.; Hussain, S.; Leszczynska, D.; Puzyn, T.; Leszczynski, J. Towards understanding mechanisms governing cytotoxicity of

metal oxides nanoparticles: Hints from nano-QSAR studies. *Nanotoxicology* **2015**, *9*, 313-325.

(70) Singh, K. P.; Gupta, S. Nano-QSAR modeling for predicting biological activity of diverse nanomaterials. *RSC Advances* **2014**, *4*, 13215-13230.

(71) Puzyn, T.; Rasulev, B.; Gajewicz, A.; Hu, X.; Dasari, T. P.; Michalkova, A.; Hwang, H.-M.; Toropov, A.; Leszczynska, D.; Leszczynski, J. Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles. *Nature nanotechnology* **2011**, *6*, 175.

(72) Xia, X.-R.; Monteiro-Riviere, N. A.; Riviere, J. E. An index for characterization of nanomaterials in biological systems. *Nature nanotechnology* **2010**, *5*, 671.

(73) Laera, S.; Ceccone, G.; Rossi, F.; Gilliland, D.; Hussain, R.; Siligardi, G.; Calzolai, L. Measuring protein structure and stability of protein–nanoparticle systems with synchrotron radiation circular dichroism. *Nano letters* **2011**, *11*, 4480-4484.

(74) Shang, W.; Nuffer, J. H.; Dordick, J. S.; Siegel, R. W. Unfolding of ribonuclease A on silica nanoparticle surfaces. *Nano letters* **2007**, *7*, 1991-1995.

(75) Zaman, M.; Ahmad, E.; Qadeer, A.; Rabbani, G.; Khan, R. H. Nanoparticles in relation to peptide and protein aggregation. *International journal of nanomedicine* **2014**, *9*, 899.

(76) Arnaudov, L. N.; de Vries, R. Thermally induced fibrillar aggregation of hen egg white lysozyme. *Biophysical Journal* **2005**, *88*, 515-526.

(77) Bellova, A.; Bystrenova, E.; Koneracka, M.; Kopcansky, P.; Valle, F.; Tomasovicova, N.; Timko, M.; Bagelova, J.; Biscarini, F.; Gazova, Z. Effect of Fe3O4 magnetic nanoparticles on lysozyme amyloid aggregation. *Nanotechnology* **2010**, *21*, 065103.

(78) Togashi, D. M.; Ryder, A. G.; Mc Mahon, D.; Dunne, P.; McManus, J. In *European Conference on Biomedical Optics*; Optical Society of America, 2007, p 6628_6661.

(79) Liu, L.; Zeng, L.; Wu, L.; Jiang, X. Revealing the Effect of Protein Weak Adsorption to Nanoparticles on the Interaction between the Desorbed Protein and its Binding Partner by Surface-Enhanced Infrared Spectroelectrochemistry. *Analytical chemistry* **2017**, *89*, 2724-2730.

(80) Shrivastava, S.; Nuffer, J. H.; Siegel, R. W.; Dordick, J. S. Position-specific chemical modification and quantitative proteomics disclose protein orientation adsorbed on silica nanoparticles. *Nano letters* **2012**, *12*, 1583-1587.

(81) Zeng, S.; Yu-ming, M. H.; Chia-en, A. C.; Zhong, W. Protein binding for detection of small changes on a nanoparticle surface. *Analyst* **2014**, *139*, 1364-1371.

(82) Srinivasu, B. Y.; Bose, B.; Mitra, G.; Kurpad, A. V.; Mandal, A. K. Adsorption induced changes of human hemoglobin on ferric pyrophosphate nanoparticle surface probed by isotope exchange mass spectrometry: An implication on structure–function correlation. *Langmuir* **2017**, *33*, 8032-8042.

(83) Lomenick, B.; Hao, R.; Jonai, N.; Chin, R. M.; Aghajan, M.; Warburton, S.; Wang, J.; Wu, R. P.; Gomez, F.; Loo, J. A. Target identification using drug affinity responsive target stability (DARTS). *Proceedings of the National Academy of Sciences* **2009**, pnas. 0910040106.

(84) Smits, A. H.; Vermeulen, M. Characterizing protein–protein interactions using mass spectrometry: challenges and opportunities. *Trends in biotechnology* **2016**, *34*, 825-834.

CHAPTER 2: Develop a High-Throughput Fluorescamine Labeling Method to Profile Nanoparticle-Protein Interactions

2.1: Introduction

Although promising nanomaterial-based biosensors, imaging probes, drug carriers, *etc*, have been developed, implementation of these materials in biomedical fields is still hindered by the lack of a thorough understanding about the complex impacts of nanomaterials on biological systems.[1-4] It has been established that the behaviors of nanomaterials in biosystems is strongly influenced by their interaction with proteins.[5,6] The protein interaction behaviors of nanomaterials could be indicative of their biological activity.[7-9] Nanomaterials can be produced with distinct or subtle differences in chemical composition, size, shape, surface modification, etc.[10-12] Proteins are highly diverse in their properties as well. Changes in properties of both proteins and particles strongly impact on protein-nanoparticle interaction, making it necessary to conduct such studies on large sample sets with a fast and high throughput assessment method.

Nanoparticle-protein interactions have been evaluated by separation[13,14] and spectroscopic techniques,[15,16] but problems exist with these methods, including requirements for target immobilization, multi-step sample processing, or modification. Alternatively, screening of protein adsorption on nanomaterials in biological matrices has been achieved using proteomic techniques;[6] and the adsorption profile was found to be dependent on the properties of the nanomaterial used.[17] Proteins undergo structural changes upon interacting with nanomaterials,[18] and the changes could be strongly influenced by the properties of nanomaterials, such as the surface curvature,[10,19,20] the chemical structure of

surface coating,[12,21-24] and the core material.[25,26] Protein structural changes during interaction with nanomaterials can be assessed by circular dichroism,[27-31] enzymatic activity measurement,[32] or dynamic light scattering;[33,34] but they are time consuming, tedious, and low throughput.

Herein, we developed a high-throughput assay for rapid screening of the interactions between proteins and nanomaterials. The assay relies on the quick and simple protein labeling by a fluorogenic dye, fluorescamine, which fluoresces upon reaction with a primary amine.[35] This property has been employed to label peptides and small molecules prior to chromatographic analysis, as well as quantifying proteins in a sample.[36-41] Fluorescamine's rapid reaction can restrict labeling to solely amines on the protein's surface, the number of which would be strongly affected by protein conformation and its interaction with other substances. In the present study, we tested whether fluorescamine labeling could be used to detect protein-nanoparticle interaction, and how the resulting fluorescence profile was related to the properties of nanomaterials.

2.2: Materials and Methods

Reagents used in the study. Fluorescamine was purchased from either Life Technologies (Carlsbad, CA) or Sigma-Aldrich (St. Louis, MO). Solid HEPES was purchased from CalBioChem (EMD Millipore, Darmstadt, Germany). Sodium phosphate monobasic monohydrate, anhydrous sodium phosphate dibasic, sodium chloride, sodium tetraborate, tris base, sodium hydroxide, hydrochloric acid, ethanol, dimethylformamide (DMF), dithiothreitol, iodoacetamide and glycine were all purchased from Thermo Fisher Scientific (Waltham, MA). All proteins were obtained from Sigma-Aldrich.

Nanomaterials investigated. Polystyrene (PS) nanoparticles with a carboxylated surface and core diameter of 42, 48 or 85 nm were purchased from Polysciences (Warrington, PA). Silica particles were synthesized in-house (synthesis in Supporting Methods) with varying degrees of surface carboxylation.

Circular dichroism (CD) to determine conformational change of proteins. To confirm that the fluorescence increase was due to protein conformational change, circular dichroism (CD) was conducted on two selected proteins to assay changes in secondary structure. CD spectra were collected on a Jasco J-815 spectrometer (details in Supporting Methods), and the secondary structure determined by the method outlined in Raussens *et al*.[42] As positive control, Apha-1-antitrypsin, succinyl concanavalin A, transferrin and human serum albumin were incubated in 50% ethanol (EtOH) aqueous solution to induce extreme conformational changes. After denaturation, the protein was incubated with fluorescamine (final protein and dye concentrations were 100 nM and 1 mM, respectively) for 10 minutes before fluorescence detection on the Victor II plate reader. As a control, proteins were treated under identical conditions with the 50% EtOH replaced by the 10 mM phosphate buffer (pH 8.0 with 50 mM NaCl).

On-plate fluorescamine assays for determining changes in the physicochemical properties of nanoparticles. On a 96-well PCR plate, 400 nM of each protein and 40 nM of one of the polystyrene or silica particles were incubated in 10 mM phosphate (pH 8.0), containing 50 mM NaCl. For each protein-particle pair, two controls were included. One contained only the protein; and the other had only the particles in the buffer. The final volume in each well was 100 μL. The plates were covered with a coverslip and incubated

at 37 °C in a water bath for 2 hrs. After incubation, 5 μL fluorescamine was added to each well (final concentration 1 mM), and reacted for 10 min at room temperature. The samples were then analyzed using the Victor II plate reader. Experiments were conducted in triplicate.

Afterward, the fluorescence intensity for each protein-particle pair was normalized to the particle-free signal, and the normalized values for all pairs were subject to principal components analysis (PCA). The scores plot was prepared using the first two principal components to show the grouping effect of nanoparticles. The loadings plot was used to determine the relative contribution of each protein to the particle's location on the scores plot.

2.3: Results and Discussion

Nanoparticles and proteins used in the study. All nanoparticles chosen for this study are listed in Table 2.1. The carboxylated PS and silica particles, representing optically transparent nanomaterials, were used in the on-plate assay. Two of the PS samples were from different batches of the same product, with an average diameter of 45 nm. The third PS sample had a larger diameter of 85 nm. The silica particles, fabricated by a modified Stöber synthesis[43], also carried carboxyl groups. Particles were aminated for various times prior to the 24 hour carboxylation process. DLS and zeta-potential measurements indicated that all silica particles had similar hydrodynamic diameters of 85-95 nm with small difference in surface charge density. Using these particles, we explored the capability of fluorescamine in detecting the protein's interaction with transparent

particles having difference in core material, size, and synthesis process, in the simple on-plate assay format.

Table 2.1. Physical parameters of nanoparticles investigated as part of the study, we well as their abbreviations in PCA analysis.

Particle	Abbrev. in PCA	Average Hydrodyn. Diameter	Average Zeta Potential
42 nm carboxylated polystyrene (PS)	PS-1	42 ± 6 nm	-45 mV
48 nm carboxylated polystyrene	PS-2	48 ± 7 nm	-87 mV
85 nm carboxylated polystyrene	PS-3	85 ± 7 nm	-67 mV
Silica (Si) particle aminated for 1 hr before carboxylation	Si-1	87 ± 18 nm	-44 mV
Silica particle aminated for 4 hrs before carboxylation	Si-2	94 ± 16 nm	-42 mV
Silica particle aminated for 24 hrs before carboxylation	Si-3	84 ± 12 nm	-51 mV

To prove the principle of our design, succinyl concanavalin A (ConA), alpha-1-antitrypsin (A1AT), serum albumin (HSA), transferrin, haptoglobin, and apolipoprotein A1 (APOA1) were employed. Screening of protein interaction on all of the selected nanoparticles was done on proteins with different size/Mw, isoelectric point, and hydrophobicity: cytochrome C (cyt C), hemoglobin, catalase, HSA, transferrin, fibrinogen and thyroglobulin. Protein hydrophobicity was represented by the grand average of hydropathy (GRAVY) scores calculated using ProtParam, a tool available in the SIB ExPASy Bioinformatics Resources Portal.[44] The selected proteins covered a wide range of Mw (from 11.5 to 300 kDa), pI (from 5.2 to 9.5), hydrophobicity (GRAVY score from -

0.004 to -0.885), and number of primary amines (from 17 to 255). The interaction of proteins possessing diverse properties with nanoparticles aims to reveal the key parameters that influence protein-nanoparticle interaction.

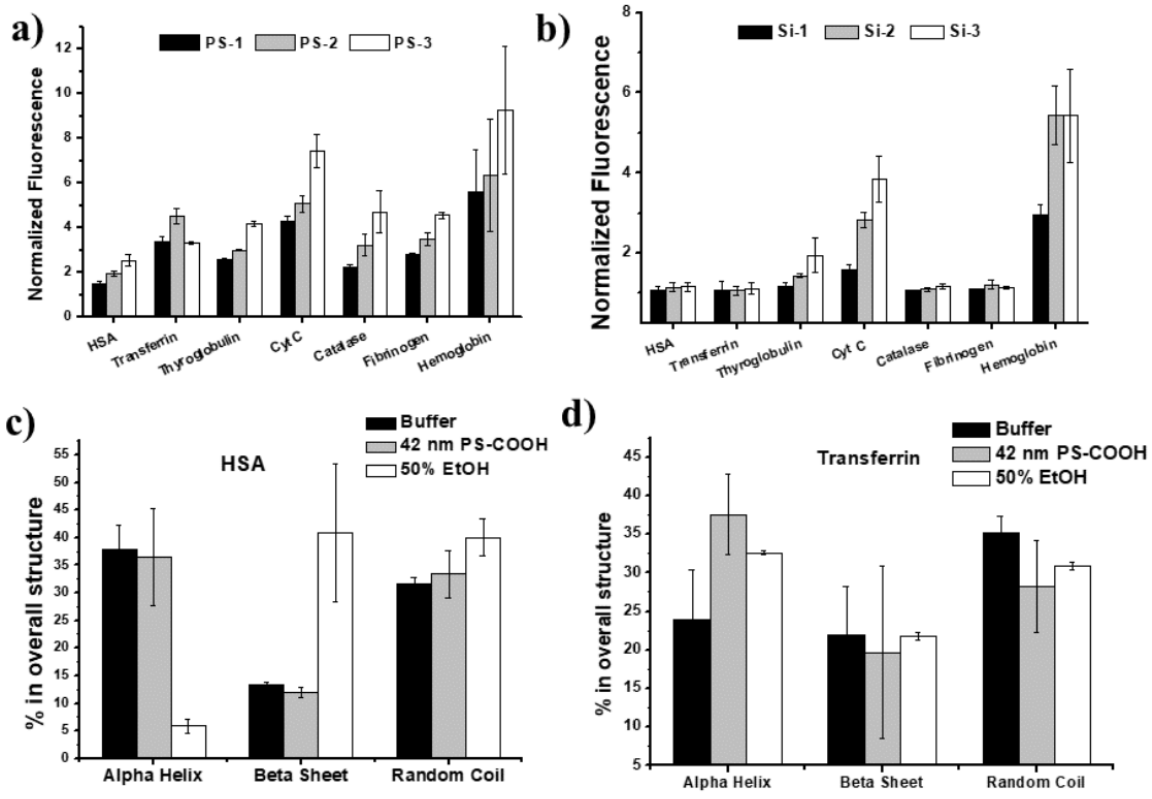


Figure 2.1. Normalized fluorescence changes of proteins after incubation with either (a) polystyrene or (b) silica nanoparticles. Fluorescence intensity were normalized to the buffer-only protein fluorescence. The percent changes in the protein secondary structures upon binding with the 42-nm PS nanoparticles and upon denatured by 50% EtOH for (c) HSA or (d) transferrin. The percentage of each secondary structure was obtained by Circular Dichroism (CD).

Fluorescamine assay for revealing interactions between proteins and nanoparticles. Specific labeling of a protein's surface by fluorescamine could be useful for detection of interaction between proteins and nanoparticles. Once the protein is adsorbed onto the nanoparticle, the number of surface amines would change due to

disturbances in the protein's tertiary structure and formation of the binding interface, altering the labeling result. We examined the interaction of HSA and transferrin with the 42-nm carboxylated PS particles by fluorescamine labeling. Compared to the protein itself, the resulting fluorescence upon interaction with the nanoparticles increased by 1.5 and 3 fold for HSA and transferrin, respectively (Figure 2.1 a). CD measurement confirmed that interaction with the PS particles significantly increased the content of alpha helix and slightly reduced random coils in transferrin (Figure 2.1 d). Such conformational changes in protein structure may have exposed more amines to the surface for fluorescamine to label. Agreeing with the fluorescamine result, conformational change in HSA detected by CD was not as large as in transferrin when adsorbed by the PS particles. No noticeable change was detected by CD in HSA (Figure 2.1 c). A large molar ratio of protein to particle (at a ratio of 100:1) was used in CD measurement for reduction of the background UV absorbance from the nanoparticles. The number of HSA adsorbed by the nanoparticles was too small to induce sufficient change in the bulk protein population to be detected by CD.

Relationship between fluorescence profile and particle properties. We further explored the relationship between protein labeling and nanoparticle properties. Interactions between various PS and silica nanoparticles and a group of proteins were examined using the fluorescamine assay. A total of four types of nanoparticles incubated with seven proteins were screened per experiment, plus the corresponding particle-only and protein-only controls. A 96-well plate was used to simultaneously test such a high number of samples. The screening was performed in triplicate. All fluorescence signals of the protein-particle mixtures were normalized to that of the protein-only blank (Figure 2.1 a/b). The

fluorescence was found to increase for most proteins upon binding to the nanoparticles, reflecting the exposure of more surface amines due to conformational change in protein structure. The particles could react with fluorescamine, but the resulting signal was negligible compared to that observed in the protein-particle incubations.

To interpret the effects of particle properties on protein interactions, we focused on nanoparticles differing in one property aspect and compared their fluorescence profiles. Figure 2.1a compared the average normalized fluorescence observed from proteins when incubated with the carboxylated PS particles of different diameters. Fluorescence profile comparison of the silica particles with varying amination durations before the carboxylation step during synthesis was shown in Figure 2.1b. For both the PS and silica particle groups, larger normalized fluorescence were obtained with cyt C and hemoglobin, the two proteins with higher pIs than others. Hemoglobin is neutral and cyt C is slightly positively charged at pH 8.0, which could induce stronger electrostatic interaction with the negatively charged carboxylated particles. The rest 5 proteins also showed significant difference in the normalized fluorescence when interacting with PS particles of different sizes. On contrary, the labeling situation did not exhibit noticeable difference when these proteins incubated with the three types of silica particles, except for the largest protein, thyroglobulin.

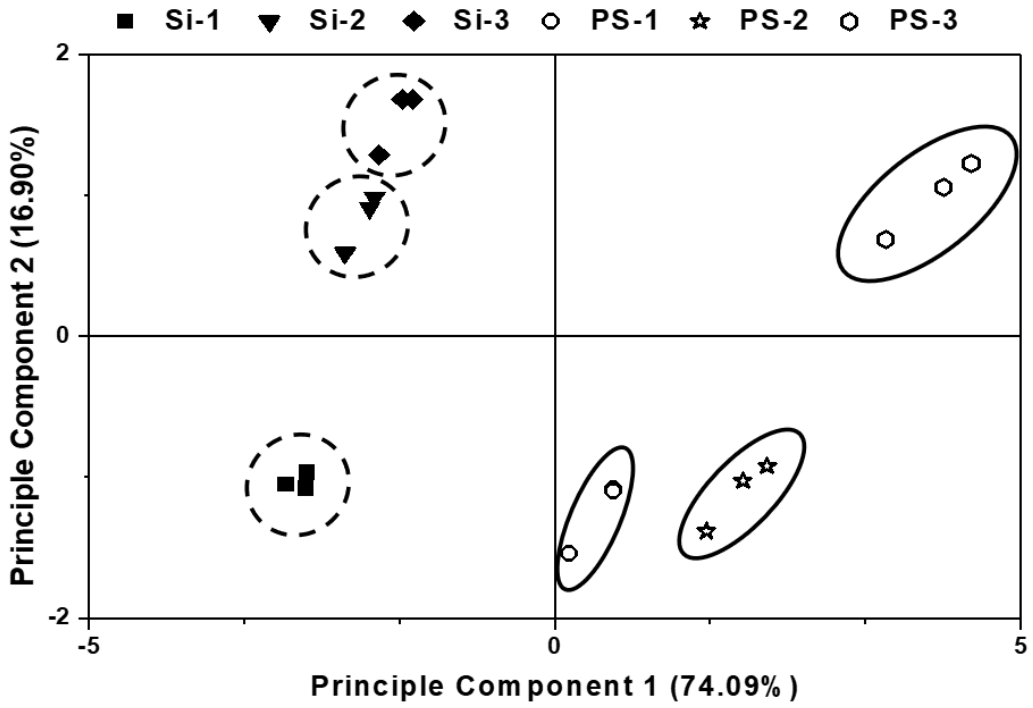


Figure 2.2. PCA scores plot showing the ability of the fluorescamine assay to differentiate between six different particles of varying physiochemical parameters.

To summarize the overall variations in fluorescamine labeling on different protein-nanoparticle pairs, the normalized fluorescence dataset was subjected to principal components analysis (PCA). The free data mining software TANAGRA developed by Ricco Rakotomalala[28] was used and the PCA results were exported to Origin and plotted. During PCA, each repetition (an average of three replicates on plate) of the same protein-nanoparticle pair was treated as one individual observation; and the proteins were viewed as variables. The resulting scores plot was displayed in Figure 2.2. Interestingly, the silica particles and the PS particles can be differentiated by the first principle component (PC1), which accounts for 74.09% of the overall variance in the dataset. In addition, the PS particles of different sizes were clearly separated from each other, with the repeated

measurements on the same particles clustered together. The 85-nm PS particles were separated from the two smaller particles by the 2nd principle component (PC2). The smaller PS particles both had hydrodynamic size around 45 nm but differed in their zeta-potentials. Similarly, the two silica particles experiencing longer amination process located closer to each other on the scores plot, while the one going through only 1 hr amination was farther away. Our results support that fluorescamine labeling can detect protein-particle interaction and the interaction profile is strongly dependent on particle properties.

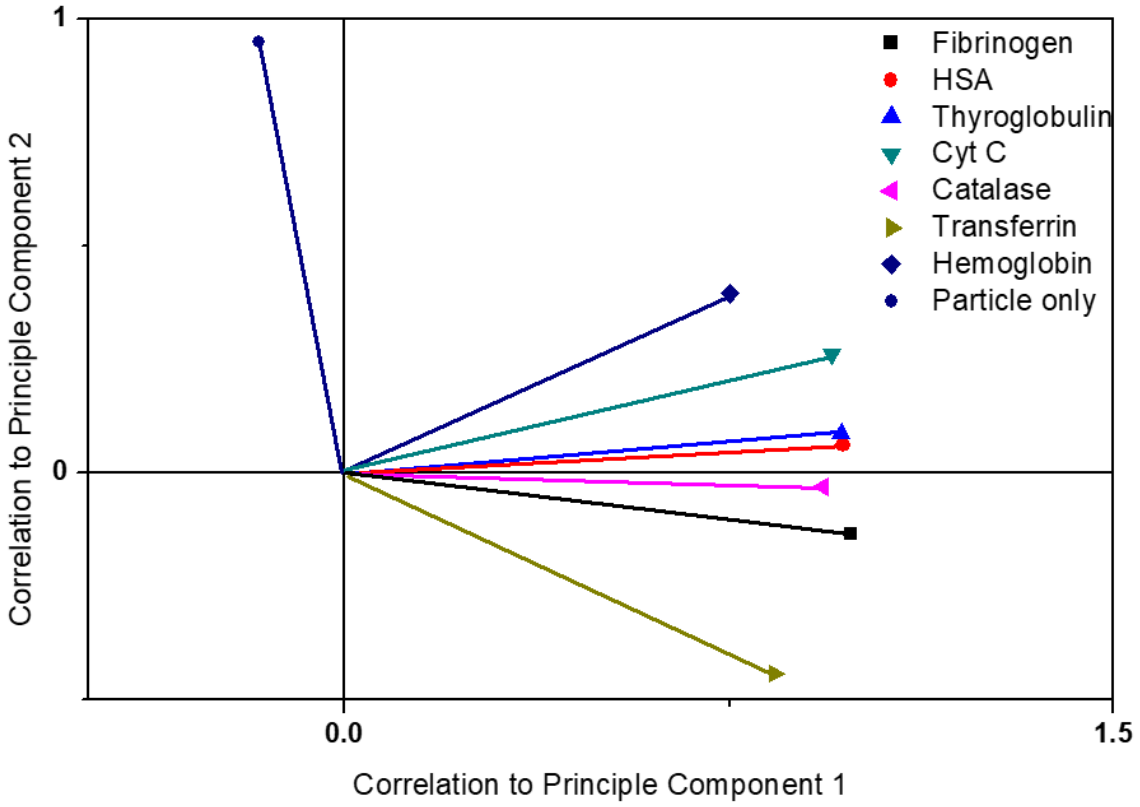


Figure 2.3. The loading plot associated with PCA sores plot shown in figure 2.2 for the PS and silica particles. The plot depicts protein's correlations to principle component 1 and 2 that determine the location of particles on the scores plot.

PCA also calculates the correlation of each variable with each principle component, and the result is displayed in the form of loading plot (Figure 2.3). The vectors of fibrinogen, HSA, thyroglobulin, and catalase all aligned with the axis of PC1 with minimal projection in the dimension of PC 2, indicating their important contribution in determination of PC1 that differentiated particles based on their core materials. Cyt c, transferrin and hemoglobin are responsible more to the separation in the dimension align with PC2 that differentiate particles by their sizes for the PS particles. The fluorescence signal resulted from the particle itself, although it was negligible in its net intensity when compared to the signals from the proteins, also played an important role in differentiating the silica particles with longer amination durations from other particles. The longer amination duration may have left more amines on the surface uncovered by carboxyl groups.

2.4: Conclusions

Our study proves that fluorescamine labeling can be used to detect nanoparticle-protein interaction, because of its capability in targeting the surface amines on proteins. Formation of the binding interface would block some surface amines from being assessed by fluorescamine; on the other hand, conformational change in protein could reveal more surface amine for labeling. Distinct differences in fluorescence post-labeling can be observed before and after incubating the protein with nanoparticles. Although the exact reason for the fluorescence change can only be understood by investigations using other analytical tools, our results have demonstrated that the fluorescamine assay can serve as a rapid screening method for probing the interaction between a large number of proteins and

nanoparticles. The obtained fluorescence profile can be used to study the relationship between particle property and protein interaction, as well as to determine batch similarity of the material prior to the more expensive characterization. Such screening could be useful for initial assessment of the biological activity and safety of nanoparticles for biomedical applications.

The technique has limitations in that it is not readily suited for testing particles coated with amine groups, or proteins low in amine content like ConA and Cyt C, which will offer low signal changes compared to proteins with a higher degree of amine content. However, such difficulties could be solved if fluorogenic dyes targeting different functional groups, such as thiol or hydrophobic residues on proteins are employed. In addition, for screening interactions on non-amine coated particles aiming to reveal changes in particle's physical parameters, suitable proteins with adequate amine contents can be chosen to ensure large signal changes can be observed.

Nevertheless, follow-up studies are needed for further exploration the applicability of our method, using proteins and particles with more diverse properties. For example, the method's efficacy on screening particle-protein interactions governed by hydrophobic forces is not clear, since the assay targets amine groups which are involved more in hydrophilic interactions. It will still be useful, if the binding results in partial unfolding of the protein to reveal the amine groups originally buried inside the tertiary structure of the protein. More studies using proteins and particles with more diverse properties are needed in order to answer this question, which should not be difficult with the high-throughput sampling capability of our method. Once the binding situations are disclosed by

fluorescamine labeling, more detailed studies with techniques that can reveal the exact degree and type of structural changes in proteins, like CD, can be performed on specific protein-particle pairs to obtain more insightful information about the interactions.

References

(1) Schulze, C.; Kroll, A.; Lehr, C. M.; Schafer, U. F.; Becker, K.; Schnekenburger, J.; Isfort, C. S.; Landsiedel, R.; Wohlleben, W. Not ready to use - overcoming pitfalls when dispersing nanoparticles in physiological media. *Nanotoxicology* **2008**, *2*, 51-U17.

(2) Mu, Q.; Jiang, G.; Chen, L.; Zhou, H.; Fourches, D.; Tropsha, A.; Yan, B. Chemical Basis of Interactions Between Engineered Nanoparticles and Biological Systems. *Chemical Reviews (Washington, DC, United States)* **2014**, *114*, 7740-7781.

(3) Nystroem, A. M.; Fadeel, B. Safety assessment of nanomaterials: Implications for nanomedicine. *Journal of Controlled Release* **2012**, *161*, 403-408.

(4) Rivera-Gil, P.; Jimenez De Aberasturi, D.; Wulf, V.; Pelaz, B.; Del Pino, P.; Zhao, Y.; De La Fuente, J. M.; Ruiz De Larramendi, I.; Rojo, T.; Liang, X.-J.; Parak, W. J. The Challenge To Relate the Physicochemical Properties of Colloidal Nanoparticles to Their Cytotoxicity. *Accounts of Chemical Research* **2013**, *46*, 743-749.

(5) Gao, H.; He, Q. The interaction of nanoparticles with plasma proteins and the consequent influence on nanoparticles behavior. *Expert Opinion on Drug Delivery* **2014**, *11*, 409-420.

(6) Mahmoudi, M.; Lynch, I.; Ejtehadi, M. R.; Monopoli, M. P.; Bombelli, F. B.; Laurent, S. Protein-Nanoparticle Interactions: Opportunities and Challenges. *Chemical Reviews (Washington, DC, United States)* **2011**, *111*, 5610-5637.

(7) Saptarshi, S. R.; Duschl, A.; Lopata, A. L. Interaction of nanoparticles with proteins: relation to bio-reactivity of the nanoparticle. *Journal of Nanobiotechnology* **2013**, *11*, 26.

(8) White, J. W.; Perriman, A. W.; McGillivray, D. J.; Lin, J.-M. Protein interfacial structure and nanotoxicology. *Nuclear Instruments & Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors, and Associated Equipment* **2009**, *600*, 263-265.

(9) You, C.-C.; De, M.; Rotello, V. M. Monolayer-protected nanoparticle-protein interactions. *Current Opinion in Chemical Biology* **2005**, *9*, 639-646.

(10) Chithrani, B. D.; Ghazani, A. A.; Chan, W. C. W. Determining the size and shape dependence of gold nanoparticle uptake into mammalian cells. *Nano Letters* **2006**, *6*, 662-668.

(11) Gagner, J. E.; Qian, X.; Lopez, M. M.; Dordick, J. S.; Siegel, R. W. Effect of gold nanoparticle structure on the conformation and function of adsorbed proteins. *Biomaterials* **2012**, *33*, 8503-8516.

(12) Walkey, C. D.; Olsen, J. B.; Guo, H. B.; Emili, A.; Chan, W. C. W. Nanoparticle Size and Surface Chemistry Determine Serum Protein Adsorption and Macrophage Uptake. *Journal of the American Chemical Society* **2012**, *134*, 2139-2147.

(13) Ashby, J.; Schachermeyer, S.; Pan, S.; Zhong, W. Dissociation-Based Screening of Nanoparticle-Protein Interaction via Flow Field-Flow Fractionation. *Analytical Chemistry (Washington, DC, United States)* **2013**, *85*, 7494-7501.

(14) Li, N.; Zeng, S.; He, L.; Zhong, W. Probing Nanoparticle-Protein Interaction by Capillary Electrophoresis. *Analytical Chemistry (Washington, DC, United States)* **2010**, *82*, 7460-7466.

(15) Piehler, J. Spectroscopic techniques for monitoring protein interactions in living cells. *Current Opinion in Structural Biology* **2014**, *24*, 54-62.

(16) Li, L.; Mu, Q.; Zhang, B.; Yan, B. Analytical strategies for detecting nanoparticle-protein interactions. *Analyst (Cambridge, United Kingdom)* **2010**, *135*, 1519-1530.

(17) Zhang, H.; Burnum, K. E.; Luna, M. L.; Petritis, B. O.; Kim, J.-S.; Qian, W.-J.; Moore, R. J.; Heredia-Langner, A.; Webb-Robertson, B.-J. M.; Thrall, B. D.; Camp, D. G.; Smith, R. D.; Pounds, J. G.; Liu, T. Quantitative proteomics analysis of adsorbed plasma proteins classifies nanoparticles with different surface properties and size. *Proteomics* **2011**, *11*, 4569-4577.

(18) Pan, H.; Qin, M.; Meng, W.; Cao, Y.; Wang, W. How Do Proteins Unfold upon Adsorption on Nanoparticle Surfaces? *Langmuir* **2012**, *28*, 12779-12787.

(19) Tenzer, S.; Docter, D.; Rosfa, S.; Wlodarski, A.; Kuharev, J.; Rekik, A.; Knauer, S. K.; Bantz, C.; Nawroth, T.; Bier, C.; Sirirattanapan, J.; Mann, W.; Treuel, L.; Zellner, R.; Maskos, M.; Schild, H.; Stauber, R. H. Nanoparticle Size Is a Critical Physicochemical Determinant of the Human Blood Plasma Corona: A Comprehensive Quantitative Proteomic Analysis. *ACS Nano* **2011**, *5*, 7155-7167.

(20) Bar-Ilan, O.; Albrecht, R. M.; Fako, V. E.; Furgeson, D. Y. Toxicity Assessments of Multisized Gold and Silver Nanoparticles in Zebrafish Embryos. *Small* **2009**, *5*, 1897-1910.

(21) Moghimi, S. M.; Muir, I. S.; Illum, L.; Davis, S. S.; Kolbbachofen, V. COATING PARTICLES WITH A BLOCK-COPOLYMER (POLOXAMINE-908) SUPPRESSES OPSONIZATION BUT PERMITS THE ACTIVITY OF DYSOPSONINS IN THE SERUM. *Biochimica Et Biophysica Acta* **1993**, *1179*, 157-165.

(22) Lundqvist, M.; Stigler, J.; Elia, G.; Lynch, I.; Cedervall, T.; Dawson, K. A. Nanoparticle size and surface properties determine the protein corona with possible implications for biological impacts. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 14265-14270.

(23) Gref, R.; Luck, M.; Quellec, P.; Marchand, M.; Dellacherie, E.; Harnisch, S.; Blunk, T.; Muller, R. H. 'Stealth' corona-core nanoparticles surface modified by polyethylene glycol (PEG): influences of the corona (PEG chain length and surface density) and of the core composition on phagocytic uptake and plasma protein adsorption. *Colloid Surf. B-Biointerfaces* **2000**, *18*, 301-313.

(24) Natte, K.; Friedrich, J. F.; Wohlrab, S.; Lutzki, J.; von Klitzing, R.; Osterle, W.; Orts-Gil, G. Impact of polymer shell on the formation and time evolution of nanoparticle-protein corona. *Colloid Surf. B-Biointerfaces* **2013**, *104*, 213-220.

(25) Asharani, P. V.; Yi, L. W.; Gong, Z. Y.; Valiyaveettil, S. Comparison of the toxicity of silver, gold and platinum nanoparticles in developing zebrafish embryos. *Nanotoxicology* **2011**, *5*, 43-54.

(26) Casals, E.; Pfaller, T.; Duschl, A.; Oostingh, G. J.; Puntes, V. F. Hardening of the Nanoparticle-Protein Corona in Metal (Au, Ag) and Oxide (Fe3O4, CoO, and CeO2) Nanoparticles. *Small* **2011**, *7*, 3479-3486.

(27) Kelly, S. M.; Jess, T. J.; Price, N. C. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2005**, *1751*, 119-139.

(28) Wang, L. M.; Li, J. Y.; Pan, J.; Jiang, X. M.; Ji, Y. L.; Li, Y. F.; Qu, Y.; Zhao, Y. L.; Wu, X. C.; Chen, C. Y. Revealing the Binding Structure of the Protein Corona on Gold Nanorods Using Synchrotron Radiation-Based Techniques: Understanding the Reduced Damage in Cell Membranes. *Journal of the American Chemical Society* **2013**, *135*, 17359-17368.

(29) Billsten, P.; Wahlgren, M.; Arnebrant, T.; McGuire, J.; Elwing, H. STRUCTURAL-CHANGES OF T4 LYSOZYME UPON ADSORPTION TO SILICA NANOPARTICLES MEASURED BY CIRCULAR-DICHROISM. *Journal of Colloid and Interface Science* **1995**, *175*, 77-82.

(30) Venyaminov, S. Y.; Vassilenko, K. S. DETERMINATION OF PROTEIN TERTIARY STRUCTURE CLASS FROM CIRCULAR-DICHROISM SPECTRA. *Analytical Biochemistry* **1994**, *222*, 176-184.

(31) Venyaminov, S. Y.; Klimtchuk, E. S.; Bajzer, Z.; Craig, T. A. Changes in structure and stability of calbindin-D-28K upon calcium binding. *Analytical Biochemistry* **2004**, *334*, 97-105.

(32) Vertegel, A. A.; Siegel, R. W.; Dordick, J. S. Silica nanoparticle size influences the structure and enzymatic activity of adsorbed lysozyme. *Langmuir* **2004**, *20*, 6800-6807.

(33) Jans, H.; Liu, X.; Austin, L.; Maes, G.; Huo, Q. Dynamic Light Scattering as a Powerful Tool for Gold Nanoparticle Bioconjugation and Biomolecular Binding Studies. *Anal. Chem.* **2009**, *81*, 9425-9432.

(34) James, A. E.; Driskell, J. D. Monitoring gold nanoparticle conjugation and analysis of biomolecular binding with nanoparticle tracking analysis (NTA) and dynamic light scattering (DLS). *Analyst* **2013**, *138*, 1212-1218.

(35) Udenfrie.S; Stein, S.; Bohlen, P.; Dairman, W. FLUORESCAMINE - REAGENT FOR ASSAY OF AMINO-ACIDS, PEPTIDES, PROTEINS, AND PRIMARY AMINES IN PICOMOLE RANGE. *Science* **1972**, *178*, 871-&.

(36) Noble, J. E.; Knight, A. E.; Reason, A. J.; Di Matola, A.; Bailey, M. J. A. A comparison of protein quantitation assays for biopharmaceutical applications. *Mol. Biotechnol.* **2007**, *37*, 99-111.

(37) Joo, W.-A.; Speicher, D. W. Protein detection in gels without fixation. *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]* **2007**, *Chapter 10*, Unit 10.16.

(38) Birlouez-Aragon, I.; Leclere, J.; Quedraogo, C. L.; Birlouez, E.; Grongnet, J. F. The FAST method, a rapid approach of the nutritional quality of heat-treated foods. *Nahr.-Food* **2001**, *45*, 201-205.

(39) Chen, Y.; Zhang, Y. Q. Fluorescent quantification of amino groups on silica nanoparticle surfaces. *Anal. Bioanal. Chem.* **2011**, *399*, 2503-2509.

(40) Stockert, J. C.; Blazquez, A.; Galaz, S.; Juarranz, A. A mechanism for the fluorogenic reaction of amino groups with fluorescamine and MDPF. *Acta Histochem.* **2008**, *110*, 333-340.

(41) Dhaunta, N.; Fatima, U.; Guptasarma, P. N-Terminal sequencing by mass spectrometry through specific fluorescamine labeling of alpha-amino groups before tryptic digestion. *Analytical Biochemistry* **2011**, *408*, 263-268.

(42) Raussens, V.; Ruysschaert, J.-M.; Goormaghtigh, E. Protein concentration is not an absolute prerequisite for the determination of secondary structure from circular dichroism spectra: a new scaling method. *Analytical Biochemistry* **2003**, *319*, 114-121.

(43) An, Y. Q.; Chen, M.; Xue, Q. J.; Liu, W. M. Preparation and self-assembly of carboxylic acid-functionalized silica. *Journal of Colloid and Interface Science* **2007**, *311*, 507-513.

(44) Ashby, J.; Pan, S.; Zhong, W. Size and Surface Functionalization of Iron Oxide Nanoparticles Influence the Composition and Dynamic Nature of Their Protein Corona. *ACS Applied Materials and Interfaces* **2014**, *6*, 15412–15419.

CHAPTER 3: Fluorescamine Labeling for Assessment of Protein Conformational Change and Binding Affinity in Protein-Nanoparticle Interaction

3.1: Introduction

Once entering the physiological environment like blood and cell cytoplasm, engineered nanomaterials (ENMs) are known to adsorb proteins and form the "protein corona" which acts as the new "biological identity" of ENMs.[1-3] The composition of protein corona can change dynamically, depending on interaction duration and the type and concentration of proteins in the environment. [8] The more abundant proteins could be adsorbed first and then displaced by the low-abundance-but-high-affinity proteins after long duration, forming the stable "hard corona".[1,4,5] Different physiological outcomes could then occur to the ENMs depending on what proteins are adsorbed by the ENMs.[6] For instance, serum albumin, found in the corona of silver, silica or polymer ENMs present in plasma could elongate their circulation time to facilitate uptake of the ENMs by the target cells, a property favored by ENMs employed as drug carriers, therapeutic reagents, or diagnostic tools.[7-9] Transferrin adsorbed by the polystyrene and poly(glycidyl methacrylate) ENMs could actively target the breast or prostate cancer cells overexpressing transferrin receptor.[10,11] However, serum albumin could reduce the targeting capability of such ENMs if co-existing in the surrounding environment,[12] probably because albumin displaces transferrin off the ENM surface or interferes with receptor binding by covering up the binding sites on transferrin. Besides the type and affinity of the adsorbed proteins, changes to protein conformation upon adsorption could alter the biological responses to ENMs. Not only protein function could be impaired,[13] but also the unfolded proteins could

enhance uptake by immune cells, activate inflammation, as well as induce other side effects like blood coagulation, membrane structure damage and complement activation.[14]

The close correlation between protein adsorption and biological responses to ENMs calls for better understanding of the interactions between proteins and ENMs and how the interaction is affected by the properties of ENMs and proteins.[13,15,16] The knowledge can guide the design of ENMs to promote the desired activity and suppress the potential adverse effects, improving the efficacy and safety of ENM implementation. Protein binding affinity and conformational change are two important aspects to be assessed in study of protein-ENM interaction. The most common tool for study of protein conformational change is circular dichroism (CD), which is widely available and simple to carry out. Hydrogen-deuterium exchange (HDX) or fast photochemical oxidation of proteins (FPOP) coupled with MS or NMR is more sensitive than CD and can provide details information about protein conformation.[17,18] Moreover, chemical crosslinking coupled with MS has been used to explore protein conformational change as well as the binding sites of protein on ENMs.[19,20] On the other hand, binding affinity can be measured using separation methods including ultracentrifugation,[4] capillary electrophoresis (CE),[21] and size exclusive chromatography (SEC).[5] Moreover, surface plasma resonance (SPR)[5], quartz crystal microbalance (QCM)[22], isothermal titration calorimetry (ITC)[5] and enzyme-linked immunosorbent assay (ELISA)[23] have been employed to study binding affinities and kinetics. However, the aforementioned techniques typically require expensive instruments, are time consuming and technically demanding, and could be compromised by the optical properties of ENMs. They are also not applicable for screening the interaction between a

large number of proteins and ENMs. Fluorescence measurement provides high sensitivity and is compatible with high-throughput screening. Although it has been applied to measure the binding affinities of proteins on some ENMs taking advantage their intrinsic property in fluorescence quenching[24], it does not work for the ENMs not capable of generating changes in the optical properties of the interacting parties.

In our previous work, we have developed a fluorescence-screening method that does not rely on the intrinsic optical properties of ENMs or proteins, but utilizes fluorescamine labeling to reveal protein-nanoparticle interaction in a rapid and high-throughput manner. We demonstrated that protein-nanoparticle binding could alter fluorescamine labeling of the protein, and the resultant fluorescence patterns could differentiate the particles based on their core composition and size and surface properties.[25] In the present study, we applied this assay to screen the interaction of a good number of proteins with two sets of nanoparticles, revealing the key protein properties that could influence protein-nanoparticle interaction and the close correlation between protein conformational change and fluorescamine labeling. We also showed that this method could be applied to evaluate the binding affinities. With the capability of assessing both protein conformational change and binding strength, our method should be very useful in revealing the key protein and ENM properties that govern protein-ENM interaction.

3.2: Materials and Methods

Chemicals and Biochemicals. Fluorescamine and all proteins were purchased from Sigma-Aldrich (St. Louis, MO). Dimethyl sulfoxide (DMSO), sodium phosphate monobasic monohydrate, anhydrous sodium phosphate dibasic, sodium chloride and

bicinchoninic (BCA) assay kit were obtained from Fisher Scientific (Waltham, MA). Ultrapure water with electric resistance > 18.2MΩ was produced in-house, by the Milipore Milli-Q water purification system (Billerica, MA).

Nanoparticles. The carboxylated polystyrene nanoparticles with an average diameter of 48 and 85 nm (PS48 and PS85) were obtained from Polysciences (Warrington, PA). The silica nanoparticles with an average diameter of 50 and 80 nm (Si50 and Si80) were purchased from nanoComposix (San Diego, CA).

Determination of protein properties. Properties of protein, including molecular weight (MW), theoretical isoelectric point (pI), and the grand average of hydropathicity (GRAVY) value, were calculated based on protein sequences via the ProtParam tool available in ExPASy. All protein sequences were downloaded from NCBI. pI was calculated using pK values of amino acids described in Bjellqvist et al. The GRAVY value was calculated as the average of hydropathicity values of all amino acids in protein sequences.

Fluorescamine screening. PS48 of 10 nM, PS85 of 3.2 nM, Si50 of 10 nM, or Si80 of 3.9 nM was incubated with 400 nM protein in the PBS buffer (10 mM phosphate at pH 7.4, 137 mM NaCl, and 2.7 mM KCl) for 1hr at 37 ˚C. The different particle molarities provided similar surface areas for protein adsorption. Then, fluorescamine was added into the mixture at a final concentration of 1 mM, and incubated for 5 min, before fluorescence detection was carried out in the Victor II plate reader.

Data processing. We defined F_0 as the fluorescence from the protein alone labeled by fluorescamine, and F as the fluorescence from the protein incubated with the NPs.

Fluorescence change ratio (F/F_0) was calculated and subject to principal components analysis (PCA) using the R package 'ggfortify'. The first two principal components were used for making the scores plots. In addition, k-means clustering was done by the same package. Significant test for fluorescence screening and PCA results were evaluated by MANOVA and linear discriminant analysis (LDA) with Jackknife prediction, with either the fluorescence change ratios or the first two principle values (PC1 and PC2) from PCA were used as the variables. For Jackknife prediction, all variables except one were used as the training set to classify proteins, and the one left out was used as the test set for test of assignment accuracy.

Measurement of NPs size. PS48 NPs of 4 nM were incubated with different concentrations of HSA (0-12.8μM) in 1×PBS buffer for 1h at 37 ºC, and were diluted 2,000 folds by 1×PBS. Nanoparticles Tracking Analysis (NTA) was used to measure the diameter of PS48 NPs after dilution, in which the Brownian motion of the NPs was monitored by a laser and converted to the hydrodynamic diameter.

Quantification of protein absorption and calculation of unit fluorescence. The same concentrations of the NPs used in the aforementioned fluorescamine screening were incubated with 4 μM protein for 1hr at 37 ˚C. The incubation was split into four aliquots. Two were labeled by fluorescamine, and subject to fluorescence measurement. One of these two labeled samples was measured directly, while the other one was filtrated by a Vivaspin 500 centrifugal filter with a MWCO of 300kD (Sartorius AG, Goettingen, Germany) and the flow-through was collected for fluorescence measurement. The flow-through contained the free unbound protein and the NP-bound protein should stay on the

membrane. The rest two aliquots were supplied with the same volume of the PBS buffer, and used for protein quantification by BCA. Similar to the fluorescamine labeled samples, one was measured without filtration and the other was filtrated and the flow-through was quantified. For both the fluorescamine labeled samples and the controls, the protein in the filtrate (the unbound protein) or the original solution (the total protein) were quantified by BCA assay after adjusting the volumes to be equal. Then the unit fluorescence was calculated by dividing the fluorescence signal by protein concentrations.

Measurement of binding affinity. PS48 at various concentrations were incubated with 400 nM of the protein for 1hr at 37 ˚C. Then, fluorescamine was added to label the protein, and fluorescence was measured by the Victor II plate reader. For each protein, the fluorescence intensity was normalized to 0-1, with the fluorescence of the control being 0 and the maximum fluorescence value as 1. Origin 8.0 was used to plot the curve and fit it with the Hills equation.

3.3: Results and Discussion

Fluorescamine screening of protein-nanoparticle (NP) interaction. Our previous work has demonstrated that fluorescamine labeling could detect protein-NP interaction.[25] Fluorescamine is a fluorogenic dye that can rapidly react with the primary amines on proteins and become fluorescent. The fast reaction rate ensures that most of the labeling events occur to the solvent accessible primary amines on protein surface. Thus, protein-NP interaction could potentially change fluorescamine labeling by blocking the solvent accessible primary amines. Or, the interaction could induce protein unfolding to expose more amines, as proved by CD measurement in our previous work.

If our labeling assay only relies on amine exposure from extensive protein unfolding, its applicability would be very limited. However, fluorescamine labeling is closely dependent on lysine reactivity, which is sensitive to solution pH as well as the microenvironment surrounding the lysine residues. While carrying out the labeling reaction in PBS buffer at pH 7.4 to match physiological conditions, we actually reduce the nucleophilicity of the amine groups on the lysine residues by protonation, yielding a reaction efficiency ~60%.[26] But, when the lysine residue is in a hydrophobic environment or surrounded by amino acid residues with negative charges, its amine group would become less protonated and more reactive to fluorescamine. Since the surrounding of lysine residues could be altered when protein conformation varies, we expect our assay should be very sensitive to subtle protein conformational change induced by NP binding and is capable of evaluating NP binding to proteins with highly diverse properties.

To prove this, in the present study, we screened the interactions of the polystyrene particles (PS48 and PS 80) or the silica particles (Si50 and Si80) with a total of 21 proteins. The PS and silica NPs have been widely applied in drug delivery or biosensor designs,[27] and thus understanding their behaviors in protein binding could help improve their effectiveness in biomedical applications. The selected proteins span a wide range of pI (isoelectric point), Mw (molecular weight) and hydrophobicity (represented by the GRAVY scores). We also included the compact, globular proteins and the intrinsically disordered proteins to see how protein tertiary structure could impact on NP interaction.[28] Most of those proteins are abundant in biological fluids,[29] and thus expected to interact with ENMs entering biological systems.

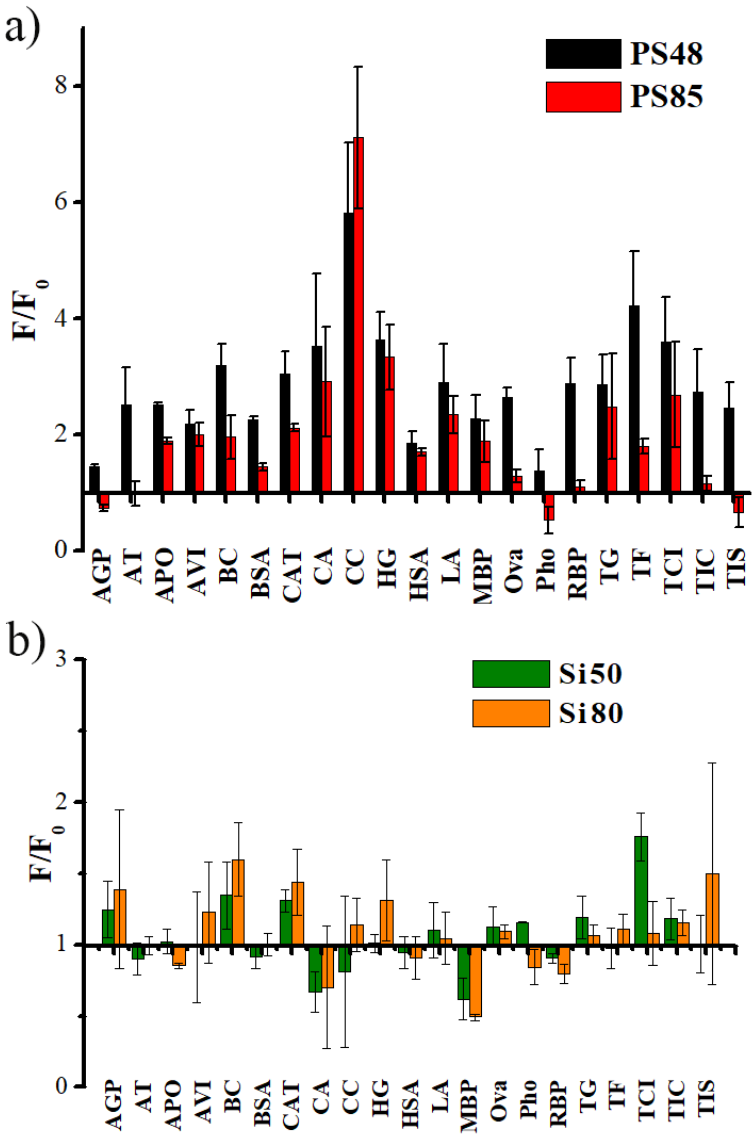


Figure 3.1. Fluorescamine labeling profiles of proteins with a) PS and b) Si nanoparticles. PS48 of 10 nM, PS85 of 3.2 nM, Si50 of 10 nM, or Si80 of 3.9 nM was incubated with 400 nM protein in the PBS buffer (10 mM phosphate at pH 7.4, 137 mM NaCl, and 2.7 mM KCl) for 1h at 37 ˚C. Then fluorescamine was added to final concentration of 1mM, and incubated for 5min before fluorescence signals being measured. Fluorescence signals change ratio (F/F_0) were calculated by dividing the signal of protein-NP pairs (F) by that of controls (F_0).

Figure 3.1 shows the changes in fluorescamine labeling induced by incubation with the PS (Figure 3.1a) or silica NPs (Figure 3.11b). F_0 and F were the fluorescence resulted from fluorescamine labeling before and after particle incubation, respectively; and the ratio of F/F_0 from different proteins incubated with the PS or silica NPs were compared. When incubated with the PS NPs, most proteins exhibited F/F_0 larger than 1, indicating that more amines were labeled by fluorescamine upon binding to the NPs. In addition, PS48 typically induced larger fluorescence change than PS80, with most of the F/F_0 larger than 2. Larger variations in fluorescence change were observed when the proteins were incubated with the silica NPs. Most of the proteins experienced less than 50% change in fluorescence (i.e. $F/F_0 > 0.5$ or < 2). Multivariate analysis of variance (MANOVA) also proved that the fluorescence changes induced by both PS48 and PS85 were significant, but those induced by Si50 and Si80 were not as significant as the PS NPs.

Dependence of protein-NP interaction on NP properties. Protein-NP interactions are affected by electrostatic interaction, Van der Waals force, solvation, Brownian motion, etc. [44] Thus, differences in fluorescamine labeling could be caused by variations in NP properties. To reveal the correlation between the fluorescamine labeling profiles and NP properties, the statistical pattern recognition tool of PCA (Principal Component Analysis) was applied to visualize NP grouping based on the fluorescence change profiles of all 21 proteins shown in Figure 1. Each repetition of one protein-nanoparticle pair was viewed as one individual observation, and the proteins were treated as the variables. After PCA, 21 original variables were reduced to two principal components (PCs), PC1 and PC2, that summarized 67.6% and 17.2% of the total variance

in the data set, respectively. The score plot of PC1 vs. PC2 (Figure 3.2a) indeed showed satisfactory grouping of the NPs: PC1 that represents the majority of the variance in our data separates the PS NPs from the silica particles based on core material difference; and PC2 separates the two PS NPs of different particle diameters. However, the two silica NPs of different diameters were not well separated (Figure 3.2a). The significant difference between the two PS particles and that between the PS and silica NPs were also confirmed by MANOVA on the two PC values, with the resultant p values for comparison between different NPs being < 0.001 except for that between the two silica NPs.

The grouping effect illustrated by the PCA scores plot can be explained by differences in surface charge and core materials of the NPs. Zeta-potential measurement showed that the PS NPs carried much more negative charges on their surface than the Si NPs, which may cause stronger attachment of the cationic residues on the protein. Besides, the base material of PS NPs is composed of benzene rings that could interact more strongly with the hydrophobic regions on protein compared to the silica NPs covered by the hydrophilic silanol groups.[45] Moreover, the hydrophobicity of the PS core can decrease the pKa values of lysine residues getting close, which leads to less protonation on the amine groups and thus stronger nucleophilicity to react with fluorescamine.[46,47]

The difference observed between the two PS particles could be mainly attributed to the variation in size. It has been reported that the sharper surface curvature on particles with a smaller diameter could induce more protein conformational change, which could lead to more primary amines to be labeled by fluorescamine.[48] Our previous work demonstrated the silica NPs with similar diameters but different surface charges could be

differentiated. The outcomes from the represent and past studies support that, particle surface charge may play a more important role than their diameter in protein-NP interaction. The small fluorescence change in the silica NPs also subsidized any signal difference between the two silica NPs of different diameters.

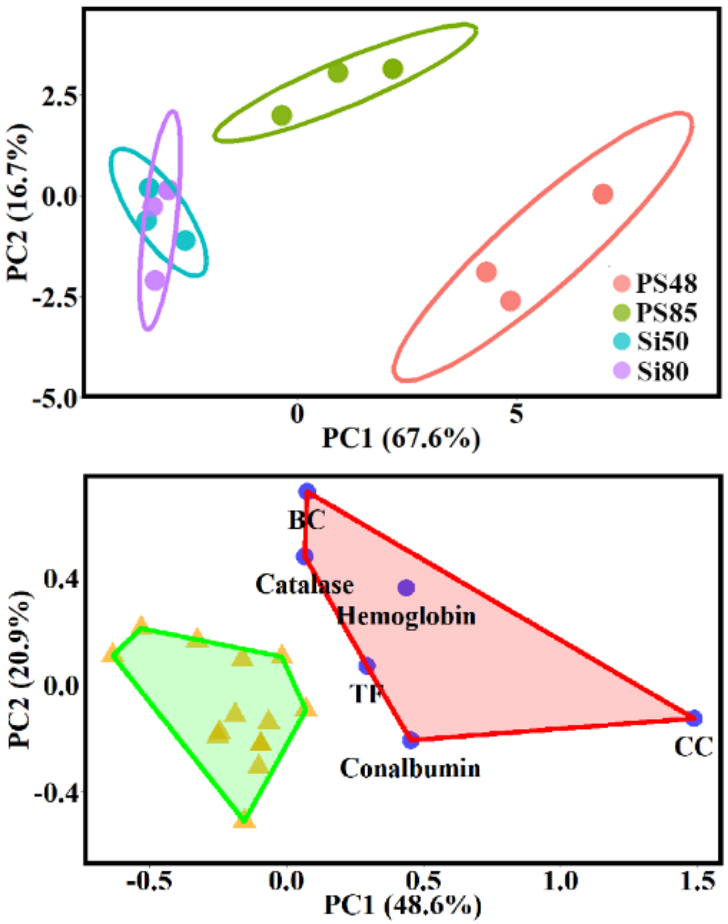


Figure 3.2. PCA scores plots obtained by a) treating the proteins as the variables to group the NPs; and b) treating the NPs as the variables to group the proteins and assigning the proteins by the boundaries determined by k-means clustering. Three repeats were evaluated. Ellipses shown were at 95% confidence.

Dependence of protein-NP interaction on protein properties. Protein-NP interaction should be affected by protein properties as well. Compared to the standard NPs, proteins are more diverse in properties because they have large variations in amino acid sequences, secondary and tertiary structures, and surface properties, increasing the difficulty in sorting out the dominant protein characteristics that govern protein-NP interaction. To explore whether the fluorescamine labeling profiles resulted from protein-NP interaction could differentiate the proteins by their Mw, pI, or hydrophobicity, we carried out PCA but with the NPs being the variables. The two PC values obtained could summarize close to 70% of the overall variance of the data (Figure 3.2b), but no clear grouping of the proteins was observed. The loading factors analysis showed that both the PS NPs had higher loadings in PC1, while the Si NPs were more decisive for PC2. Supervised clustering based on one protein property, i.e. Mw, pI, or GRAVY score did not show clear separation of the proteins, either. The poor grouping of proteins based on the properties listed in Table S1 was also verified by the poor prediction accuracies of Jackknife that only used Mw, pI, or GRAVY to build the LDA model, with < 60% of the proteins assigned accurately.

This is conceivable, because proteins are varied in many characteristic values, with no proteins only different by one or two properties while keeping the other(s) similar. In addition, the values of Mw, pI and GRAVY were calculated from the primary structures of the proteins, with no consideration of the secondary and tertiary structures nor the post translational modifications (PTMs). Since protein-NP interaction occurs between the surface of proteins and NPs, the surface properties are more important than the overall

properties calculated from protein's amino acid sequence, which could be strongly affected by protein folding status and PTM and not be easily obtained.[26,49]

We then applied the unsupervised clustering method, k-means clustering, on the PCA results with the NPs as the variables. This method can allow to classify the input dataset into k groups and visualize the relationship between samples in our dataset. Simply speaking, the distance between each data point and randomly selected centers was calculated, and the data points with the smallest square of the distance were viewed as similar data points and grouped together. Interestingly, the proteins were assigned into two groups, separated by PC 1 = 0 on the scores plot (Figure 3.2b). If examining the separation more closely, we found that the proteins grouped on the left panel of the PCA scores plot (PC1 < 0) are much more diverse in property, with no common structure features easily identified. On contrary, the group with PC1 > 0 included the proteins showing high fluorescence changes when incubated with NPs, the majority of which are with high structural flexibility. The most obvious one is beta- casein, an intrinsically disordered protein with the largest instability index among the proteins tested. It is less ordered in structure and more flexible than globular proteins like human serum albumin (HSA). Transferrin and conalbumin share high sequence similarity, and the crystal structures of both proteins show two lobes linked by flexible loops. The relative position between the lobes is changing and unstable, as reflected by the high B factors found for the amino acids (a.a.) involved in the C terminal lobe of transferrin: a higher B factor corresponds to a higher mobility of the a.a. residue. Hemoglobin and catalase can form tetramers in solution, which are considered as a flexible structure as well.

The only exception is cytochrome c that contains high alpha helix contents and low B factors in its structure. However, it is worth noting that, ligand binding to cytochrome c could induce slight conformational change and cause disruption and rearrangement of the interactions such as hydrogen bonding and salt bridging between lysine residues. Interestingly, hemoglobin and transferrin could have similar phenomena upon ligand binding. The well-known Bohr effect of hemoglobin describes that binding of oxygen could induce conformational change in hemoglobin and disrupt the salt bridges involving histamine and lysine residues, along with decrease of their pKa values and loss of protonation. Upon binding or releasing of iron ions, the conformation of transferrin changes and is accompanied with deprotonation on lysine residues. Therefore, we expect that when these proteins bind to NPs, similar effects could be induced and thus change the reactivity of the lysine residues towards fluorescamine, altering the labeling profile.

Protein adsorption on NPs and fluorescamine labeling. With better understanding on how particle or protein properties could affect fluorescamine labeling, we went on to evaluate whether the fluorescence signal could reflect the amount of proteins adsorbed by the NPs. Such a correlation could allow affinity measure using our assay. We employed centrifugation to separate the free and NP-bound proteins. The centrifugal filter has a MWCO of 300 kDa that should pass the free proteins through but keep the large NPs with the adsorbed proteins on the filter top. We chose to evaluate the adsorption of 9 proteins on PS48, including transferrin, conalbumin, catalase, cytochrome C and hemoglobulin. These proteins showed higher fluorescence change than others when

incubated with PS48. HSA was also tested to represent the proteins exhibiting medium-to-low fluorescence change.

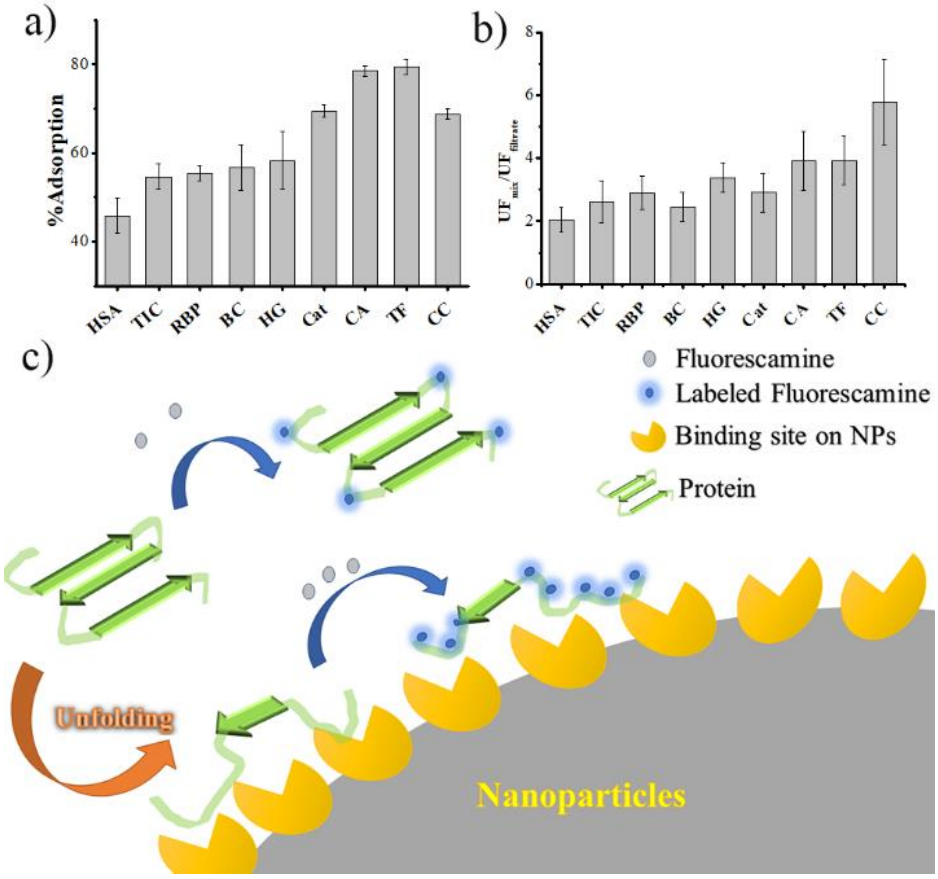


Figure 3.3. Comparison of a) %Adsorption of each protein on PS48, and b) ratio of unit fluorescence of each protein in the incubation mixture and in the filtrate. c) Illustration of fluorescence signal increase upon protein binding to the NPs. The bound protein would undergo conformational change that increases lysine reactivity with fluorescamine, and results in increased fluorescence signal.

The incubation procedure was carried out as done in the screening assay. The protein and NP mixture was added to the filter, and the free proteins eluted to the filtrate were quantified by BCA. The amount of proteins remained on the filter was then calculated and divided by the total amount of the protein used in the incubation to give out the

%Adsorption plotted in Figure 3.3a. We can see from this plot that transferrin and conalbumin led to the highest %Adsorption, with close to 80% of the protein absorbed onto the PS48 NPs upon incubation. HSA yielded the lowest %Adsorption (< 50%), agreeing with its lower fluorescence changes than the other proteins during fluorescamine screening. Pearson correlation analysis showed that, except for the three proteins with high structure flexibility – transferrin, conalbumin, and catalase, the %Adsorption holds a strong linear correlation with the fluorescence change ratio F/F_0; Pearson correlation coefficient = 0.9756, p-value = 8.8E-4). This supports that, more proteins adsorbed, more amines would be labeled by fluorescamine. The three proteins with high structure flexibility also formed another linear relationship between F/F_0 and %Adsorption that had a smaller slope. They could experience a different degree of conformational change per unit mass of the adsorbed protein, than the other proteins.

The fluorescence signal detected from the protein and NP mixture is from both the free and the NP-bound proteins. To compare the degree of labeling in these two groups of proteins for better understanding of the relationship between adsorption and resultant fluorescence, we also filtered the incubation mixture that went through fluorescamine labeling. The fluorescence in both the mixture (before filtration) and the filtrate was measured and divided by the protein concentration in the corresponding solution, termed "unit fluorescence" (UF), which should reflect the degree of fluorescamine labeling per unit mass of the protein in the free or NP-bound proteins. The UF in the protein-PS48 mixture was much higher than that in the filtrate (Figure 3.3b), i.e. UFmix/UFfree >> 1, proving that the protein adsorbed on the NPs indeed experienced a higher degree of

labeling compared to the free proteins, further proving that adsorption to the NPs changed protein conformation and activated more amine groups to be reactive with fluorescamine (Figure 3.3c).

Influence of protein or NP concentrations on the fluorescence profiles. For our assay to properly reflect protein adsorption and conformational change, the molar ratio between the protein and NPs should be carefully chosen. If the protein amount is much higher than NPs, only a small portion of the protein could bind to the NPs, resulting in a small F/F_0. Hence, influence of protein concentration on fluorescamine labeling was investigated. HSA was chosen as the model protein and the concentrations of HSA were changed from 40 nM to 8 μM, with the NP concentration fixed at either 4 or 40 nM. After incubating HSA with PS48 for 1 hr at 37 ºC, fluorescamine was added to the mixture for protein labeling. Agreeing with our assumption, the fluorescence change ratios (F/F_0) decreased with increasing protein concentration for both NP concentrations, indicating binding saturation at higher protein concentrations.

A clearer relationship of fluorescamine labeling and adsorption saturation can be viewed by the normalized fluorescence change $(F-F_0)/(F_{max}-F_{min})$, with F_{max} and F_{min} being the highest or lowest fluorescence signals in the dataset, respectively (Figure 3.4a). The value of $(F-F_0)$ should represent the difference in the number of labeled amines before and after interaction with the NPs, and the normalization eliminates the random difference between data sets for clearer comparison. At lower protein concentrations between 40-200 nM, fluorescence increased gradually with protein concentrations (Figure 3.4a). This suggested that the surface areas of PS48 were enough to accommodate all HSA molecules,

and with increasing protein concentrations, more proteins were adsorbed and experienced conformational change. When the concentrations of HSA increased further, such an increase slowed down, indicating that the proportion of HSA interacting with PS48 decreased, with fewer adsorption sites available. With the concentration of HSA higher than 2 μM, the curve reached a plateau if the PS48 concentration was 4 nM (500:1 molar ratio of HSA:NPs, Fig. 4a). At this point, all active adsorption sites of PS48 were occupied by HSA and no more proteins could interact with PS48. To monitor the adsorption of HSA on PS48 NPs, size information of NPs was obtained via NTA under different protein/NPs ratios (Figure 3.4b). We can see the diameter of NPs increased with protein concentration and reached a plateau at [HSA] = 2 μM. With a higher PS48 concentration of 40 nM, the fluorescence signal kept increasing, although with a lower slope. It was expected that a plateau would be reached with HSA concentration increasing beyond 20 μM, but this concentration was outside of the linear range of fluorescamine labeling for primary amines, and was not tested. The results indicate that the fluorescence signal should reflect the proportion of protein adsorbed and can possibly be used for evaluation of binding affinity if keeping the ratio of protein to NP below the saturation level.

NP could also influence fluorescence measurement. Similar to other organic fluorophores, the fluorescence of fluorescamine can be influenced by NPs, by either the inner filter effect (IFE) or the near-field effects including dynamic or static quenching, surface enhancement, and quantum yield variation, which depends on the position or distance of fluorescamine relative to the NPs.[30] In fact, these are the same concerns shared by using other optical methods to assess protein-NP interaction, especially when analyzing

the luminescent NPs or those inducing strong light absorption or scattering. To evaluate the impact from NPs to fluorescence measurement, we incubated 400 nM of the fluorescamine-labeled HSA with different concentrations of the PS NPs, and indeed observed some quenching of the fluorescence from the NPs: the fluorescence change ratio F/F_0 (with F being the fluorescence at the presence of NPs) increased linearly with particle concentration when [NP] > 5 nM. The bigger PS NPs showed larger influence, by increasing the fluorescence for ~ 50% at [PS85] = 10 nM; while < 20% change in fluorescence was observed for PS40 at 10 nM. Such an increase may be attributed to inter-particle scattering. In our screening, we kept the NP concentrations at 10 nM for PS48 or 3.2 nM for PS85, to keep the impact from NPs low (increasing the fluorescence by 1.1 or 1.2 folds) so that the measured fluorescence is only correlated to protein binding. Correction can also be applied for situations that large influence from NPs is observed taking advantage of the linear relationship between fluorescence change and NP concentration.

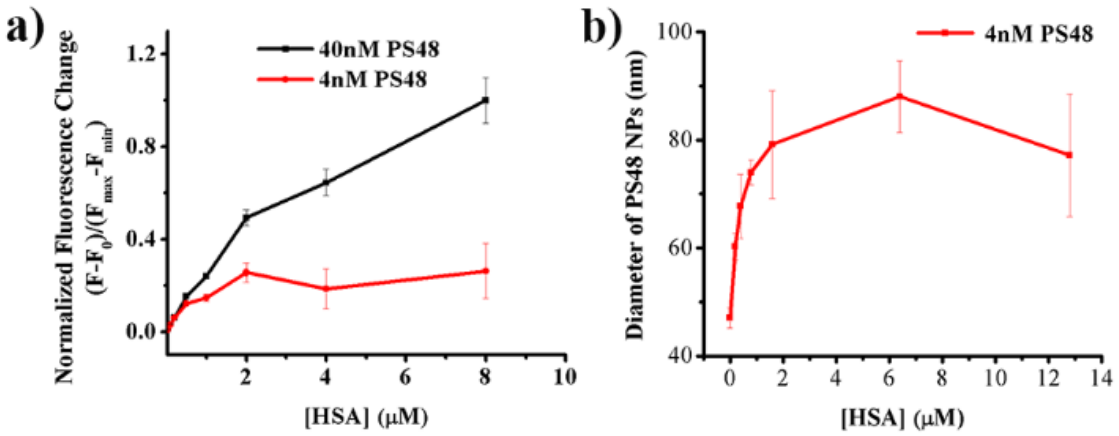


Figure 3.4. a) Normalized fluorescence change of HSA incubated with PS48 at 40 or 4 nM. Fluorescamine was added at a final concentration of 1mM. The change between fluorescence of

HSA incubated with PS48 NPs (F) and that of HSA itself (F_0) was normalized to 0-1 using F_{max} and F_{min} in the dataset for better visualization. b) Hydrodynamic diameter of PS48 NPs changed along with HSA concentrations. Sizes were measured by Nanosight™ (Malvern Instruments) using the Nanaoparticle Tracking Analysis software, after PS48 NPs were incubated with different concentrations of HSA (0-12.8 μM) in PBS buffer for 1 hr at 37 °C.

Assessment of the affinity of protein-NP interaction. Fluorescamine labeling is sensitive to the reactivity of primary amines on the protein which could be strongly impacted by protein conformational change. Protein-NP interaction is likely to induce protein conformational change, making our assay valuable in interaction assessment. Our above results prove that if the appropriate protein and NP concentrations are used in our assay, the fluorescence change can be viewed as the signal of binding. Thus, the ratio can be plotted against NP concentration to obtain the binding curve of protein-NP interaction. We chose 13 proteins to explore their binding curves with PS48. These protein-particle pairs showed higher fluorescence changes than others. The fluorescence of avidin, aprotinin, and conalbumin increased with increasing PS48 concentration (from 0 – 6 nM) and reached a plateau at different [PS48].

By plotting the fluorescence change at any NP concentration, i.e. ($F - F_0$) after normalization with the maximum change at the plateau, i.e. ($F_{max} - F_0$), verse NP concentration, we obtained the binding curves that can be fitted with the Hill equation (Figure 3.5). The K_d (the macroscopic dissociation constant) was also calculated and listed in Table 3. In this fitting, the protein was viewed as the receptor, and the binding site on NPs were treated as the ligand, assuming that each NP were identical and carried the same number of binding sites. By monitoring the fluorescence resulted from fluorescamine labeling of the protein upon binding to the NPs, we measured the change in the "receptor"

when interacting with the ligand, matching well with the receptor-ligand binding model illustrated by the Hill equation.

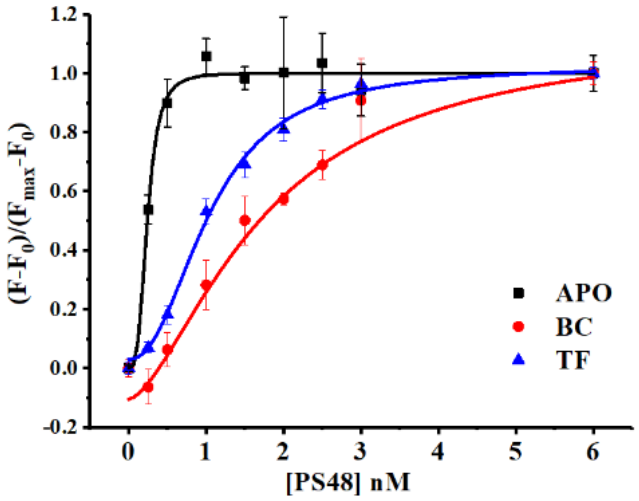


Figure 3.5. Normalized fluorescence of aprotinin, beta casein, and transferrin, under different PS48 concentrations. All proteins were at 400nM. Fluorescamine was added to a final concentration of 1 mM after proteins were incubated with PS48 in PBS for 1 hr at 37°C. For each protein, fluorescence signals were normalized to that with highest PS48 concentrations. Hills fitting was conducted with Origin 8.0 (OriginLab Corp.).

Table 3.1. Dissociation constants (K_d) calculated from the binding curves. All proteins were in 400nM.

Protein	K_d(μM)	Protein	K_d(μM)
APO	0.24±0.16	HG	0.81±0.24
AT	0.45±0.23	LA	0.61±0.06
AVI	1.07±0.25	MBP	0.96±0.12
BC	1.75±0.33	OVA	1.03±0.07
BSA	1.43±0.56	RBP	1.57±0.24

CA	0.63±0.08	TCI	0.39±0.23
Cat	0.91±0.05	TF	1.05±0.07
CC	0.95±0.07	TIC	0.56±0.09
HSA	1.29±0.07		

From Table 3.1, we noticed that, the K_d values were not simply related to the size, pI, or hydrophobicity of the proteins, agreeing with the above PCA results (Table 3.1). Aprotinin has the smallest K_d among all proteins tested. It is a basic protein and carries positive charges to have electrostatic interaction with the negatively charged NPs. But the K_d values of two acidic proteins, antitrypsin and TCI, are also smaller than most of the other proteins. Conalbumin and transferrin are proteins with high homology, but different K_d values were obtained from our screening. It is worth noting that proteins with higher flexibilities, like beta-casein (an IDP protein), catalase, and transferrin, was proved to be weak binders with larger K_d values, although they showed large fluorescence change ratio in the screening at a fixed protein concentration (Figure 3.1). This indicates that a higher degree of conformal change in protein does not directly reflect stronger binding affinity. However, our assay is a convenient tool to assess both conformational change and binding affinity.

3.4: Conclusions

The present work demonstrated that fluorescamine labelling can be applied to evaluate binding between proteins with diverse properties and NPs, and the fluorescence should reflect protein conformational change caused by interaction with NPs. The impacts

on the fluorescent signals from protein or particle concentrations were also evaluated to reveal the suitable protein and NP concentration ranges for affinity measurement using our assay. The study proves that, our method is convenient and rapid for assessment of protein-NP interaction, and fluorescamine labeling should reflect protein conformational change upon NP adsorption because of the high sensitivity of amine reactivity with the surrounding environment. Moreover, our assay can be used for evaluation of binding affinity using simple instrumentation. It is valuable to screen for proteins with higher binding affinities to NPs but experiencing lower conformational change which are useful for particle functionalization in biomedical research. On the other hand, conformational change may alter protein functions. For example, the metalloproteins screened in our study all underwent obvious conformational changes upon interacting with the negatively charged nanoparticles, indicating the potential adverse impact on protein functions by the charged PS nanoparticles. Follow-up studies on impact of NPs on the functions of such proteins are needed for toxicity evaluation.

References

(1) Walkey, C. D.; Olsen, J. B.; Song, F.; Liu, R.; Guo, H.; Olsen, D. W. H.; Cohen, Y.; Emili, A.; Chan, W. C. Protein corona fingerprinting predicts the cellular interaction of gold and silver nanoparticles. *ACS nano* **2014**, *8*, 2439-2455.

(2) Lundqvist, M.; Stigler, J.; Elia, G.; Lynch, I.; Cedervall, T.; Dawson, K. A. Nanoparticle size and surface properties determine the protein corona with possible implications for biological impacts. *Proceedings of the National Academy of Sciences* **2008**.

(3) Monopoli, M. P.; Walczyk, D.; Campbell, A.; Elia, G.; Lynch, I.; Baldelli Bombelli, F.; Dawson, K. A. 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*, 2525-2534.

(4) Casals, E.; Pfaller, T.; Duschl, A.; Oostingh, G. J.; Puntes, V. F. Hardening of the nanoparticle–protein corona in metal (Au, Ag) and oxide (Fe3O4, CoO, and CeO2) nanoparticles. *Small* **2011**, *7*, 3479-3486.

(5) Cedervall, T.; Lynch, I.; Lindman, S.; Berggård, T.; Thulin, E.; Nilsson, H.; Dawson, K. A.; Linse, S. Understanding the nanoparticle–protein corona using methods to quantify exchange rates and affinities of proteins for nanoparticles. *Proceedings of the National Academy of Sciences* **2007**, *104*, 2050-2055.

(6) Tenzer, S.; Docter, D.; Kuharev, J.; Musyanovych, A.; Fetz, V.; Hecht, R.; Schlenk, F.; Fischer, D.; Kiouptsi, K.; Reinhardt, C. Rapid formation of plasma protein corona critically affects nanoparticle pathophysiology. *Nature nanotechnology* **2013**, *8*, 772.

(7) Yan, Y.; Gause, K. T.; Kamphuis, M. M.; Ang, C.-S.; O'Brien-Simpson, N. M.; Lenzo, J. C.; Reynolds, E. C.; Nice, E. C.; Caruso, F. Differential roles of the protein corona in the cellular uptake of nanoporous polymer particles by monocyte and macrophage cell lines. *ACS nano* **2013**, *7*, 10960-10970.

(8) Lesniak, A.; Fenaroli, F.; Monopoli, M. P.; Åberg, C.; Dawson, K. A.; Salvati, A. Effects of the presence or absence of a protein corona on silica nanoparticle uptake and impact on cells. *ACS nano* **2012**, *6*, 5845-5857.

(9) Gebauer, J. S.; Malissek, M.; Simon, S.; Knauer, S. K.; Maskos, M.; Stauber, R. H.; Peukert, W.; Treuel, L. Impact of the nanoparticle–protein corona on colloidal stability and protein structure. *Langmuir* **2012**, *28*, 9673-9679.

(10) Singh, R.; Norret, M.; House, M. J.; Galabura, Y.; Bradshaw, M.; Ho, D.; Woodward, R. C.; St Pierre, T. G.; Luzinov, I.; Smith, N. M.; Lim, L. Y.; Iyer, K. S. Dose-Dependent Therapeutic Distinction between Active and Passive Targeting Revealed Using Transferrin-Coated PGMA Nanoparticles. *Small* **2016**, *12*, 351-359.

(11) Kelly, P. M.; Åberg, C.; Polo, E.; O'connell, A.; Cookman, J.; Fallon, J.; Krpetić, Ž.; Dawson, K. A. Mapping protein binding sites on the biomolecular corona of nanoparticles. *Nature nanotechnology* **2015**, *10*, 472.

(12) Salvati, A.; Pitek, A. S.; Monopoli, M. P.; Prapainop, K.; Bombelli, F. B.; Hristov, D. R.; Kelly, P. M.; Åberg, C.; Mahon, E.; Dawson, K. A. Transferrin-functionalized nanoparticles lose their targeting capabilities when a biomolecule corona adsorbs on the surface. *Nature nanotechnology* **2013**, *8*, 137.

(13) Wang, J.; Jensen, U. B.; Jensen, G. V.; Shipovskov, S.; Balakrishnan, V. S.; Otzen, D.; Pedersen, J. S.; Besenbacher, F.; Sutherland, D. S. Soft interactions at nanoparticles alter protein function and conformation in a size dependent manner. *Nano letters* **2011**, *11*, 4985-4991.

(14) Deng, Z. J.; Liang, M.; Monteiro, M.; Toth, I.; Minchin, R. F. Nanoparticle-induced unfolding of fibrinogen promotes Mac-1 receptor activation and inflammation. *Nature nanotechnology* **2011**, *6*, 39.

(15) Wang, H.; Duan, Y.; Zhong, W. ZrO2 nanofiber as a versatile tool for protein analysis. *ACS applied materials & interfaces* **2015**, *7*, 26414-26420.

(16) Baptista, P.; Pereira, E.; Eaton, P.; Doria, G.; Miranda, A.; Gomes, I.; Quaresma, P.; Franco, R. Gold nanoparticles for the development of clinical diagnosis methods. *Analytical and bioanalytical chemistry* **2008**, *391*, 943-950.

(17) Chen, J.; Rempel, D. L.; Gau, B. C.; Gross, M. L. Fast photochemical oxidation of proteins and mass spectrometry follow submillisecond protein folding at the amino-acid level. *Journal of the American Chemical Society* **2012**, *134*, 18724-18731.

(18) Konermann, L.; Pan, J.; Liu, Y. H. ChemInform Abstract: Hydrogen Exchange Mass Spectrometry for Studying Protein Structure and Dynamics. *Cheminform* **2011**, *40*, 1224-1234.

(19) Sinz, A. Divide and conquer: cleavable cross-linkers to study protein conformation and protein-protein interactions. *Analytical and bioanalytical chemistry* **2017**, *409*, 33-44.

(20) Zeng, S.; Yu-ming, M. H.; Chia-en, A. C.; Zhong, W. Protein binding for detection of small changes on a nanoparticle surface. *Analyst* **2014**, *139*, 1364-1371.

(21) Li, N.; Zeng, S.; He, L.; Zhong, W. Probing nanoparticle− protein interaction by capillary electrophoresis. *Analytical chemistry* **2010**, *82*, 7460-7466.

(22) Hoshino, Y.; Koide, H.; Furuya, K.; Lee, S. H.; Kodama, T.; Kanazawa, H.; Oku, N.; Shea, K. J. The rational design of a synthetic polymer nanoparticle that neutralizes a toxic peptide in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 33-38.

(23) Yonamine, Y.; Hoshino, Y.; Shea, K. J. ELISA-mimic screen for synthetic polymer nanoparticles with high affinity to target proteins. *Biomacromolecules* **2012**, *13*, 2952-2957.

(24) Zhang, X.; Chytil, P.; Etrych, T. s.; Liu, W.; Rodrigues, L.; Winter, G.; Filippov, S. K.; Papadakis, C. M. Binding of HSA to Macromolecular p HPMA Based Nanoparticles for Drug Delivery: An Investigation Using Fluorescence Methods. *Langmuir* **2018**, *34*, 7998-8006.

(25) Ashby, J.; Duan, Y.; Ligans, E.; Tamsi, M.; Zhong, W. High-Throughput Profiling of Nanoparticle–Protein Interactions by Fluorescamine Labeling. *Analytical chemistry* **2015**, *87*, 2213-2219.

(26) Muz, M.; Ost, N.; Kuhne, R.; Schuurmann, G.; Brack, W.; Krauss, M. Nontargeted detection and identification of (aromatic) amines in environmental samples based on diagnostic derivatization and LC-high resolution mass spectrometry. *Chemosphere* **2017**, *166*, 300-310.

(27) Slowing, I. I.; Vivero-Escoto, J. L.; Wu, C.-W.; Lin, V. S.-Y. Mesoporous silica nanoparticles as controlled release drug delivery and gene transfection carriers. *Advanced drug delivery reviews* **2008**, *60*, 1278-1288.

(28) M Redwan, E.; Xue, B.; A Almehdar, H.; N Uversky, V. Disorder in milk proteins: Caseins, intrinsically disordered colloids. *Current Protein and Peptide Science* **2015**, *16*, 228-242.

(29) Stempfer, R.; Kubicek, M.; Lang, I. M.; Christa, N.; Gerner, C. Quantitative assessment of human serum high-abundance protein depletion. *Electrophoresis* **2010**, *29*, 4316-4323.

(30) Zhang, D.; Nettles, C. B. A Generalized Model on the Effects of Nanoparticles on Fluorophore Fluorescence in Solution. *Journal of Physical Chemistry C* **2015**, *119*, 7941-7948.

CHAPTER 4: Structure-Activity Model for Prediction of Protein Corona Formation on Nanomaterials Using Fluorescamine Labeling Data

4.1: Introduction

The diagnostic and therapeutic applications of engineered nanomaterials(ENMs) have been in great prosperity since the first of its kind being proved for clinical usage 23 years ago.[1-4] However, correlations between properties of ENMs and their biological outcomes are still not well understood, hindering the further optimization and application of ENMs in biological systems.[5,6] One important reason is that ENMs will adsorb proteins on their surface and form "protein corona", immediately after entering biological fluids.[7-9] Protein corona will change the synthetic identity of ENMs, and could be seen by other biological molecules and cells.[6] As a result, the fate of ENMs, including kinetics, stability, distribution, and toxicity, would also be changed by protein corona.[7,10-14] Despite its importance, identifying the composition of protein corona still remains challenging and time consuming, due to the complexity of the biological matrix and the dynamic property of protein corona. As the standard and most widely used method, mass spectrometry usually needs sufficient separation and robust sample preparation, which is not suitable for testing the protein corona composition rapidly and routinely.[15,16]

An alternative way is to establish the correlations between the properties of known ENMs, which can be measured rapidly and routinely, and the composition of the protein corona of ENMs.[17] Based on these correlations, the protein corona composition of new ENMs could be predicted by feeding the model with the properties of the new ENMs. A similar paradigm has been used to predict biological outcomes of ENMs using their

physicochemical properties and protein corona fingerprints as descriptors.[18-22] However, not much progress has been made to predict protein corona composition, so far. There was only one quantitative prediction model for this purpose, based on the structure activity relationship (SAR), in which the size and charge of ENMs were used as their descriptors.[17] Although this model could provide reasonable prediction about the identities of proteins enriched in the corona, the descriptors of the ENMs were too few to make significant contribution to the model compared to the much more diverse protein properties, which would limit the model's prediction ability for new ENMs. In other research, the physicochemical properties of ENMs, including size, surface charge, and functional groups, were precisely controlled; and comparison of protein corona composition was done by changing only one property but keeping the others the same.[23-28] Although these works succeeded in drawing certain conclusions about the preference of some proteins over selected ENMs' properties, the correlations were not quantitative, and some were even contradictory to each other, which made the correlations unreliable for prediction. One common drawback for these prediction models is that the physicochemical properties of the ENMs are not enough to discriminate the protein interaction behaviors between ENMs. The binding of ENMs to proteins is based on the cooperative effect of multiple factors, including various types of molecular interactions distributed on the curved surface of ENMs. Moreover, proteins could undergo conformational changes during interactions with ENMs, which makes the binding more difficult to evaluate.[29-31] Various kinds of novel descriptors for ENMs, including compositions, intrinsic/extrinsic properties, and even topology, have been proposed and applied for predicting interactions with small molecules

recently; however, none of them could directly solve the problem with protein.[5,32,33] As a result, new descriptors for ENMs are critical for the improvement of these models to predict protein-ENM interactions.

One possible solution is to use the interactions between one single protein and the ENM as the descriptors, which can be obtained from different proteins. The formation of protein corona is based on the equilibria between proteins with different binding affinities and kinetics, and each protein can be used as an independent probe to characterize the binding behavior of the ENM. A similar methodology has been used in the biological surface adsorption index (BSAI) of ENMs, in which the adsorption processes of small molecules were measured and used to represent molecular forces involved in ENMs interactions.[34,35] However, no successful endeavor has been made to use proteins as the probes, partially due to the lack of methods that can screen such interactions rapidly and robustly. Recently, a high throughput method using fluorescamine labeling to screen the interactions between proteins and ENMs has been developed in our group, and the fluorescence profiles obtained can not only differentiate either proteins or ENMs by their properties, but also quantify the binding affinity.[36,37] The fluorescence profiles obtained with this screening method for each ENM should represent the protein binding behaviors of the ENM, and thus can be viewed as a new set of feasible descriptors of the ENMs to be used for the prediction model. Herein, the present work tests this hypothesis by obtaining the fluorescence profiles of various kinds of ENMs binding to several standard proteins with fluorescamine labeling, and employing the resultant profiles as the descriptors for ENMs to build the SAR model for prediction of protein corona compositions. Our result

proves that these descriptors can greatly improve the performance of the prediction model, compared to the benchmark work.

4.2: Materials and Methods

General. Proteins, fluorescamine, SYPRO Orange, urea, ammonium bicarbonate, 1,4-dithiothreitol (DTT), iodoacetamide (IAA), formic acid (FA), and trypan blue were purchased from Sigma-Aldrich (St. Louis, MO). ENMs were obtained from the HSPH-NIEHS Nanosafety Center. Ultrapure water with electric resistance > 18.2MΩ was produced in-house, by the Milipore Milli-Q water purification system (Billerica, MA).

Fluorescamine labeling at different temperatures. The ENMs at the concentration of 0.1 mg/ml were mixed with 0.2 mg/ml protein in 1×PBS buffer (pH 7.4) and incubated at 37 ^{0}C for 1 hr. The mixture was then incubated at 37, 60, or 80 ^{0}C for 5 minutes. After quickly cooled to room temperature, an aliquot of fluorescamine in acetone was added to the solution at a final concentration of 1 mM. The solution was then diluted by 10 times with 1xPBS, and its fluorescence was measured in the Victor II plate reader.

Thermal stability screening. The same incubation step as stated above was conducted by mixing 0.1 mg/ml of ENMs and 0.2 mg/ml of the protein at 37 ^{0}C for 1 hr. Next, SYPRO Orange dye was added into the solution at a final concentration of 4×. Then, the mixture was transferred into the CFX Real-Time PCR instrument (Bio-Rad) and subject to a temperature gradient increasing from 37 to 98 ^{0}C with an incubation period of 20s at each temperature before recording the fluorescence intensity. The laser for excitation was 488 nm, and the range of emission filter was 515-545 nm.

Serum protein corona identification. The ENMs at 0.1 mg/ml in 1x PBS was mixed with the same volume of human serum, and the mixture was incubated for 1 hr at 37 ^{0}C. Centrifugation at 15,000 ×g for 15 minutes was used to pellet the ENMs. After washing the ENMs by 1xPBS twice, 8 M urea in 50 mM ammonium bicarbonate was used to re-suspend ENMs and the proteins. DTT was added into the solution at a final concentration of 5 mM and was incubated for 40 minutes at 56 ^{0}C. The solution was cooled down to room temperature before 10 mM IAA was added. Then 50 mM of ammonium bicarbonate was used to dilute the solution by 8 times, and trypsin was added to digestion the proteins at a trypsin: protein mass ratio of 1:50. The ENMs were removed by centrifugation at 15,000 ×g for 15 minutes. After being lyophilized and desalted, the resultant peptides sample was injected into the Waters CapLC system, which was connected to a Finnigan LTQ MS with a nano-ESI ion source. Collision induced dissociation (CID) was used for fragmentation, and the mass range was set to 300-2000 Da. MSGF+ was used to search against the human proteome downloaded from UniProt. Reversely ordered protein sequences were used as decoys, and false discovery rate (FDR) was set to 0.1%. Spectra counting (SC), as a label-free semi-quantitative method, was used to calculate the relative abundance (RA) of the corresponding protein i by Equation 1:

$$\%RA_i = \frac{SC_i}{\sum SC} \times 100\% \quad (1)$$

The similarity or overlap of protein corona between two ENMs (a, b) was calculated by Equation 2:

$$Similarity = \sum \frac{\min(\%RA_a, \%RA_b)}{\frac{1}{2} \times (\%RA_a + \%RA_b)} \quad (2)$$

Cytotoxicity test. Human T lymphoblast (CEM) cells were cultured in RPMI 1640 media with 10% FBS. For different purposes, cells were pelleted down by brief centrifugation and fresh media with or without DBS were used to resuspend cells to a concentration of $1x10^6$/ml. Aliquots of these cells were transferred to a 24-well cell culture plate, in which ENMs were added to a final concentration of 50 μg/ml. After 24-hr incubation, trypan blue was used to quantify the number of cells and the percentage of living cells.

Correlation and clustering. The fluorescamine labeling intensity of each protein-ENM mixture was divided by the intensity of the corresponding protein itself for the normalization, after removing the background. For each protein, the normalized intensities with all ENMs tested were put into one individual array named X. The relative abundance of each protein in the protein corona of all ENMs tested were combined into another array named Y. Correlation coefficient was calculated by Equation 3:

$$Correl(X,Y) = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}} \quad (3)$$

Hierarchical classification was performed by R. Normalized fluorescence in fluorescamine labeling under heating pressure and the thermal stability signal matrix were used as the independent variables, respectively. The bottom-up manner based agglomerative clustering, complete linkage, was used to find out similar clusters.

SAR model for serum protein corona prediction. The isoelectric point (pI), molecular weight (Mw)**,** grand average of hydropathy (GRAVY), percentage of negative/positive/aromatic amino acids, and the relative abundance of the proteins identified in the protein corona were used as the descriptors for each protein. The

fluorescence profiles of each ENM with all standard proteins measured at 37, 60, 80 ^{0}C were used as descriptors for ENMs. The %RA of each protein in the corona was compared to that in the serum, and the change, i.e. the enrichment degree, was used as the target for the regression model. For the classification model, the proteins with enrichment (i.e. higher %RA in the protein corona compared to that in the serum) were considered as the corona protein (i.e. considered as "positive"). The data was randomly split into two sets: training set containing 90% of the data, and testing set with the other 10%. Random forest was used for either classification or regression. The running environment included python 2.7, scikit-learn v0.19.1, NumPy v1.15.0, and Pandas v0.23.4. The minimum number of samples in each leaf node was set to 3. One thousand trees were grown for the bootstrap. A 5-fold cross validation was performed.

4.3: Results and Discussion

Fluorescamine labeling at different temperatures. The fluorescamine labeling method developed in our group has been proved to not only differentiate ENMs by their core composition, size, and surface modification, but also quantify the binding affinities of various proteins to the ENMs tested.[36,37] This method is based on the rapid reaction of fluorescamine to the surface lysine residues, and the protein-ENM interaction would change the number of accessible lysine residues on protein surface, due to blockage or protein unfolding, giving out varied fluorescence upon fluorescamine labeling. Thus, rapid screening of protein-ENM interaction can be achieved, providing the characteristic interaction profile. Using this method, we screened the fluorescence profiles between different standard proteins and various ENMs obtained from the HSPH-NIEHS Nanosafety

Center. These ENMs include metal and metal oxide nanoparticles (NPs), cellulose nanofiber (NF) and nanocrystal (NC), citrated silver and gold NPs, as well as silver dosed silica NPs. The proteins we chose covered a wide range of the protein property space, including molecular weight (Mw), isoelectric point (pI), and hydrophobicity (GRAVY). For example, two representative proteins, lysozyme and lactalbumin, have similar MW (14.3 and 14.1 kDa), but different pI values (11.3 and 4.5).

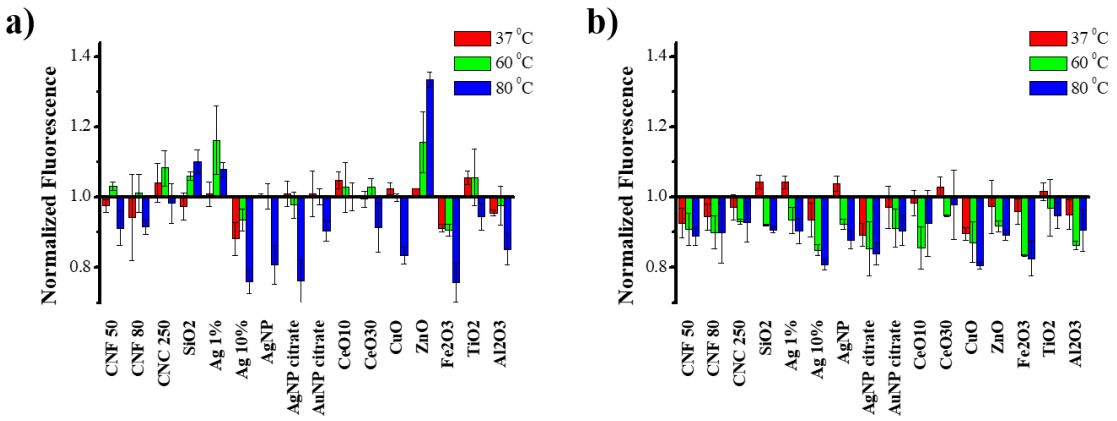


Figure 4.1. Normalized fluorescence after fluorescamine labeling for (a) lysozyme and (b) lactalbumin. The fluorescence signals of proteins alone at different temperatures were used as controls, and signals of protein-ENMs were divided by the controls for normalization.

At first, proteins were incubated with ENMs in a 2:1 mass ratio at 37 ^{0}C for 1 hr. Since there was only one single protein used at one time and no exchange with other proteins, 1 hr incubation was deemed sufficient for the formation of protein corona.[38] After being labeled by fluorescamine, the fluorescence intensity of protein-ENM mixture was measured and divided by the intensity of each protein control, for normalization. The background and the inner filter effect (IFE) of ENMs on the fluorescence were measured by the pre-labeled human serum albumin (HSA), and they were removed from all measurements before normalization. The fluorescence change should be mainly caused by

either the blockage of the protein surface upon binding to ENMs (usually indicated by a decrease for the fluorescence), or the unfolding induced by ENMs (usually indicated by an increase for the fluorescence). Fluorescence profiles of lysozyme and lactalbumin showed various patterns for different proteins and ENMs, but these changes were not big enough (Figure 4.1), probably due to the weak and transient interactions between proteins and these ENMs. In previous work, we have discovered that ENMs with strong negative zeta potential and hydrophobic core materials, such as polystyrene nanoparticles (PS NPs), could induce higher change on fluorescamine labeling signals, while ENMs with moderate surface charges and hydrophilic core materials, including silica nanoparticles, only induced week changes.[36] ENMs used in this study did not show high surface charges nor have ligands with hydrophobicity, and thus moderate changes were expected.

To enlarge signal changes, an external pressure, i.e. heating, was applied to proteins to induce protein unfolding, which could expose more amino groups to the protein surface for labeling and increase the signal. This strategy has been successfully used to study small molecule/protein binding, in which heating was able to enlarge the difference between the free proteins and ligand-bound proteins.[39] If the ligand binds to the protein, the hydrophobic or electric interactions will change the enthalpy of the complex, which would shift the thermal stability of protein. The interaction between protein and ENM is similar as ligand-protein binding, but with more complicated outcomes. The overall effect will be different structural changes of proteins at higher temperatures, as well as larger changes on fluorescamine labeling sites and fluorescence signals.

The temperatures of 60 and 80 ^{0}C were chosen for heating pressure, because the denaturation temperatures for most of proteins are in this range, leading to the largest degrees of unfolding.[40] After heating for 5 minutes, the protein-ENMs mixture was cooled to room temperature, and fluorescamine was added into it immediately. The protein incubated alone at each temperature was used as the control, and the fluorescence intensity of the protein with each ENM was normalized against that of this control value, representing the labeling change induced by binding to the ENM. As expected, the fluorescence changes were enlarged for most of the protein-ENM pairs (Figure 4.1). Moreover, the fluorescence changes were different among proteins. Two proteins, lysozyme and lactalbumin, were used for illustration. At higher temperatures, the normalized fluorescence intensities for lactalbumin with all ENMs tested were below 1, which indicated the binding between lactalbumin and ENMs would block the protein surface and decrease the number of lysine residues for fluorescamine labeling. Furthermore, binding to ENMs seemed to stabilize lactalbumin, because the diminished fluorescence at higher temperatures also indicated less unfolding.

In contrast, lysozyme showed both large increases and decreases in the normalized fluorescence, which indicated the binding would either stabilize or destabilize lysozyme, depending on the types of ENMs. The different pI values might be one of the reasons for these variances: lysozyme is a basic protein and carries positive charges in the neutral pH, making it easier to be interfered by these negatively charged ENMs used in our study. To illustrate the impact of ENMs' negative surface charges on the fluorescence profiles of lysozyme, the NPs with similarities in their core compositions, including SiO_2, 1%

Ag/SiO_2, 10% Ag/SiO_2, and Ag NPs, were taken as examples. We found that the silica NP and silica NP with a lower percent content (1%) of Ag exhibited comparable trend of change in fluorescamine labeling, i.e. increasing fluorescence at higher temperatures; while the silica NP with a higher percent content (10%) of Ag displayed similar patterns as the Ag NP, reducing fluorescamine labeling at higher temperatures (Figure 4.1a). Silica NP is hydrophilic and has negatively charged silanol groups in the neutral 1xPBS solution.[41,42] The silver decoration reduced its surface negative charges, and thus, silica NPs with higher silver components would form weaker interactions with the positively charged lysozyme, causing lower degrees of unfolding and destabilization.

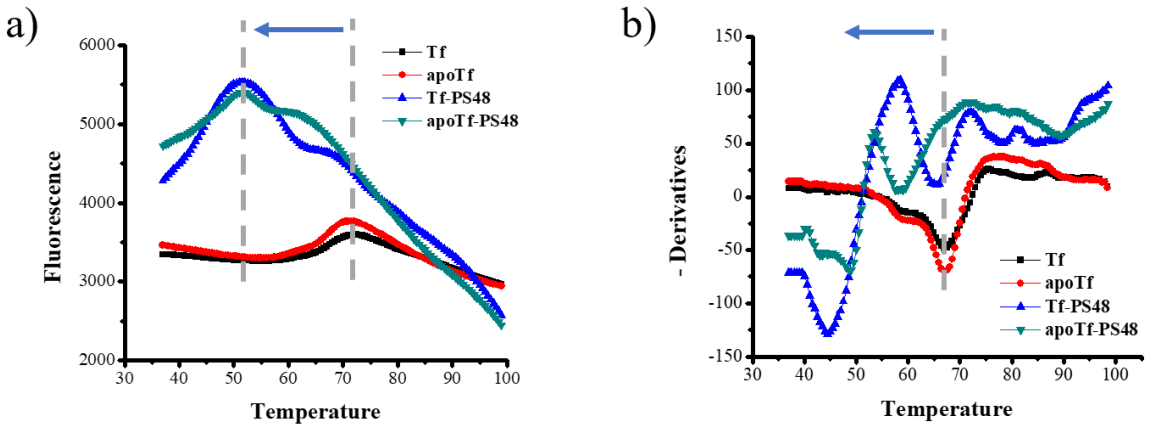


Figure 4.2. (a) Fluorescence changes of SYPRO Orange added into Tf/apoTf, with or without PS48 NPs, at various temperatures. The peaks were marked with gray dash lines. Based on the fluorescence signal plot, the negative derivative for (b) Tf/aoTf with PS48 NP was plotted. T_m in (b) was marked with a gray dash line.

Thermal stability screening for protein-ENMs. We hypothesized that more differences in fluorescamine labeling detected at higher temperatures was due to ENM incubation changing protein thermal stability. To verify this hypothesis, we employed the method of SYPRO Orange labeling, which has been widely used to explore protein

conformational changes but seldom applied to study protein-ENM interaction.[39,43-46] SYPRO Orange can bind to the hydrophobic regions of the protein, which increases its quantum yield. Protein unfolding could expose more hydrophobic areas to be bound by SYPRO Orange and result in increasing fluorescence. The fluorescence would reach a peak and then start to decrease, because of protein aggregation that reduces the numbers of labeling sites for SYPRO Orange. The temperature with the highest increasing rate (ΔF) of fluorescence intensity is termed as the melting temperatures, T_m, which can be used to quantify the thermal stability of the protein. Since this method monitors the shape, i.e. the shift of the T_m, instead of directly quantifying the fluorescence intensity, it can eliminate any potential influence from the ENMs to fluorescence intensity, a general pitfall for any fluorescence-based screening method applied on ENMs.[47]

We first evaluated the SYPRO Orange labeling profile on transferrin (Tf) when incubated with PS NPs, because Tf exhibited enhanced fluorescence in fluorescamine labeling upon interacting with PS NPs which was assumed to be because of protein unfolding.[36] After 1 hr incubation at 37 ^{0}C, 4× SYPRO Orange was added into the mixture of Tf and PS NPs. CFX Real-Time PCR (Bio-Rad) was used as both the temperature controller and the recorder for fluorescence intensity. As shown in Figure 4.2 a-b, the original T_m for Tf was at 67 ^{0}C, but it shifted to a lower temperature after binding to PS NPs. Interestingly, two new T_m values were observed: one was around 65 ^{0}C, and the other one was around 45 ^{0}C. Since Tf binding to the PS NPs reached an equilibrium in the solution, the NP-bound Tf and the free Tf could possess different T_m values. Judged from the results of fluorescamine labeling, Tf would unfold upon binding to the PS NPs. Thus,

it is highly possibly that the T_m at 65 ^{0}C is related to proteins interacting transiently with PS NPs or even free in the solution, and the T_m at 45 ^{0}C is related to Tf bound to the PS NPs and destabilized.[36] Besides Tf, we also tested apo-Tf, which is a Tf molecule without iron ions bound and thus less stable. The low stability was reflected in its SYPRO Orange labeling profile shown in Figure 4.2: a larger decrease in the T_m value induced by PS NPs was observed in apo-Tf compared to Tf. All these results verified that SYPRO Orange was practicable to measure the thermal stability shift of proteins induced by ENMs.

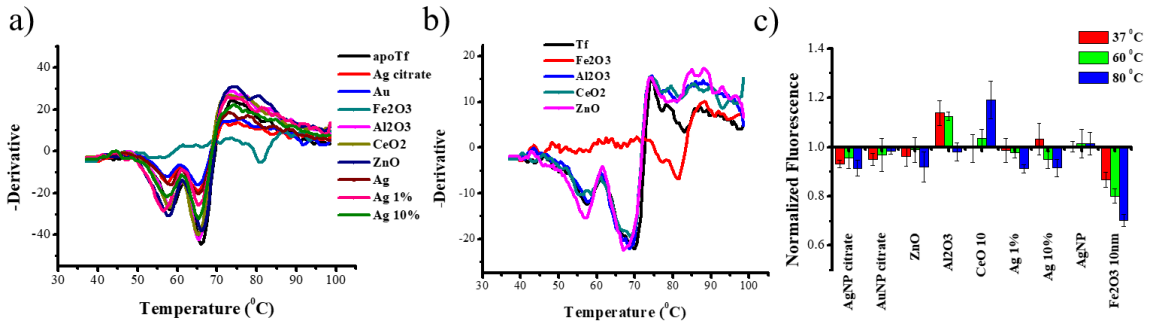


Figure 4.3. Fluorescence change of SYPRO Orange for (a) apo-Tf and (b) Tf with different ENMs. (c) Normalized fluorescence profiles of apo-Tf with various kinds of ENMs labeled by fluorescamine at different temperatures. The fluorescence intensities were normalized to the signal of the protein control.

We continued to apply this method to study the protein stability changes induced by other ENMs. Interestingly, we observed that, while some ENMs reduced the thermal stability of the proteins, the other could enhance the stability as well. Figure 4.3 a-b shows the result of Tf/apo-Tf with different ENMs. It is clearly seen that the Fe_2O_3 NPs could increase the T_m from the range of 55-70 ^{0}C observed with the majority of the ENMs screened to a higher value of 83 ^{0}C (Figure 4.3 a-b). This result indicates that the binding between Tf/apo-Tf and Fe_2O_3 NPs could stabilize Tf/apo-Tf, and less unfolding of the

protein could happen at temperatures below 83 ^{0}C. Since protein unfolding could enhance fluorescamine labeling by exposing more surface amines, and binding could diminish the fluorescence by blocking the surface sites, the degree of fluorescamine labeling for Tf or apo-Tf with Fe_2O_3 NPs should decrease at 60 or 80 ^{0}C, both still lower than the T_m measured by the SYPRO Orange method, which was well verified in Figure 4.3c: obvious decrease was observed in the fluorescence resulted from fluorescamine labeling of apo-Tf when incubated with the Fe_2O_3 NPs; and more decrease occurred at higher temperatures. When this ENM was incubated with other proteins, small changes in the fluorescence were observed, excluding the possibility that the decrease was caused by the quenching effect of Fe_2O_3 NPs on the dye. Therefore, both the SYPRO Orange and the fluorescamine labeling methods support that, Tf/apo-Tf could be stabilized by the Fe_2O_3 NPs.

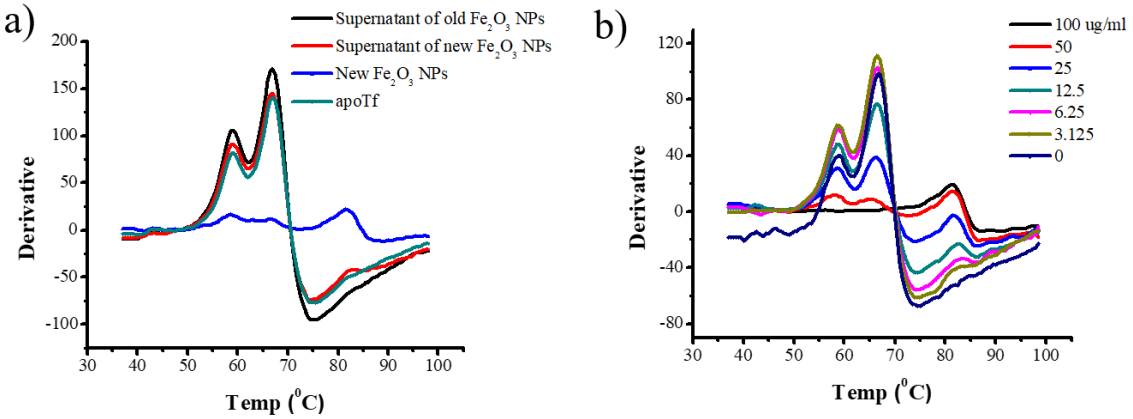
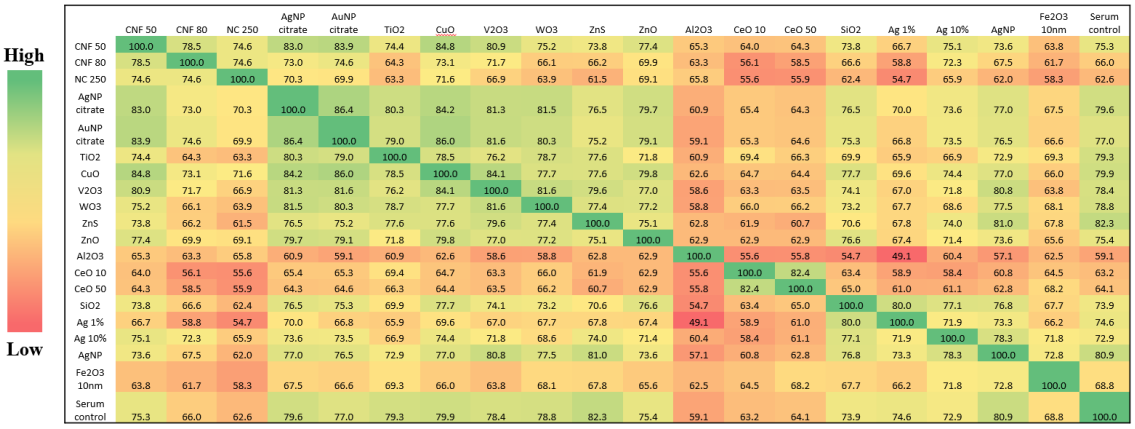


Figure 4.4. (a) Derivative plot of SYPRO Orange fluorescence signal against temperature, for apo-Tf itself (olive green line), apo-Tf incubated with either iron oxide NPs (blue line) and supernatant of old or new iron oxide NPs (black or red line). (b) Derivative plot of SYPRO Orange fluorescence signal against temperature, for 200 μg/ml apo-Tf incubated with different amounts of iron oxide NPs (0-100 μg/ml).

To confirm that the stabilization effect was not resulted from the free Fe (III) ions leaked by the NPs,[48-50] we removed the Fe_2O_3 NPs from the stock solution by centrifugation, and then added apo-Tf to the supernatant for thermal stability testing (Figure 4.4a). The stock solution was prepared and stored in a fridge for either one day (new) or one month (old). There was no obvious influence of the supernatant to the thermal stability of apo-Tf, which indicated the amount of Fe (III) dissolved in solution was not enough to induce the T_m shift. The stabilization effect could be the result of the strong binding of Tf or apo-Tf to the Fe-rich ENM. This is further confirmed by the experiment that changed the concentrations of the Fe_2O_3 NPs in the incubation with apo-Tf. Indeed, higher NP concentrations caused more T_m shift with the concentration of NPs (Figure 4.4 b). The new T_m peak at > 80 ^{0}C appeared with a very small amount (3.125 µg/ml) of NPs, and it became higher with increasing concentration of NPs.



	CNF 50	CNF 80	NC 250	AgNP citrate	AuNP citrate	TiO2	CuO	V2O3	WO3	ZnS	ZnO	Al2O3	CeO 10	CeO 50	SiO2	Ag 1%	Ag 10%	AgNP	Fe2O3 10nm	Serum control
CNF 50	100.0	78.5	74.6	83.0	83.9	74.4	84.8	80.9	75.2	73.8	77.4	65.3	64.0	64.3	73.8	66.7	75.1	73.6	63.8	75.3
CNF 80	78.5	100.0	74.6	73.0	74.6	64.3	73.1	71.7	66.1	66.2	69.9	63.3	56.1	58.5	66.6	58.8	72.3	67.5	61.7	66.0
NC 250	74.6	74.6	100.0	70.3	69.9	63.3	71.6	66.9	63.9	61.5	69.1	65.8	55.6	55.9	62.4	54.7	65.9	62.0	58.3	62.6
AgNP citrate	83.0	73.0	70.3	100.0	86.4	80.3	84.2	81.3	81.5	76.5	79.7	60.9	65.4	64.3	76.5	70.0	73.6	77.0	67.5	79.6
AuNP citrate	83.9	74.6	69.9	86.4	100.0	79.0	86.0	81.6	80.3	75.2	79.1	59.1	65.3	64.6	75.3	66.8	73.5	76.5	66.6	77.0
TiO2	74.4	64.3	63.3	80.3	79.0	100.0	78.5	76.2	78.7	77.6	71.8	60.9	69.4	66.3	69.9	65.9	66.9	72.9	69.3	79.3
CuO	84.8	73.1	71.6	84.2	86.0	78.5	100.0	84.1	77.7	77.6	79.8	62.6	64.7	64.4	77.7	69.6	74.4	77.0	66.0	79.9
V2O3	80.9	71.7	66.9	81.3	81.6	76.2	84.1	100.0	81.6	79.6	77.0	58.6	63.3	63.5	74.1	67.0	71.8	80.8	63.8	78.4
WO3	75.2	66.1	63.9	81.5	80.3	78.7	77.7	81.6	100.0	77.4	77.2	58.8	66.0	66.2	73.2	67.7	68.6	77.5	68.1	78.8
ZnS	73.8	66.2	61.5	76.5	75.2	77.6	77.6	79.6	77.4	100.0	75.1	62.8	61.9	60.7	70.6	67.8	74.0	81.0	67.8	82.3
ZnO	77.4	69.9	69.1	79.7	79.1	71.8	79.8	77.0	77.2	75.1	100.0	62.9	62.9	62.9	76.6	67.4	71.4	73.6	65.6	75.4
Al2O3	65.3	63.3	65.8	60.9	59.1	60.9	62.6	58.6	58.8	62.8	62.9	100.0	55.6	55.8	54.7	49.1	60.4	57.1	62.5	59.1
CeO 10	64.0	56.1	55.6	65.4	65.3	69.4	64.7	63.3	66.0	61.9	62.9	55.6	100.0	82.4	63.4	58.9	58.4	60.8	64.5	63.2
CeO 50	64.3	58.5	55.9	64.3	64.6	66.3	64.4	63.5	66.2	60.7	62.9	55.8	82.4	100.0	65.0	61.0	61.1	62.8	68.2	64.1
SiO2	73.8	66.6	62.4	76.5	75.3	69.9	77.7	74.1	73.2	70.6	76.6	54.7	63.4	65.0	100.0	80.0	77.1	76.8	67.7	73.9
Ag 1%	66.7	58.8	54.7	70.0	66.8	65.9	69.6	67.0	67.7	67.8	67.4	49.1	58.9	61.0	80.0	100.0	71.9	73.3	66.2	74.6
Ag 10%	75.1	72.3	65.9	73.6	73.5	66.9	74.4	71.8	68.6	74.0	71.4	60.4	58.4	61.1	77.1	71.9	100.0	78.3	71.8	72.9
AgNP	73.6	67.5	62.0	77.0	76.5	72.9	77.0	80.8	77.5	81.0	73.6	57.1	60.8	62.8	76.8	73.3	78.3	100.0	72.8	80.9
Fe2O3 10nm	63.8	61.7	58.3	67.5	66.6	69.3	66.0	63.8	68.1	67.8	65.6	62.5	64.5	68.2	67.7	66.2	71.8	72.8	100.0	68.8
Serum control	75.3	66.0	62.6	79.6	77.0	79.3	79.9	78.4	78.8	82.3	75.4	59.1	63.2	64.1	73.9	74.6	72.9	80.9	68.8	100.0

Figure 4.5. The similarity of serum protein corona composition for different ENMs, based on Equation 2. Serum itself was used as a control. The cells were colored from green to red, based on their values (100-49.1).

Correlation of fluorescence profiles with protein corona compositions. Agreement between SYPRO Orange and fluorescamine labeling in screening protein

structure change upon binding to ENMs supports the validity of fluorescamine labeling in probing protein-ENM interaction. Compared to SYPRO Orange labeling, fluorescamine labeling is more straightforward to be carried out, with only fluorescence intensity measured, while SYPRO Orange labeling requires substantial data processing to get the T_m value. Since fluorescamine labeling is a rapid method to screen the interaction between standard proteins and different ENMs, it would be the most valuable if the labeling profiles obtained between different protein-ENM pairs could be used to predict protein corona formation. To explore this possibility, we studied the protein corona formed on several ENMs when incubated in human serum, and evaluated the correlation between corona composition and the ENM's fluorescamine labeling profiles with various standard proteins.

Agreement between SYPRO Orange and fluorescamine labeling in screening protein structure change upon binding to ENMs supports the validity of fluorescamine labeling in probing protein-ENM interaction. Compared to SYPRO Orange labeling, fluorescamine labeling is more straightforward to be carried out, with only fluorescence intensity measured, while SYPRO Orange labeling requires substantial data processing to get the T_m value. Since fluorescamine labeling is a rapid method to screen the interaction between standard proteins and different ENMs, it would be the most valuable if the labeling profiles obtained between different protein-ENM pairs could be used to predict protein corona formation. To explore this possibility, we studied the protein corona formed on several ENMs when incubated in human serum, and evaluated the correlation between corona composition and the ENM's fluorescamine labeling profiles with various standard proteins.

To get the data of protein corona composition, we employed the standard centrifugation approach to precipitate the corona proteins with the ENMs, and then identified the proteins by LC-MS/MS. ENMs were incubated in human serum for 1 hr at 37 ^{0}C to allow protein corona to form and reach equilibrium. After centrifuging down ENMs and washing away the free proteins, a full trypsin digestion was performed on the proteins that were adsorbed strongly on ENMs, and the ENMs were removed by centrifugation, leaving the resultant peptides in the supernatant for LC-MS/MS analysis. The peptides were identified by running MS-GF+ against the human proteome database downloaded from UniProt. For each of the identified protein, the relative abundance (RA) was calculated based on Equation 1.

We evaluated the difference in protein corona between different ENMs. The similarity of the protein corona between every two ENMs was calculated based on Equation 2. All the similarity data were shown in Figure 4.5 and highlighted from red to green based on their values, with red being the lowest value and green being the highest one. ENMs screened here were very different in their core composition, size, and charge, thus high heterogeneity for their protein corona compositions was expected. Still, there were several pairs of ENMs that could allow assessment of how the property of ENM could affect corona formation. For example, among the ENMs tested, there were the cesium oxide NPs of two different diameters. They formed distinct protein corona from other ENMs, but shared very similar protein compositions with each other, despite their size differences. Another group of ENMs were the Silica NPs and the silver-dosed silica NPs. They also shared high similarity in their protein corona composition, due to the common

base structure of silica. Besides the core material, surface chemistry also strongly influenced corona formation. For instance, the citrate-coated Au or Ag NPs shared a high similarity in their protein corona. We also compared the protein composition in the corona of different ENMs with that in the serum matrix, and the similarity values were found between 60-80%, which indicated that the corona composition did not resemble what were in the matrix. Instead, distinct adsorption patterns of different ENMs yielded heterogenous corona compositions, which is consistent with previous research.[22]

High

Low

	serum albumin	transferrin	Ig G1	Ig M	*Ig*	Apo-AI	Apo-AII	Apo-E	*Clusterin*	*Hemopexin*	*Histidine-rich glycoprotein*	*Alpha-1-antitrypsin*	*Alpha-2-macroglobulin*
CEM	0.19	0.11	0.11	0.38	0.15	0.19	0.46	0.41	0.31	0.23	0.44	0.10	0.27
CEM_FBS	0.38	0.03	0.24	0.16	0.32	0.19	0.30	0.07	0.16	0.12	0.35	0.01	0.06
TG-37	0.03	0.26	0.40	0.21	0.39	0.49	0.28	0.41	0.12	0.30	0.11	0.15	0.40
TG-60	0.20	0.21	0.41	0.41	0.33	0.44	0.36	0.19	0.07	0.03	0.16	0.31	0.04
TG-80	0.00	0.02	0.03	0.21	0.16	0.19	0.02	0.21	0.04	0.11	0.20	0.14	0.09
aCd-37	0.28	0.48	0.12	0.14	0.20	0.13	0.17	0.05	0.05	0.46	0.33	0.32	0.19
aCd-60	0.49	**0.54**	0.09	0.12	0.17	0.10	0.05	0.07	0.00	0.39	0.02	0.22	0.05
aCd-80	0.47	0.09	0.18	0.14	0.07	0.19	0.13	0.18	0.01	0.07	0.11	0.34	0.35
Ova-37	0.09	0.00	0.44	0.18	0.37	0.28	0.14	0.34	0.38	0.09	0.42	0.00	0.39
Ova-60	0.02	0.18	0.02	0.05	0.01	0.14	0.20	0.01	0.01	0.45	0.04	0.40	0.06
Ova-80	0.01	0.09	0.09	0.25	0.12	0.15	0.43	0.17	0.19	0.02	0.05	0.45	0.27
aC-37	0.13	0.02	0.04	0.24	0.19	0.10	0.38	0.08	0.03	0.03	0.02	0.04	0.08
aC-60	0.37	0.06	0.19	0.16	0.21	0.31	0.11	0.03	0.09	0.08	0.22	0.02	0.00
aC-80	0.23	0.00	0.02	0.26	0.16	0.07	0.38	0.15	0.18	0.02	0.04	0.01	0.04
FG-37	0.16	0.04	0.16	0.15	0.22	0.20	0.29	0.15	0.20	0.13	0.17	0.25	0.04
FG-60	0.36	0.07	0.03	0.34	0.14	0.31	0.39	0.21	0.33	0.17	0.26	0.18	0.12
Fg-80	0.39	0.13	0.07	0.40	0.38	0.17	**0.53**	0.18	0.46	0.05	0.37	0.01	0.05
His-37	0.09	0.09	**0.61**	0.45	**0.58**	0.36	0.37	0.34	0.21	0.21	0.19	0.18	0.19
His-60	0.04	0.04	0.39	0.20	0.28	0.11	0.23	0.12	0.09	0.12	0.20	0.28	0.07
His-80	0.34	0.30	0.38	**0.64**	**0.64**	0.42	0.28	0.31	0.17	0.18	0.14	0.15	0.38
LA-37	0.16	0.42	0.21	0.20	0.18	0.06	0.15	0.08	0.41	0.38	**0.56**	**0.51**	0.29
LA-60	0.25	0.14	0.31	0.14	0.18	0.38	0.01	0.09	0.21	0.02	0.31	0.16	0.08
La-80	0.19	0.06	0.38	0.26	0.15	**0.53**	0.33	0.02	0.15	0.01	0.11	0.47	0.14
Lys-37	**0.50**	0.17	0.27	0.07	0.02	0.09	0.02	0.40	**0.51**	0.25	0.19	0.15	0.31
Lys-60	0.20	0.16	0.13	0.10	0.11	0.09	0.01	0.08	0.08	0.17	0.30	0.25	0.07
Lys-80	0.09	0.17	0.05	0.08	0.08	0.19	0.01	0.28	0.04	0.33	0.28	0.16	0.02
rG-37	0.19	0.24	0.18	0.30	0.31	0.39	0.32	0.05	0.05	0.35	0.07	0.02	0.04
rG-60	0.30	0.41	0.17	0.45	**0.55**	0.43	**0.50**	0.13	0.21	**0.68**	**0.54**	0.47	**0.52**
rG-80	0.21	0.28	0.11	0.25	0.40	0.35	0.15	0.10	0.19	**0.68**	0.25	0.45	0.37
Tf-37	0.32	**0.52**	0.17	0.29	0.24	0.27	0.48	0.28	0.15	0.22	0.16	0.37	0.13
Tf-60	0.17	0.43	0.27	**0.59**	**0.57**	0.49	**0.61**	0.27	0.14	0.19	0.08	0.09	0.23
Tf-80	0.36	0.06	0.05	0.25	0.03	0.40	0.35	0.13	0.09	0.13	0.15	0.24	0.07
H-37	0.33	0.34	0.04	0.01	0.07	0.01	0.22	0.00	0.21	0.05	0.23	0.21	0.00
H-60	**0.55**	0.02	0.23	0.34	0.14	0.15	0.24	0.29	0.32	0.02	0.16	0.07	0.05
H-80	0.19	0.04	0.12	0.22	0.11	0.20	0.31	0.10	0.12	0.17	0.12	0.27	0.16

Figure 4.6. The absolute value of the correlation coefficient between the fluorescamine labeling profiles of one protein to ENMs and the %RA of 13 most abundant proteins identified in corona. For the first column, characters before the dash line represent protein name, and the number behind

it represent the temperature, at which fluorescamine labeling was performed. Each cell was colored from dark green to white, based on their values. The values above 0.5 were also highlighted in red.

Next, we tested the correlations between fluorescamine labeling and the composition of serum protein corona. For each protein identified in the protein corona, its %RA values for all ENMs were viewed as one group of data, and compared with the group composed of the normalized fluorescence of one standard protein incubated with all ENMs (obtained in previous section). The absolute value of the correlation coefficient (|r|) was calculated for each pair of the array that involved the same ENM. In Figure 4.6, each column represented one of the 13 proteins with the highest abundance in the protein corona formed in human serum, and each row represented the standard protein used for fluorescamine labeling and the temperature condition. Interestingly, high correlations (e.g. $|r| > 0.5$) have been observed for some pairs. For example, the composition of serum albumin in protein corona showed a high correlation ($|r| = 0.55$) with the fluorescamine labeling profile of HSA itself at 60 ^{0}C. A similar phenomenon was also observed for the composition of transferrin in the serum protein corona, showing the second highest correlation ($|r| = 0.52$) to the transferrin labeling profile at 37 ^{0}C. Immunoglobulin (Ig), including Ig G and Ig M, showed high correlations with the γ globulin labeling profile at 60 ^{0}C, despite the relatively high correlations with the labeling profiles of Tf and histone at the same time. These correlations indicated that fluorescamine labeling profiles at different temperatures, which have been proved to show the different binding behavior of proteins to ENMs, could also be useful to differentiate the composition of protein corona formed in a complex matrix, e.g. serum. However, poor correlations for other pairs, including these between HSA labeling profiles at 37 or 80 ^{0}C and the serum albumin

compositions in corona, indicated one single factor was not enough to fully reveal the complicated interactions in the matrix. More sophisticated measurements towards protein adsorption on individual ENMs were needed to improve the correlation.

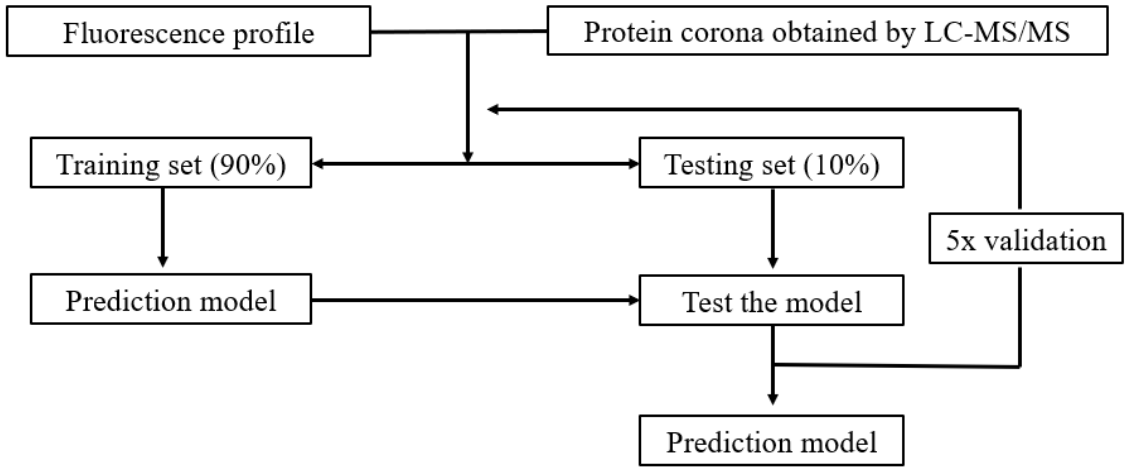


Figure 4.7. The workflow to build the prediction modeling. The fluorescence profiles of ENMs with different proteins at 37, 60, and 80 ^{0}C were used as descriptors for ENMs. The %RA of each protein identified in serum protein corona was used as the target.

Prediction models based on classification and regression. Since fluorescamine labeling profiles showed differentiations on ENMs and reasonable correlations with the protein composition in serum protein corona, we tested whether the fluorescamine labeling profile could be of any value in prediction of corona formation in biological samples. We initiated the prediction modeling effort by incorporating the fluorescence obtained with fluorescamine labeling as the descriptors for the ENMs. The properties of the proteins in the corona, including MW, pI, GRAVY, percentage of negative/positive/aromatic amino acids, and their abundances in the serum control, were used as the descriptors for proteins. For the classification model, the proteins enriched in the corona of ENMs was considered as the positive hit. For the regression model, the change of %RA of each protein in corona,

compared to the serum control, was used as the target value for prediction. In total, 19 ENMs and the 100 most abundant proteins in corona were chosen. All data was randomly split into two datasets: the training set containing 90% of the data, and the testing set with the remaining 10%. The prediction model was built based on the training set, and the algorithm was random forest (RF). One thousand bootstraps and a minimum value for each leaf node (>=3) were used to help minimizing overfitting. The performance of the obtained prediction model was verified on the testing set, which was viewed as the unknown dataset. After 5-fold cross validation, the final model was obtained.

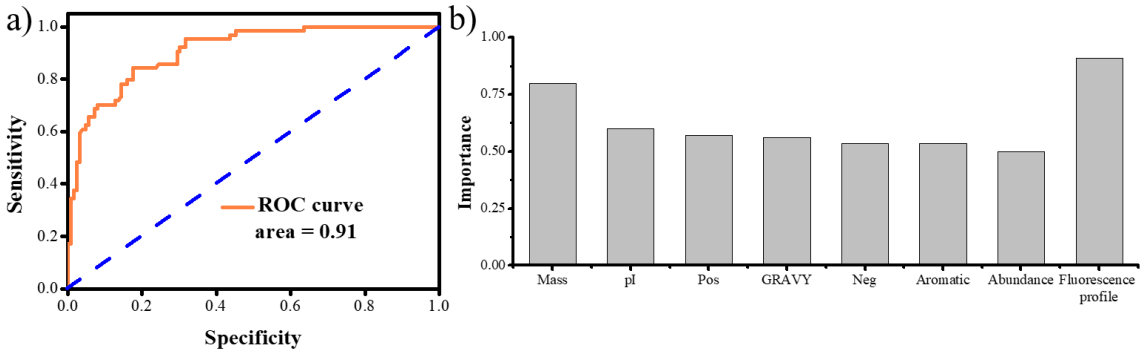


Figure 4.8. (a) Receptor operation characteristic (ROC) curve of the classification model. (b) The importance of each descriptor used in the classification model.

The ROC curve of the classification model was shown in Figure 4.8 a, and the prediction ability of this model was very good, as the curve was much deviated from the random guess line (blue dash line). The area under the curve (AUC) was 0.91, which indicated the excellent ability of our model to discriminate the corona protein. As for the standard performance metrics, the overall precision and recall by this model were both 0.85, and the f1 score was 0.84. The improvements in both accuracy and AUC were achieved in our model, comparing to the benchmark work reported previously.[17] It is worth

noting that the ENMs used here were made from rather different core materials, instead of some ENMs with the same core composition but different sizes or surface modifications. Previous reports have already shown the difficulty to incorporate ENMs with different core materials for prediction.[22] One reason was that the descriptors of ENMs, such as size and charge, could not differentiate ENMs with different compositions very well, and that was also why new descriptors were heavily needed. From Figure 4.8 b, it was interesting to see the importance of the total fluorescence profile was higher than the descriptors of protein, which indicated that incorporating fluorescence profiles as the descriptors of ENMs could improve the performance of this prediction model to a large extent.

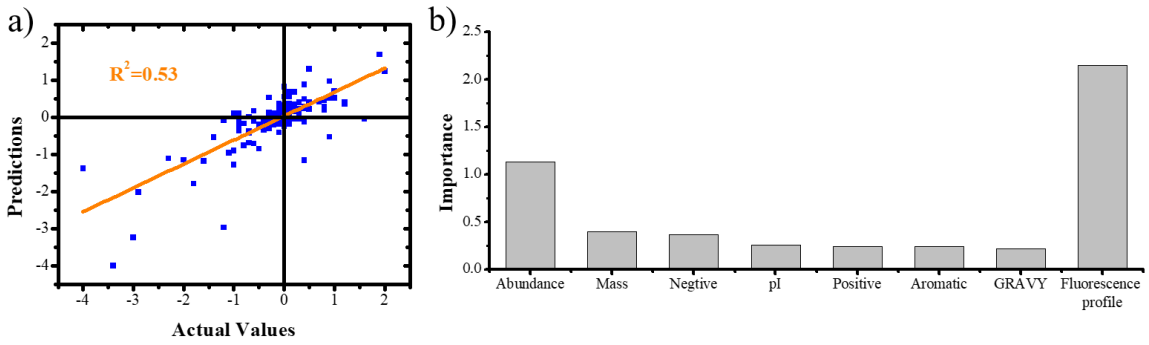


Figure 4.9. (a) The prediction performance of the regression model on the testing set. The linear regression line is in orange. (b) The importance of each descriptor used in the regression model.

To predict "yes" or "no", i.e. whether the protein could be present in the corona, is not enough. We further challenged our model with a more difficult task: to predict the value change for %RA of each protein in the protein corona, using the regression model. For each protein, the %RA in the serum protein corona of ENMs was compared to that in the serum control. The prediction results from the testing set were shown in Figure 4.9 a, and the mean absolute error was 0.11. Although the R^2 was only 0.53, most of the datapoints were in the 1st and 3rd quadrants, which indicated the correct prediction on the signs of

%RA changes. Most importantly, the importance of the total fluorescence profile was much higher than all other descriptors, including the abundance of proteins in serum (Figure 4.9 b). This indicated the contribution of the fluorescence profile as the descriptor for the ENM was the most critical for this regression model to be successful.

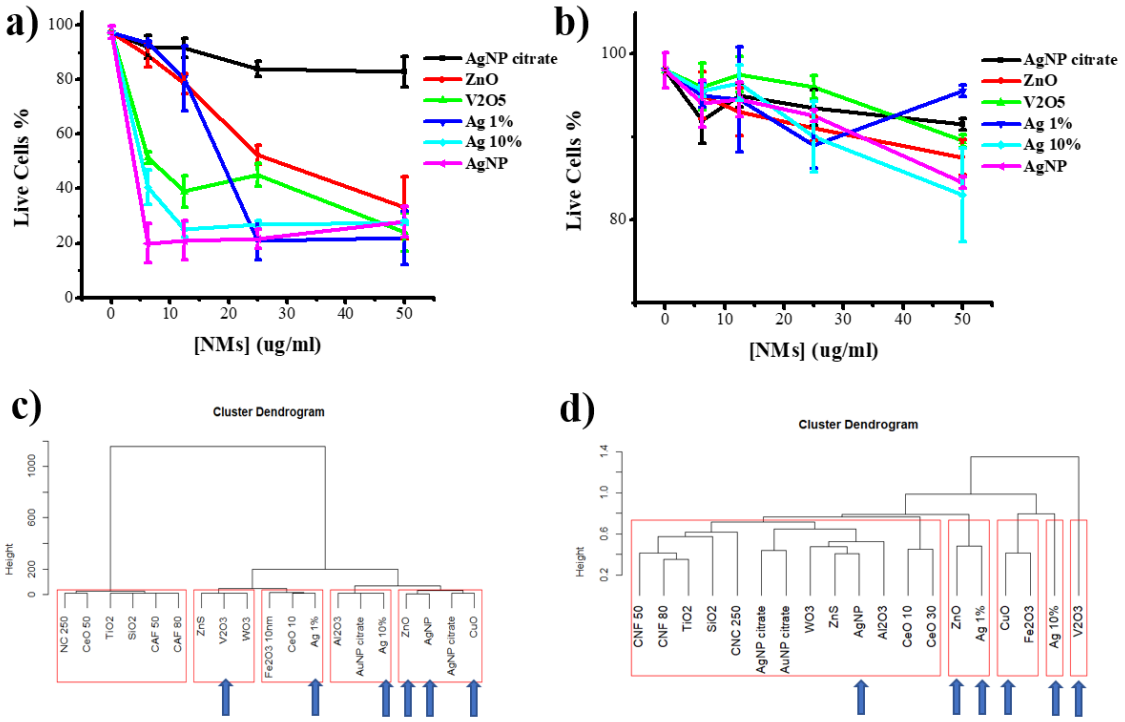


Figure 4.10. The percentage of live cells identified in the trypan blue test, after CEM cells were treated by different amount of ENMs in media without (a) or with (b) 10% FBS for 1 day. The hierarchical clustering result using (c) thermal stability screening data or (d) fluorescamine labeling profiles. Five clusters were circled by red boxes. ENMs identified as toxic in our screening were marked with arrows.

Correlation of the fluorescence screening data with ENMs' cytotoxicity. Owing to its important role in modulating biological outcomes of the ENMs, protein corona composition has been used for the SAR model to predict cellular interactions of ENMs, which showed big improvements in both accuracy and interpretability over the models only

using physicochemical properties of ENMs as the descriptors.[22] Since fluorescamine labeling profiles could differentiate protein-ENMs interactions and predict serum protein corona compositions, it is very promising to use them as the descriptors to predict biological outcomes of ENMs. Comparing with the time-consuming identification of protein corona compositions by LC-MS/MS, fluorescamine labeling is a high throughput screening method and could greatly improve the efficiency of the biological evaluation of ENMs. To verify this assumption, we first measured cytotoxicity of these ENMs on human T lymphoblast (CEM) cells, using trypan blue staining. At 50 μg/ml, most of these ENMs didn't show cytotoxicity in the culture media without fetal bovine serum (FBS); furthermore, only some of them, including the silica/silver NPs, zinc oxide NPs, and vanadium oxide NPs, showed high cytotoxicity at even lower doses (figure 4.10 a). Moreover, the cytotoxicity of these ENMs could be alleviated by adding FBS in the media (figure 4.7 b), which was consistent with the literature that reported protein corona formed in FBS could decrease the cytotoxicity of ENMs.

To show the correlation between cytotoxicity of ENMs and our fluorescence screening data, unsupervised hierarchical clustering was used to group ENMs based on the fluorescamine labeling obtained in the previous sections. Since the thermal stability screening data shown correlations with the fluorescamine labeling profiles, they were also used as variables for the clustering of ENMs. The bottom-up manner based agglomerative clustering, e.g. complete linkage, was used to find out similar clusters.[51] Both clusters showed rough discriminations on the cytotoxicity of ENMs (figure 4.10 c-d). All cellulose based NPs, silica NPs, and titanium oxide NPs were grouped into one cluster, and these

ENMs are usually considered as less toxic. However, ENMs with high cytotoxicity, as identified in our test, were not grouped together. The limitation of using trypan blue staining to quantify cytotoxicity may contribute to the failure of grouping the highly toxic ENMs: trypan blue staining only labels the cells with membrane damages, such as those undergoing necrosis. However, cytotoxicity induced by ENMs could also be a result from apoptosis that may not induce membrane damages or from other cellular effects, e.g. oxidative stress and inflammasome activation. Cell death caused by these factors could not be stained by trypan blue.[52,53] Despite these limitations, the preliminary clustering result indicated fluorescamine labeling and thermal stability screening profiles could have potential for the prediction of biological outcomes of ENMs.

4.4: Conclusions

Fluorescamine labeling at different temperatures has been shown to be able to successfully discriminate ENM-protein interactions. The stress of higher temperatures would induce protein unfolding, which exposes more labeling areas for fluorescamine and enlarge the fluorescence signals. Besides, binding to ENMs could alter the thermal stability of the protein, which was verified by SYPRO Orange labeling. Both stabilization and destabilization effects were observed; and the alteration in thermal stability introduces differences in fluorescamine labeling at varied temperatures. Moreover, good correlations were observed between fluorescamine labeling profiles and protein compositions in serum protein corona of ENMs. The result supports the feasibility of using these fluorescence profiles as the novel descriptors of ENMs to predict protein corona compositions. By incorporating fluorescamine labeling profiles as the descriptors, the prediction model based

on classification showed improvement in performance, comparing to the benchmark work. The prediction model based on regression was also proved to be successful and robust, which could provide quick and routine evaluation of the protein corona formation for ENMs development. The easily accessible fluorescamine labeling data can dramatically attenuate the workload of traditional protein corona study, and it will provide new insights for protein-ENM interactions. Finally, these fluorescence profiles, as rapid, robust, and bio-related screening data, also showed great potentials to predict biological outcomes of ENMs.

References

(1) Wagner, V.; Dullaart, A.; Bock, A.-K.; Zweck, A. The emerging nanomedicine landscape. *Nature biotechnology* **2006**, *24*, 1211.

(2) Shi, J.; Kantoff, P. W.; Wooster, R.; Farokhzad, O. C. Cancer nanomedicine: progress, challenges and opportunities. *Nature Reviews Cancer* **2017**, *17*, 20.

(3) von Roemeling, C.; Jiang, W.; Chan, C. K.; Weissman, I. L.; Kim, B. Y. Breaking down the barriers to precision cancer nanomedicine. *Trends in biotechnology* **2017**, *35*, 159-171.

(4) Riehemann, K.; Schneider, S. W.; Luger, T. A.; Godin, B.; Ferrari, M.; Fuchs, H. Nanomedicine—challenge and perspectives. *Angewandte Chemie International Edition* **2009**, *48*, 872-897.

(5) Lynch, I.; Weiss, C.; Valsami-Jones, E. A strategy for grouping of nanomaterials based on key physico-chemical descriptors as a basis for safer-by-design NMs. *Nano Today* **2014**, *9*, 266-270.

(6) Walczyk, D.; Bombelli, F. B.; Monopoli, M. P.; Lynch, I.; Dawson, K. A. What the cell "sees" in bionanoscience. *J Am Chem Soc* **2010**, *132*, 5761-5768.

(7) Van Hong Nguyen, B.-J. L. Protein corona: a new approach for nanomedicine design. *International journal of nanomedicine* **2017**, *12*, 3137.

(8) Corbo, C.; Molinaro, R.; Tabatabaei, M.; Farokhzad, O. C.; Mahmoudi, M. Personalized protein corona on nanoparticles and its clinical implications. *Biomaterials science* **2017**, *5*, 378-387.

(9) Chen, F.; Wang, G.; Griffin, J. I.; Brenneman, B.; Banda, N. K.; Holers, V. M.; Backos, D. S. Complement proteins bind to nanoparticle protein corona and undergo dynamic exchange in vivo. **2017**, *12*, 387-393.

(10) Monopoli, M. P.; Walczyk, D.; Campbell, A.; Elia, G.; Lynch, I.; Baldelli Bombelli, F.; Dawson, K. A. 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*, 2525-2534.

(11) Lesniak, A.; Fenaroli, F.; Monopoli, M. P.; Åberg, C.; Dawson, K. A.; Salvati, A. Effects of the presence or absence of a protein corona on silica nanoparticle uptake and impact on cells. *ACS nano* **2012**, *6*, 5845-5857.

(12) Yan, Y.; Gause, K. T.; Kamphuis, M. M.; Ang, C.-S.; O'Brien-Simpson, N. M.; Lenzo, J. C.; Reynolds, E. C.; Nice, E. C.; Caruso, F. Differential roles of the protein corona in the cellular uptake of nanoporous polymer particles by monocyte and macrophage cell lines. *ACS nano* **2013**, *7*, 10960-10970.

(13) Saha, K.; Moyano, D. F.; Rotello, V. M. Protein coronas suppress the hemolytic activity of hydrophilic and hydrophobic nanoparticles. *Materials horizons* **2014**, *1*, 102-105.

(14) Lee, Y. K.; Choi, E.-J.; Webster, T. J.; Kim, S.-H.; Khang, D. Effect of the protein corona on nanoparticles for modulating cytotoxicity and immunotoxicity. *International journal of nanomedicine* **2015**, *10*, 97.

(15) Li, L.; Mu, Q.; Zhang, B.; Yan, B. Analytical strategies for detecting nanoparticle-protein interactions. *Analyst (Cambridge, U. K.)* **2010**, *135*, 1519-1530.

(16) Smits, A. H.; Vermeulen, M. Characterizing protein–protein interactions using mass spectrometry: challenges and opportunities. *Trends in biotechnology* **2016**, *34*, 825-834.

(17) Findlay, M. R.; Freitas, D. N.; Mobed-Miremadi, M.; Wheeler, K. E. Machine learning provides predictive analysis into silver nanoparticle protein corona formation from physicochemical properties. *Environmental Science: Nano* **2018**, *5*, 64-71.

(18) Puzyn, T.; Rasulev, B.; Gajewicz, A.; Hu, X.; Dasari, T. P.; Michalkova, A.; Hwang, H.-M.; Toropov, A.; Leszczynska, D.; Leszczynski, J. Using nano-QSAR

to predict the cytotoxicity of metal oxide nanoparticles. *Nature nanotechnology* **2011**, *6*, 175.

(19) Singh, K. P.; Gupta, S. Nano-QSAR modeling for predicting biological activity of diverse nanomaterials. *RSC Advances* **2014**, *4*, 13215-13230.

(20) Pan, Y.; Li, T.; Cheng, J.; Telesca, D.; Zink, J. I.; Jiang, J. Nano-QSAR modeling for predicting the cytotoxicity of metal oxide nanoparticles using novel descriptors. *RSC Advances* **2016**, *6*, 25766-25775.

(21) Papa, E.; Doucet, J.; Sangion, A.; Doucet-Panaye, A. Investigation of the influence of protein corona composition on gold nanoparticle bioactivity using machine learning approaches. *SAR and QSAR in Environmental Research* **2016**, *27*, 521-538.

(22) Walkey, C. D.; Olsen, J. B.; Song, F.; Liu, R.; Guo, H.; Olsen, D. W. H.; Cohen, Y.; Emili, A.; Chan, W. C. Protein corona fingerprinting predicts the cellular interaction of gold and silver nanoparticles. *ACS nano* **2014**, *8*, 2439-2455.

(23) Ashby, J.; Pan, S.; Zhong, W. Size and Surface Functionalization of Iron Oxide Nanoparticles Influence the Composition and Dynamic Nature of Their Protein Corona. *ACS Applied Materials and Interfaces* **2014**, *6*, 15412–15419.

(24) Durán, N.; Silveira, C. P.; Durán, M.; Martinez, D. S. T. Silver nanoparticle protein corona and toxicity: a mini-review. *Journal of nanobiotechnology* **2015**, *13*, 55.

(25) Clemments, A. M.; Botella, P.; Landry, C. C. Protein adsorption from biofluids on silica nanoparticles: corona analysis as a function of particle diameter and porosity. *ACS applied materials & interfaces* **2015**, *7*, 21682-21689.

(26) Lundqvist, M.; Stigler, J.; Cedervall, T.; Berggård, T.; Flanagan, M. B.; Lynch, I.; Elia, G.; Dawson, K. The evolution of the protein corona around nanoparticles: a test study. *ACS nano* **2011**, *5*, 7503-7509.

(27) Natte, K.; Friedrich, J. F.; Wohlrab, S.; Lutzki, J.; von Klitzing, R.; Osterle, W.; Orts-Gil, G. Impact of polymer shell on the formation and time evolution of nanoparticle-protein corona. *Colloids and Surfaces B-Biointerfaces* **2013**, *104*, 213-220.

(28) Hadjidemetriou, M.; Kostarelos, K. Nanomedicine: evolution of the nanoparticle corona. *Nature nanotechnology* **2017**, *12*, 288.

(29) Shang, W.; Nuffer, J. H.; Dordick, J. S.; Siegel, R. W. Unfolding of ribonuclease A on silica nanoparticle surfaces. *Nano letters* **2007**, *7*, 1991-1995.

(30) Deng, Z. J.; Liang, M.; Monteiro, M.; Toth, I.; Minchin, R. F. Nanoparticle-induced unfolding of fibrinogen promotes Mac-1 receptor activation and inflammation. *Nature nanotechnology* **2011**, *6*, 39.

(31) Dominguez-Medina, S.; Kisley, L.; Tauzin, L. J.; Hoggard, A.; Shuang, B.; Indrasekara, A. S.; Chen, S.; Wang, L. Y.; Derry, P. J.; Liopo, A.; Zubarev, E. R.; Landes, C. F.; Link, S. Adsorption and Unfolding of a Single Protein Triggers Nanoparticle Aggregation. *ACS Nano* **2016**, *10*, 2103-2112.

(32) Apul, O. G.; Wang, Q.; Shao, T.; Rieck, J. R.; Karanfil, T. Predictive model development for adsorption of aromatic contaminants by multi-walled carbon nanotubes. *Environmental science & technology* **2012**, *47*, 2295-2303.

(33) Mahmoudi, M. Debugging Nano–Bio Interfaces: Systematic Strategies to Accelerate Clinical Translation of Nanotechnologies. *Trends in biotechnology* **2018**.

(34) Xia, X.-R.; Monteiro-Riviere, N. A.; Riviere, J. E. An index for characterization of nanomaterials in biological systems. *Nature nanotechnology* **2010**, *5*, 671.

(35) Chen, R.; Zhang, Y.; Darabi Sahneh, F.; Scoglio, C. M.; Wohlleben, W.; Haase, A.; Monteiro-Riviere, N. A.; Riviere, J. E. Nanoparticle surface characterization and clustering through concentration-dependent surface adsorption modeling. *ACS nano* **2014**, *8*, 9446-9456.

(36) Duan, Y.; Liu, Y.; Shen, W.; Zhong, W. Fluorescamine Labeling for Assessment of Protein Conformational Change and Binding Affinity in Protein–Nanoparticle Interaction. *Analytical chemistry* **2017**, *89*, 12160-12167.

(37) Ashby, J.; Duan, Y.; Ligans, E.; Tamsi, M.; Zhong, W. High-Throughput Profiling of Nanoparticle–Protein Interactions by Fluorescamine Labeling. *Analytical chemistry* **2015**, *87*, 2213-2219.

(38) Tenzer, S.; Docter, D.; Kuharev, J.; Musyanovych, A.; Fetz, V.; Hecht, R.; Schlenk, F.; Fischer, D.; Kiouptsi, K.; Reinhardt, C. Rapid formation of plasma protein corona critically affects nanoparticle pathophysiology. *Nature nanotechnology* **2013**, *8*, 772.

(39) Lomenick, B.; Hao, R.; Jonai, N.; Chin, R. M.; Aghajan, M.; Warburton, S.; Wang, J.; Wu, R. P.; Gomez, F.; Loo, J. A. Target identification using drug affinity responsive target stability (DARTS). *Proceedings of the National Academy of Sciences* **2009**, pnas. 0910040106.

(40) Schön, A.; Clarkson, B. R.; Jaime, M.; Freire, E. Temperature stability of proteins: Analysis of irreversible denaturation using isothermal calorimetry. *Proteins: Structure, Function, and Bioinformatics* **2017**, *85*, 2009-2016.

(41) Keskin, O.; Gursoy, A.; Ma, B.; Nussinov, R. Principles of protein− protein interactions: What are the preferred ways for proteins to interact? *Chemical reviews* **2008**, *108*, 1225-1244.

(42) van Dun, S.; Ottmann, C.; Milroy, L.-G.; Brunsveld, L. Supramolecular chemistry targeting proteins. *Journal of the American Chemical Society* **2017**, *139*, 13960-13968.

(43) Leuenberger, P.; Ganscha, S.; Kahraman, A.; Cappelletti, V.; Boersema, P. J.; von Mering, C.; Claassen, M.; Picotti, P. Cell-wide analysis of protein thermal unfolding reveals determinants of thermostability. *Science* **2017**, *355*, eaai7825.

(44) Mateus, A.; Määttä, T. A.; Savitski, M. M. Thermal proteome profiling: unbiased assessment of protein state through heat-induced stability changes. *Proteome science* **2016**, *15*, 13.

(45) Layton, C. J.; Hellinga, H. W. Quantitation of protein–protein interactions by thermal stability shift analysis. *Protein Science* **2011**, *20*, 1439-1450.

(46) Niesen, F. H.; Berglund, H.; Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature protocols* **2007**, *2*, 2212.

(47) Debnath, M.; Farace, J. M.; Johnson, K. D.; Nesterova, I. Quantitation without calibration: response profile as an indicator of target amount. *Analytical chemistry* **2018**.

(48) Misra, S. K.; Dybowska, A.; Berhanu, D.; Luoma, S. N.; Valsami-Jones, E. The complexity of nanoparticle dissolution and its importance in nanotoxicological studies. *Science of the total environment* **2012**, *438*, 225-232.

(49) Keller, A. A.; Wang, H.; Zhou, D.; Lenihan, H. S.; Cherr, G.; Cardinale, B. J.; Miller, R.; Ji, Z. Stability and aggregation of metal oxide nanoparticles in natural aqueous matrices. *Environmental science & technology* **2010**, *44*, 1962-1967.

(50) Nel, A. E.; Mädler, L.; Velegol, D.; Xia, T.; Hoek, E. M.; Somasundaran, P.; Klaessig, F.; Castranova, V.; Thompson, M. Understanding biophysicochemical interactions at the nano–bio interface. *Nature materials* **2009**, *8*, 543.

(51) Suzuki, R.; Shimodaira, H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **2006**, *22*, 1540-1542.

(52) Mirshafiee, V.; Sun, B.; Chang, C. H.; Liao, Y.-P.; Jiang, W.; Jiang, J.; Liu, X.; Wang, X.; Xia, T.; Nel, A. E. Toxicological Profiling of Metal Oxide Nanoparticles in Liver Context Reveals Pyroptosis in Kupffer Cells and Macrophages versus Apoptosis in Hepatocytes. *ACS nano* **2018**, *12*, 3836-3852.

(53) Li, R.; Guiney, L. M.; Chang, C. H.; Mansukhani, N. D.; Ji, Z.; Wang, X.; Liao, Y.-P.; Jiang, W.; Sun, B.; Hersam, M. C. Surface oxidation of graphene oxide determines membrane damage, lipid peroxidation, and cytotoxicity in macrophages in a pulmonary toxicity model. *ACS nano* **2018**, *12*, 1390-1402.

CHAPTER 5: Limited Proteolysis for Exploration of the Binding Site of Synthetic Receptors on Protein

5.1: Introduction

Proteins are key regulators on signal pathways, cell functions and communications in the biological system, which make them the primary targets for diagnostics and therapeutics.[1,2] However, the complexity of protein and the competitive environment in cells could both make the specific recognition and protein targeting *in vivo* very challenging.[3] Antibodies have been used for protein recognition and are pivotal to the pharmaceutical industry today; but the high cost, prolonged production time, and issues to penetrate biological barriers, all limit their applications.[4] As for small drug molecules, the limited number of pockets in druggable proteins has become the bottleneck for the discovery of novel targeting compounds, despite successful endeavors using pocket docking and rational designs.[5] On the other hand, synthetic receptors with proper sizes, easy accessibility and versatile binding forces, have shown themselves to be promising tools for protein recognition and modulation, making them the forefront of success in life science and pharmaceutics.[6-8]

Synthetic receptors, including calixarenes and cucurbiturils, interact with proteins differently from small molecules, but more similarly to protein-protein interactions (PPIs).[8] With common features such as preorganized structures, hydrophobic cavities and polar rims, synthetic receptors will bind to protein surfaces with lower affinity but higher avidity, comparing to small molecules that usually dock into the pocket of protein selectively and specifically.[9,10] As the building block of protein, amino acids have been explored

thoroughly for the interactions with synthetic receptors in either aqueous or gas phase, via isothermal titration calorimetry (ITC), fluorescence, ultraviolet (UV), NMR spectroscopy, and X-ray diffraction analysis.[11-15] Although the binding affinities of synthetic receptors to amino acids and small peptides seem to be straightforward, neighboring moieties on the protein surface complicate synthetic receptors' behaviors when they interact with proteins.[8,16] Moreover, post-translational modifications (PTMs) make it even more challenging to predict synthetic receptor-protein interactions, and multiple PTMs could exist at various locations on the same protein, interfering with the recognition and targeting of synthetic receptors.[8,17,18] Furthermore, the complicated PPIs in biological systems can also make the identification of the real protein binding sites of synthetic receptors challenging.[8,19,20] Thus, robust and routine methods are desired to analyze the protein binding behaviors of synthetic receptors and harness their biomedical applications.

Similar to methods used for amino acids and peptides, ITC, capillary electrophoresis, absorbance and fluorescence have been used to quantify binding affinities or thermodynamics of synthetic receptors on different proteins.[12,21-23] These methods are easily accessible, and can be used in routine work. However, not enough molecular details, such as the binding sties, could be revealed in these methods. Currently, NMR spectroscopy and X-ray crystallography are two successful ways to obtain molecular details of proteins bound to synthetic receptors, and multiple crystal structures have been successfully obtained, including insulin, lysozyme (with or without methylations), cytochrome c (with or without methylations), luciferase, methylated lectin, and 14-3-3 proteins.[24-30] With these crystal structures, it is remarkable to see that synthetic receptors

do not bind to all potential binding sites identified in study of amino acid sequence. Such phenomena strongly support the assumption that influences from protein surface topology and PTMs could alter the binding behaviors of synthetic receptors. As a result, information at the molecular level about how synthetic receptors interact with proteins is critical to support further rational design of synthetic receptors with better protein targeting properties. Unfortunately, NMR or X-ray crystallography requires high protein concentrations, has specific requirements on the size or type of the protein, and is time consuming and labor intensive. These technical difficulties result in the very small numbers of structural data available today, in contrast to the huge numbers and types of synthetic receptors developed.[8,31]

Electro spray ionization (ESI) with low energy collision-induced dissociation (CID) and matrix assisted lase desorption ionization (MALDI) mass spectrometry (MS) are alternative methods which have been applied to study binding sites of cucurbiturils on peptides and ubiquitin based on the fragmentation patterns.[32,33] These methods also take advantage of the enhanced ionization from cucurbiturils in MS. However, its application is limited in this special scenario. MS coupled with limited proteolysis, on the other hand, has been proven to be a versatile and robust way of studying protein-drug interactions and PPIs. However, no application has been reported on structural exploration of the complex formed between synthetic receptors and proteins.[34,35] Since the hydrophobic or charged residues that are targeted by synthetic receptors are also the cutting sites of various kinds of proteases, binding of synthetic receptors on proteins may block these residues from being accessed by the proteases. Moreover, the short assay time in limited proteolysis could

minimize the impact of digested peptides on synthetic receptors, despite their quick exchange rates.[12] By choosing proper proteases, including trypsin (cutting positively charged residues), chymotrypsin (cutting aromatic residues), and proteinase K (nonspecific), specific and precise information could be obtained for residues blocked by synthetic receptors. Here, we show the proof-of-concept work on how limited proteolysis can provide information about the binding sites of cucurbit[7]uril (CB7) or 4-sulfonatocalix[4]arene (sclx4) on insulin or lysozyme at the molecular level.

5.2: Materials and Methods

General. Human insulin, lysozyme from chicken egg white, trypsin from porcine pancreas, α-chymotrypsin from bovine pancreas, α-Cyano-4-hydroxycinnamic acid (CHCA), Tris(2-carboxyethyl)phosphine hydrochloride (TCEP), formic acid (FA), cucurbit[7]uril hydrate and 4-sulfonatocalix[4]arene were purchased from Sigma-Aldrich (St. Louis, MO). Proteinase K was obtained from Promega (Madison, WI). ZipTips with C18 resin were purchased from Millipore Sigma (Burlington, MA). Ultrapure water (18 MΩ) was obtained from a Direct-Q Water Purification System (Millipore Sigma, Billerica, MA).

Proteolysis by trypsin, chymotrypsin, or proteinase K in solution. Proteins were incubated with CB7 or sclx4 in the molar ratios of 1:0-4, in 10 mM PBS buffer (pH 7.4) at 37 ^{0}C for 1 hr, and the protease of trypsin, chymotrypsin or proteinase K was added into the mixture at a 1: 30 mass ratio. $CaCl_2$ was also supplied at the concentration of 5 mM if chymotrypsin or proteinase K was used. The digestion lasted for 30-60 minutes at 37 ^{0}C, and terminated by the addition of 0.1% FA. The disulfide bonds were reduced by the

treatment of 5 mM TCEP for 30 minutes at 56 ^{0}C. ZipTips were used for peptide desalting. If digestion of the entire protein was needed, after TCEP treatment, trypsin was added at a mass ratio of 1:30 to digest the remaining protein for 6 hrs.

Limited proteolysis of lysozyme with sclx4 on surface. Fifty µl of 0.5 mg/ml lysozyme in 1× PBS pH 7.4 were added to each well of the Microfluor-2 96-well immune plate for 1 hr. After depleting the well, 200 µl 1× PBS were added for washing three times. Fifty µl of 0.1 mg/ml sclx4 in PBS was added to each well and incubated for 1 hr. 1× PBS was used to wash the well afterwards. Fifty µl of 0.05 mg/ml trypsin in 1× PBS was added into each well for 30 minutes at room temperature. After rinsing with 100 µl of 1× PBS twice, all solutions were collected and prepared for further analysis.

LC-MS/MS to identify cutting sites. Digested peptides were lyophilized, cleaned, and reconstituted in 0.1% FA before being injected into Waters CapLC system connected to Finnigan LTQ MS with a nano-ESI ion source. A 10-cm homemade column with 10 µm tip packed with 5 µm diameter C18 beads was used for separation. Flow rate at tip was set to between 0.2-0.3 µl/min. 0.1% FA in water and 0.1% FA in ACN were used for solvent A and B, respectively. The separation gradient was from 2% B to 40% B for 1 hr duration. Collision induced dissociation (CID) was used for fragmentation, and the mass range was set to 300-2000 Da.

MALDI-TOF-MS to identify peptides. The lyophilized and cleaned samples were re-dissolved in 10 µl of 0.1% FA and mixed with 10 µl CHCA saturated solution in 0.1% FA / 50% ACN. One ul of the mixture was spotted onto the Opti-TOF 96 plate and allowed

to completely dry. AB Sciex 5800 TOF/TOF MS analyzer was used to identify peptides. Positive reflection mode was used.

5.3: Results and Discussion

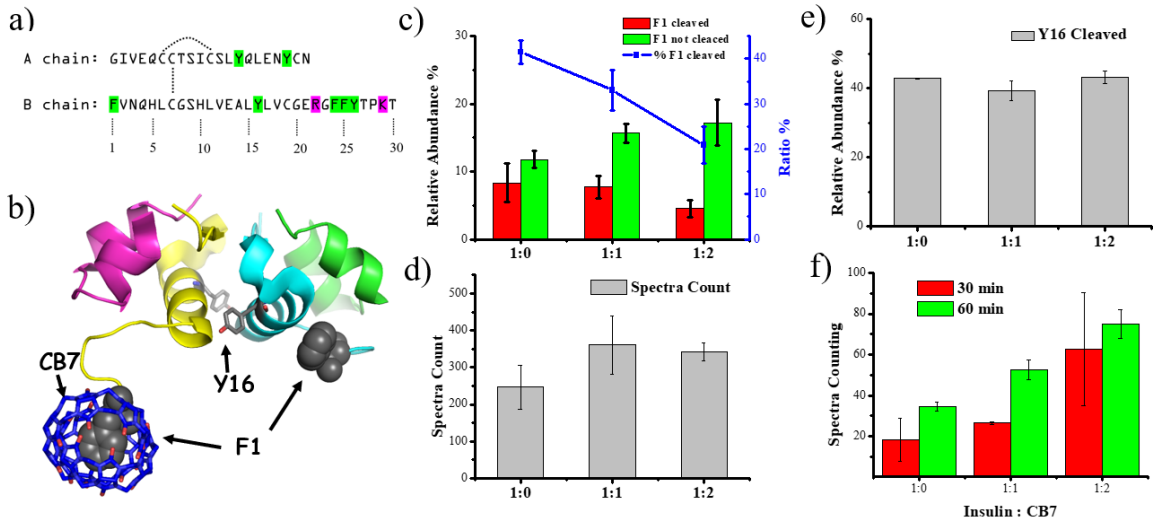


Figure 5.1. (a) The sequence of human insulin. Cutting sites for trypsin and chymotrypsin are highlighted in magenta or green, respectively. (b)The crystal structure of insulin bound with CB7 (PDB ID: 3Q6E). CB7 is shown in blue stick, and insulin is shown in cartoon, colored for different chains. F1 is shown in gray ball. (c) The % relative abundance for peptides with or without F1 cleaved by chymotrypsin are shown in bar plots. The ratio of F1 cleaved is shown in blue line. (d) Total spectra counting for insulin with different amount of CB7. (e) The % relative abundance of Y16 being cleaved by chymotrypsin. (f) Spectra counting for peptide identified for trypsin digestion.

Limited proteolysis for insulin with CB7. To prove that limited proteolysis can provide molecular details for the binding site of synthetic receptors on proteins, we firstly chose human insulin-CB7 complex as the model system. Human insulin is a relatively simple protein, containing 2 chains and 51 residues (Figure 5.1 a). The short sequence makes insulin good for limited proteolysis study, because even partially digested fragments can still be in the detection range of mass spectrometry. One of the special properties of

human insulin is the rare N-terminal phenylalanine (F1, in chain B) that makes it an ideal target for CB7. Because of previous amino acids studies, the mechanism for CB7 binding has been studied very thoroughly, and it shows that aromatic amino acids can bind to CB7 with high affinities.[21] As a seven-membered macrocycle in the cucurbit[n]uril family, CB7 contains a hydrophobic cavity and polar rims. Apolar aromatic rings of phenylalanine will be buried inside of the cavity, and oxygen atoms on the portal will stabilize positively charged atoms in vicinity. Because those two different interactions need be perfectly balanced for stable binding, all possible binding sites (including 3 phenylalanine and 4 tryptophan) on insulin would not have equal affinities to CB7. In reality, CB7 shows selective and specific recognition on F1, which has been revealed by its crystal structure (PDB ID: 3Q6E, Figure 5.1 b). The good solvent accessibility of F1 and its cationic amino group (N-terminal) contribute together to this result.

Based on these previous findings, chymotrypsin was chosen as the protease for limited proteolysis, because of its relatively high activity and specificity for the aromatic residues like phenylalanine and tryptophan that were potential binding sites of CB7. Freshly reconstituted insulin was mixed with CB7 in different molar ratios for 1 hr at 37 ^{0}C. The binding of CB7 is rapid and would reach equilibrium during 1-hr incubation.[12] The digestion time was 30 minutes. After being lyophilized and desalted, the digested fragments and cutting sites were identified by LC-MS/MS. The spectra counting of each identified peptide was also used to quantify cutting sites. To minimize the influence of instrumental variance, relative abundance (RA) was calculated and used to normalize the quantification.[36] Among all cutting sites of chymotrypsin, F1 showed the most obvious

changes. Without CB7, there were about 9% peptides with F1 being cleaved and 11% were not, which made around 40% of F1 were cut by chymotrypsin in 30 minutes (Figure 5.1 c). With CB7 mixed with insulin, the percentage of F1 cleaved peptides decreased, along with the increasing ratio of CB7 (Figure 5.1 c). This clearly indicated that CB7 would bind to F1 and block it from being digested by chymotrypsin. The activity of chymotrypsin was not decreased by CB7, because the total spectra counting for CB7-insulin mixture were even higher than that of insulin itself (Figure 5.1 d). The reason for the enhanced digestion efficiency is not clear, but the oligomerization and unfolding of insulin could partially explain it. Although insulin was always freshly prepared to avoid fibrillation, it was expected that dimers (induced by CB7, as shown in the crystal structure, Figure 5.1 b) would form, which can increase the local protein concentration and enhance the digestion efficiency of protease. Moreover, the unfolding of N-terminal of B chain shown in the crystal structure can increase its surface accessibility for proteolysis. Other than the decreased digestion on F1, not too many changes were observed for other cutting sites. For example, the %RA of Y16, which was the cutting site next to F1 and another potential binding site for CB7, was not significantly changed (Figure 5.1 e).

Trypsin was also used for limited proteolysis to enhance the sequence coverage, due to its different cutting sites compared to chymotrypsin. One of the limitations for proteolysis is the unevenly distributed cutting sites on protein sequence, which could make the generated peptides too long or too short for detection in MS. Just like the three consecutive aromatic residues in the C-terminal of insulin B chain (Figure 5.1 a), they all can be cut by chymotrypsin, and generate peptides consisting of 1 or 2 residues, which are

too short for the LC-MS method set up for sequencing peptides longer than 300 Da. On the other hand, there are two cutting sites for trypsin on the B chain, and digestion at these two sites would generate a peptide fragment with 7 amino acid residues, ideal for our MS method (Figure 5.1 a). Thus, limited proteolysis using trypsin could provide better information for this region. As shown in Figure 5.1 f, the spectral counting (SC) for those two cutting sites were increased with CB7, which indicated no blockage or binding of CB7 in the C-terminal of B chain. All results generated by limited proteolysis with both chymotrypsin and trypsin indicated that CB7 bound to F1, but not other sites on insulin.

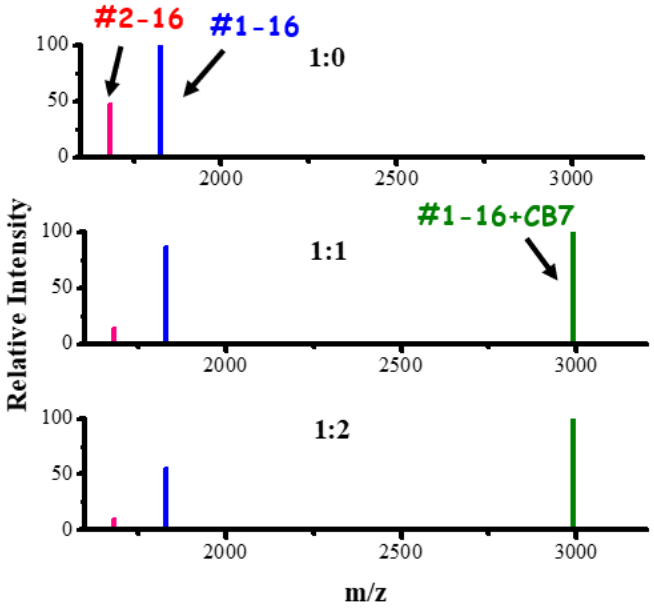


Figure 5.2. Relative intensity of peptides generated after limited proteolysis of insulin with CB7 in 1:0, 1:1, and 1:2 ratios by chymotrypsin, identified in MALDI-TOF-MS.

MALDI-TOF-MS was also used to verify the results obtained by LC-MS/MS. As shown in Figure 5.2, the intensity of peptides with F1 being cleaved (#2-16) became lower for the mixture with CB7, while the same peptide but with F1 (#1-16) was observed with

higher intensity. This was consistent with the results obtained by LC-MS/MS. More interestingly, the complex of peptide (#1-16) / CB7 could be detected in MALDI, which could provide a direct evidence for CB7 binding. The same complex was not seen in previous LC-MS/MS results. MALDI ionization is milder than the ESI used in the LC-MS/MS method; and it does not require column separation, both preserving the receptor-protein complex when analyzed by MALDI-TOF-MS.

Limited digestion of lysozyme with sclx4. The results acquired from applying limited proteolysis on the model complex of CB7-insulin has demonstrated its feasibility in locating the binding site of CB7. However, there is always debate that insulin is too short to represent other proteins with high complexity. To challenge this method, we next analyzed the binding to chicken lysozyme (referred as "lysozyme" henceforth), one globular protein with 129 amino acids, by sclx4, one of the calixarene macrocyclic molecules. Sclx4 differs from CB7 by possessing 4 negatively charged sulfonate groups on one side of the rim and a much smaller hydrophobic cavity. As a result, electrostatic interactions, rather than hydrophobic ones, dominates the binding of sclx4 to lysozyme. The crystal structure of sclx4-lysozyme revealed that one sclx4 would stably bind to R128 near the C-terminal of lysozyme through electrostatic interactions, and another sclx4 would bridge two lysozyme molecules and induce the tetramer formation (PDB ID: 4PRQ). The binding of sclx4 could passivate the lysozyme surface, as indicated by the decrease of solvent accessible surface area (SASA) at the N- and C-terminal (Figure 5.3a). Because of tetramer formation, the interfaces of each monomer also showed lower SASA (Figure 5.3

a). Furthermore, the decreased surface accessibility could also hinder the protease from approaching and lead to the decreased digestion efficiencies on these blocked areas.

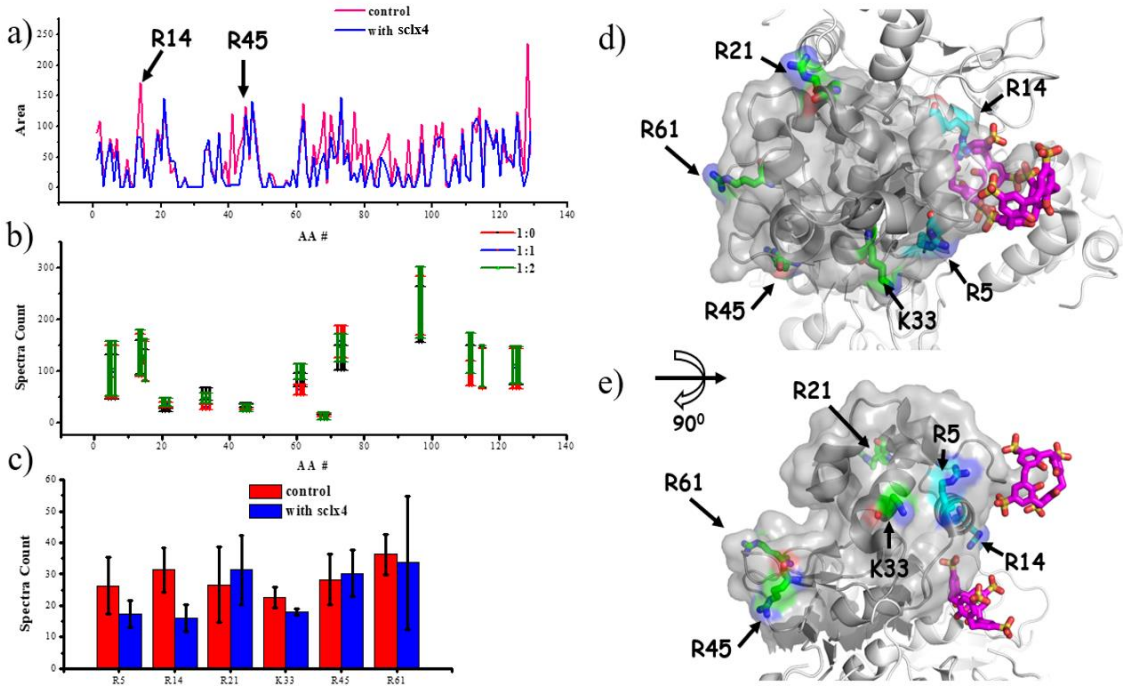


Figure 5.3. (a) Solvent accessible surface area (SASA) of lysozyme with (blue line) or without sclx4 (control, red line). Calculation is based on crystal structures of 4PRQ and 1DPX. (b) Spectra counting of each cutting site identified in LC-MS/MS, after fully trypsin digestion of lysozyme mixed with sclx4 in 1:0, 1:1, and 1:2 ratios. (c) Spectra counting of each cutting site identified in LC-MS/MS, after limited proteolysis of lysozyme with 1:0 or 1:2 ratios of sclx4. (d) Top and (e) left view of the crystal structure of lysozyme tetramer with sclx4 (PDB ID: 4PRQ). Sclx4 is shown in magenta stick, and one monomer of lysozyme is shown in gray surface. Specific residues are highlighted in either green or cyan stick.

Since the potential binding sites of sclx4 are lysine or arginine residues, trypsin is an ideal protease for digestion. The influence of sclx4 on the trypsin activity was negligible, since the fully digested lysozyme with or without sclx4 showed similar SC across the whole sequence (Figure 5.3 b). This also indicated that there was no influence of sclx4 on the peptide identifications by LC-MS/MS. As for the limited proteolysis, different changes

were observed for different cutting sites. With sclx4 bound, two cutting sites on lysozyme, R5 and R14, showed obvious decrease on spectra counting, while the other cutting sites, R21, K33, R45 and R61, did not change appreciable (Figure 5.3 c). Although the binding site of sclx4 on lysozyme should be R128, there was no confirmation of its identity in our LC-MS/MS results, due to the aforementioned limitations of this method. However, it was believed that the binding of sclx4 could not only block R128, but also influence other residues in vicinity, as shown in the SASA plot (Figure 5.3 a), and the crystal structure (Figure 5.3 d-e). R5 and R14 are located in the α-helix closest to sclx4, and the binding of sclx4 could dramatically decrease their SASA. On the other hand, R21, K33, R45 and R61 are remote to the binding site of sclx4, and they were less affected during the limited proteolysis. These results substantiate that limited proteolysis is capable of identifying the binding sites of sclx4 on lysozyme.

Digestion with nonspecific protease for synthetic receptors with unknown binding behaviors. Although trypsin and chymotrypsin have respective utility for CB7- and sclx4-protein binding studies, their success was based on the consistency between the cutting sites and the potential binding residues. If limited proteolysis is applied to another synthetic receptor with unknown binding preferences, choosing the correct protease becomes tricky. Thus, it would be beneficial to have one nonspecific protease, such as proteinase K, working in the same paradigm. To verify its performance, lysozyme with sclx4 was used as the model system, again. Because of the nonspecificity, proteinase K would cut more sites compared to trypsin or chymotrypsin. Sated another way, the coverage of the protein sequence would be enhanced. The digestion time of proteinase K

was limited to 10 minutes to prevent generation of peptide fragments that are too short for LC-MS/MS. A second step of digestion by trypsin was performed to cut peptides that were too long. The main cutting sites of trypsin, arginine and lysine, are rarely cut by proteinase K. As a result, the cutting sites of proteinase K can be easily differentiated from trypsin's. The SC of each peptide was also used for the quantification of cutting sites that generated the peptide. To eliminate the run-to-run variations, %SC was calculated by dividing the SC of each cutting site by the total number of SC.

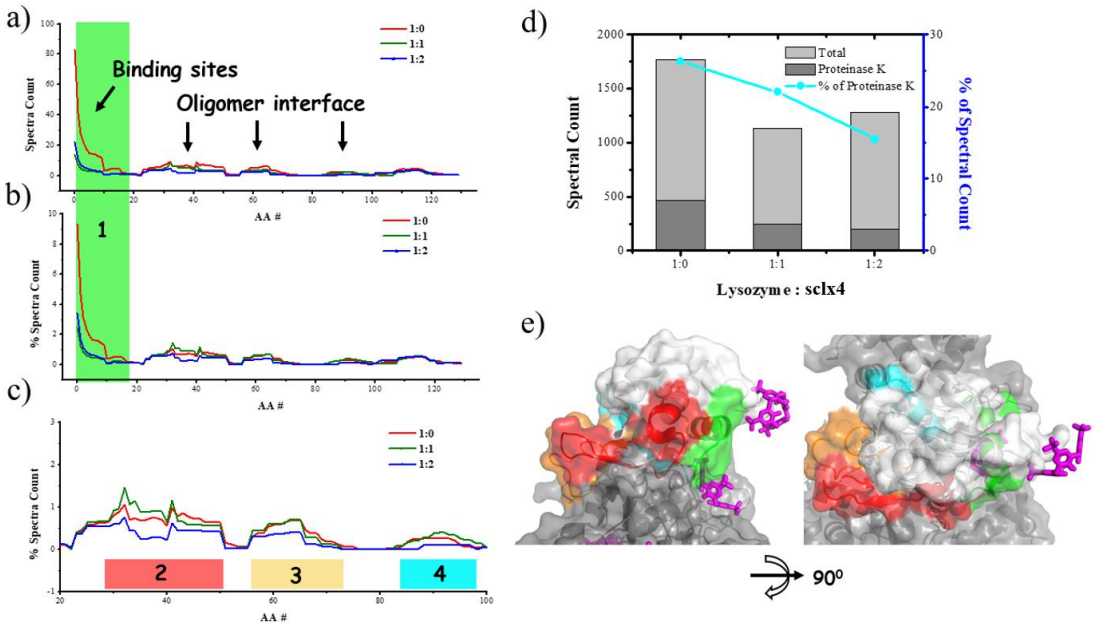


Figure 5.4. (a) Spectra counting and (b) % spectra counting for cutting sites of lysozyme/sclx4 by proteinase K. The ratios of lysozyme:sclx4 are 1:0, 1:1, and 1:2. N-terminal with large decrease is highlighted in green. (c) The zoom in view of % spectra counting changes for sequence of #20-100. Three regions are colored in red, yellow, and cyan. (d) The total spectra counting for cutting sites of proteinase K or trypsin, which are shown in gray bars. The percentage of cutting sites of proteinase K is plotted in cyan line. (e) Crystal structure of lysozyme tetramer complexed with sclx4 (PDB ID: 4PRQ). One lysozyme monomer is shown in white surface, and the other parts are shown in gray surface. Sclx4 is shown in magenta stick. Areas identified in a-c with changes on spectra counting are colored in green, red, yellow, and cyan, correspondingly.

The SC or %SC for each cutting site of proteinase K were smooth in moving average, and plotted against the protein sequence, from N- to C- terminal (Figure 5.4 a-c). With sclx4 bound to lysozyme (Figure 5.4 d), The total number of SC for proteinase K's cutting sites was decreased. indicating the binding of sclx4 would block and protect the lysozyme surface from proteinase K cleavage. By looking at the SC curve cross the whole sequence (Figure 5.4 a), the area with the most dramatic changes for lysozyme with sclx4 (1:1 or 1:2 ratios) was at the N-terminal (region 1, highlighted in green). A similar trend was also observed for the %SC plot in Figure 5.4 b, which strongly indicated that region 1 was the binding site of sclx4. As shown in the crystal structure, region 1 (in green) verified the assumption by being the closest to the binding sites of sclx4 (magenta).

Besides region 1, there were three more regions (region 2-4, shown in red, yellow, and cyan, respectively) that also showed decrease in either SC or % SC. In the zoom-in plot (Figure 5.4 c), an interesting phenomenon was observed for these three regions: there was no reduction in %SC at lower concentrations of sclx4 (1:1 ratio), but a more obvious decrease at higher concentrations of sclx4 (1:2 ratio). This indicated that these three regions were not the primary binding sites of sclx4 but could be influenced by higher concentrations of sclx4. By virtue of the crystal structure, it can be said that the tetramer formation is responsible for the reduced %SC. However, it is difficult to corroborate this conclusion based only on the limited proteolysis result

To get more supporting information, the interaction of sclx4 to lysozyme was also tested by other methods. The lysozyme surface blockage by sclx4 was observed through the fluorescamine labeling, as the fluorescence intensity for lysozyme-sclx4 mixtures

decreased (Figure 5.5 a). The intrinsic fluorescence profile of lysozyme with or without sclx4 did not show a peak shift, which indicated no obvious conformational change in lysozyme with sclx4 (Figure 5.5 b). Additionally, the thermal stability test also showed that the binding of sclx4 could stabilize lysozyme, as indicated by the higher shift of the melting temperature. However, there was no obvious difference for mixtures with different ratios of sclx4 (Figure 5.5 c). The lysozyme/sclx4 mixture peak could be directly verified by MALDI-TOF-MS, and there was also a small peak corresponding to lysozyme bound by two sclx4 (Figure 5.5 d). Collectively, these results support that (1) sclx4 could bind to lysozyme and block some of its surface lysine resides, and (2) region 1 should be the primary binding site. However, there was no obvious evidence to support the tetramer formation impacting the changes on regions 2-4.

On-surface limited proteolysis of the complex of lysozyme-sclx4. Although previous results have bolstered the usefulness of limited in identifying the binding sites of synthetic receptors on proteins, protein oligomerization also requires consideration as it complicates analysis by inducing changes on digestion efficiency. To minimize the effect of oligomerization caused by synthetic receptors, on surface digestion, in which a layer of protein could be formed on a surface via nonspecific adsorption, was adopted. In this way, there will be less chances for proteins to form oligomer during the process. Firstly, lysozyme was immobilized onto the hydrophilic surface of the 96-well plate via physical adsorption (Figure 5.6 a). Sclx4 was then added to bind to the adsorbed lysozymes. Afterwards, the lysozymes were digested by trypsin, while the sites blocked by sclx4 were protected by sclx4. In the final step, digested fragments were collected by multiple washing

and identified in LC-MS/MS. Lysozyme adsorbed on surface but not incubated with sclx4 was used as the control. As with previous sections, SC and %SC for each cutting site were used for quantification (Figure 5.6 b-c).

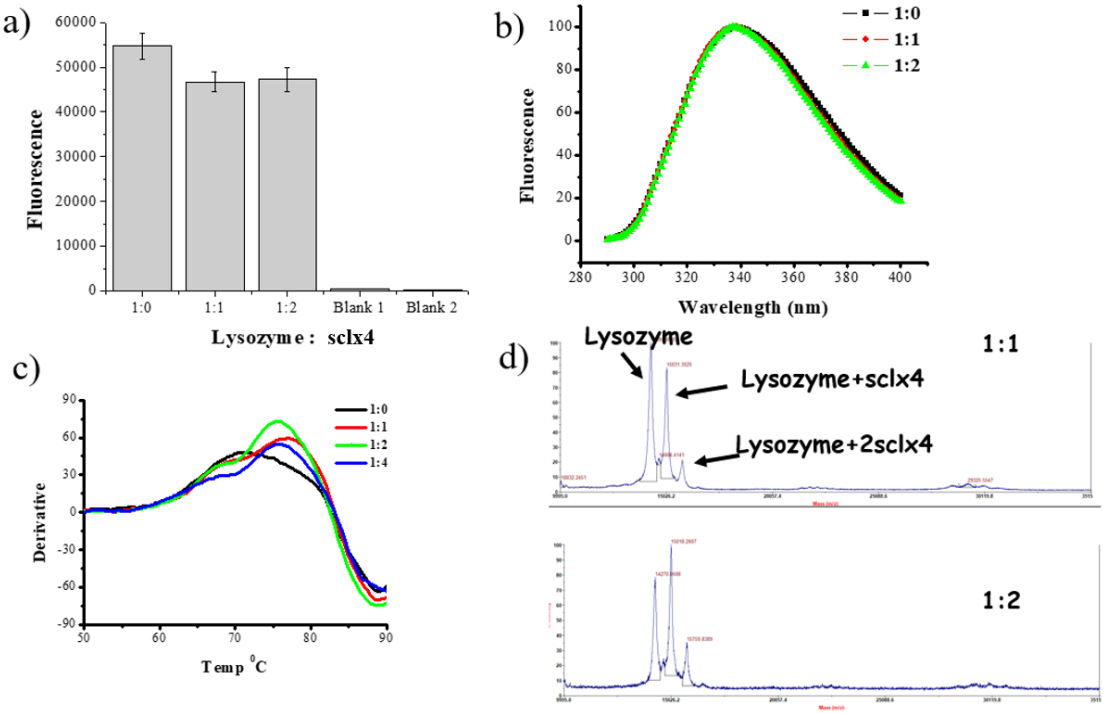


Figure 5.5. (a) Fluorescamine labeling profile of lysozyme mixed with sclx4 in 1:0, 1:1, and 1:2 molar ratios. Sclx4 in different concentrations are measured as blank 1 and 2. (b) The intrinsic fluorescence of lysozyme mixed with sclx4. Excitation is at 280 nm. (c) Thermal stability test for lysozyme/sclx4 mixture. 4× SYPRO Orange is used to monitor the real-time fluorescence changes. Derivative of fluorescence signal is plotted against temperatures. (d) MALDI-TOF results for lysozyme with or without sclx4.

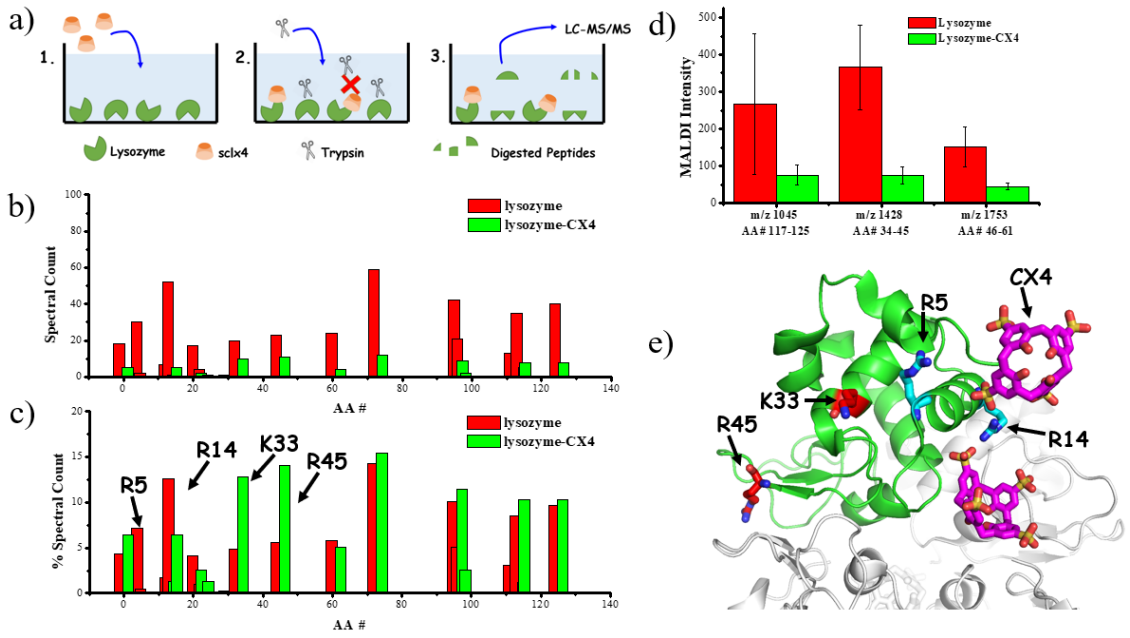


Figure 5.6. (a) The scheme for on surface limited proteolysis. Lysozyme is immobilized on surface by non-specific binding, and sclx4 was added afterwards. After washing, 1:50 mass ration of trypsin is added, and 0.1% FA is used to terminate the reaction after 10 minutes. Digested peptides are collected and loaded to LC-MS/MS. (b) SC for each cutting site identified in LC-MS/MS. Lysozyme itself is used as control and colored in red. Lysozyme/sclx4 mixture is colored in greed. (c) The normalized SC in (b). Normalization is done by dividing each cutting site's SC by the total number of SC. (d) Intensity of representative peptides by MALDI. (e) Crystal structure of sclx4 bound to lysozyme. Sclx4 is shown in magenta stick, and one monomer of lysozyme is shown in green cartoon. Other lysozymes in tetramer are shown in gray cartoon. Representative residues are shown in either cyan or red sticks.

For lysozyme incubated with sclx4, we observed decreased SC for every cutting site identified in LC-MS/MS (Figure 5.6 b). This was also further verified by MALDI-TOF-MS (Figure 5.6 d). Moreover, the activity of trypsin cannot be inhibited by sclx4, as demonstrated in the previous section. As a result, our data point out that, multiple binding events of sclx4 on one protein are present, due to either unfolding of the protein on the surface, or the high concentration of sclx4 added. In Figure 5.6 c, the SC of each cutting

site was normalized to the total number of SC to generate % SC to show the relative change of each cutting site. It was interesting to see the dramatic decrease for both R5 and R14, which were both close to the binding site of sclx4 (Figure 5.6 e). This implied that on-surface limited proteolysis could also provide accurate information for the binding sties of synthetic receptors. Compared to in-solution limited proteolysis, on-surface limited proteolysis boasts the advantage of circumventing interferences from oligomerization. However, this approach suffers from the lower SC for every cutting site. This disadvantage could be mitigated by normalization.

5.4: Conclusions

Limited proteolysis has proven to be a useful tool in identifying the binding sites of synthetic receptors on proteins, based on the changed proteases accessibility of residues near the binding sites. The purview of proteases with different cutting sites makes this method versatile for different synthetic receptors. By choosing proteases wisely based on the binding preference of synthetic receptors, more precise information about the binding sites could be obtained. If little information for synthetic receptors is available, then nonspecific proteases, such as proteinase K, could be used to deduce the position of the binding site. Moreover, on-surface limited proteolysis can also be employed to eliminate interferences stemming from oligomerization. As a rapid and routine screening method to evaluate the targeting ability of synthetic receptors, limited proteolysis will greatly facilitate the development and optimization of synthetic receptors in biomedical applications. This would be a very welcomed asset to the repertoire of life science and pharmaceutical industries that focus on the recognition and modulation of proteins.

References

(1) Teague, S. J. Implications of protein flexibility for drug discovery. *Nature reviews Drug discovery* **2003**, *2*, 527.

(2) Keskin, O.; Gursoy, A.; Ma, B.; Nussinov, R. Principles of protein− protein interactions: What are the preferred ways for proteins to interact? *Chemical reviews* **2008**, *108*, 1225-1244.

(3) Scott, D. E.; Bayly, A. R.; Abell, C.; Skidmore, J. Small molecules, big targets: drug discovery faces the protein–protein interaction challenge. *Nature Reviews Drug Discovery* **2016**, *15*, 533.

(4) Beck, A.; Goetsch, L.; Dumontet, C.; Corvaïa, N. Strategies and challenges for the next generation of antibody–drug conjugates. *Nature reviews Drug discovery* **2017**, *16*, 315.

(5) Hussein, H. A.; Geneix, C.; Petitjean, M.; Borrel, A.; Flatters, D.; Camproux, A.-C. Global vision of druggability issues: applications and perspectives. *Drug discovery today* **2017**, *22*, 404-415.

(6) Onishi, Y.; Eshita, Y.; Ji, R.-C.; Kobayashi, T.; Onishi, M.; Mizuno, M.; Yoshida, J.; Kubota, N. Supermolecular drug challenge to overcome drug resistance in cancer cells. *Drug discovery today* **2018**.

(7) Milroy, L.-G.; Grossmann, T. N.; Hennig, S.; Brunsveld, L.; Ottmann, C. Modulators of protein–protein interactions. *Chemical reviews* **2014**, *114*, 4695-4748.

(8) van Dun, S.; Ottmann, C.; Milroy, L.-G.; Brunsveld, L. Supramolecular chemistry targeting proteins. *Journal of the American Chemical Society* **2017**, *139*, 13960-13968.

(9) Webber, M. J.; Langer, R. Drug delivery by supramolecular design. *Chemical Society Reviews* **2017**, *46*, 6600-6620.

(10) Grad, J.-N.; Gigante, A.; Wilms, C.; Dybowski, J. N.; Ohl, L.; Ottmann, C.; Schmuck, C.; Hoffmann, D. Locating large, flexible ligands on proteins. *Journal of chemical information and modeling* **2018**, *58*, 315-327.

(11) Kovalenko, E.; Vilaseca, M.; Díaz-Lobo, M.; Masliy, A.; Vicent, C.; Fedin, V. P. Supramolecular adducts of cucurbit [7] uril and amino acids in the gas phase. *Journal of The American Society for Mass Spectrometry* **2016**, *27*, 265-276.

(12) Hennig, A.; Bakirci, H.; Nau, W. M. Label-free continuous enzyme assays with macrocycle-fluorescent dye complexes. *Nature methods* **2007**, *4*, 629.

(13) Liu, S.; Ruspic, C.; Mukhopadhyay, P.; Chakrabarti, S.; Zavalij, P. Y.; Isaacs, L. The cucurbit [n] uril family: prime components for self-sorting systems. *Journal of the American Chemical Society* **2005**, *127*, 15959-15967.

(14) Gamal-Eldin, M. A.; Macartney, D. H. Selective molecular recognition of methylated lysines and arginines by cucurbit [6] uril and cucurbit [7] uril in aqueous solution. *Organic & biomolecular chemistry* **2013**, *11*, 488-495.

(15) Danylyuk, O.; Fedin, V. P. Solid-state supramolecular assemblies of tryptophan and tryptamine with cucurbit [6] uril. *Crystal Growth & Design* **2011**, *12*, 550-555.

(16) Mutihac, L.; Lee, J. H.; Kim, J. S.; Vicens, J. Recognition of amino acids by functionalized calixarenes. *Chemical Society Reviews* **2011**, *40*, 2777-2796.

(17) Pejaver, V.; Hsu, W. L.; Xin, F.; Dunker, A. K.; Uversky, V. N.; Radivojac, P. The structural and functional signatures of proteins that undergo multiple events of post-translational modification. *Protein Science* **2014**, *23*, 1077-1093.

(18) Liu, Y.; Perez, L.; Mettry, M.; Easley, C. J.; Hooley, R. J.; Zhong, W. Self-aggregating deep cavitand acts as a fluorescence displacement sensor for lysine methylation. *Journal of the American Chemical Society* **2016**, *138*, 10746-10749.

(19) Attar, A.; Ripoli, C.; Riccardi, E.; Maiti, P.; Li Puma, D. D.; Liu, T.; Hayes, J.; Jones, M. R.; Lichti-Kaiser, K.; Yang, F. Protection of primary neurons and mouse brain from Alzheimer's pathology by molecular tweezers. *Brain* **2012**, *135*, 3735-3748.

(20) Prabhudesai, S.; Sinha, S.; Attar, A.; Kotagiri, A.; Fitzmaurice, A. G.; Lakshmanan, R.; Ivanova, M. I.; Loo, J. A.; Klärner, F.-G.; Schrader, T. A novel "molecular tweezer" inhibitor of α-synuclein neurotoxicity in vitro and in vivo. *Neurotherapeutics* **2012**, *9*, 464-476.

(21) Urbach, A. R.; Ramalingam, V. Molecular recognition of amino acids, peptides, and proteins by cucurbit [n] uril receptors. *Israel Journal of Chemistry* **2011**, *51*, 664-678.

(22) Hewitt, S. H.; Filby, M. H.; Hayes, E.; Kuhn, L. T.; Kalverda, A. P.; Webb, M. E.; Wilson, A. J. Protein Surface Mimetics: Understanding How Ruthenium Tris (Bipyridines) Interact with Proteins. *ChemBioChem* **2017**, *18*, 223-231.

(23) Lee, J.; Perez, L.; Liu, Y.; Wang, H.; Hooley, R. J.; Zhong, W. Separation of Methylated Histone Peptides via Host-Assisted Capillary Electrophoresis. *Analytical chemistry* **2018**, *90*, 1881-1888.

(24) Bosmans, R. P.; Briels, J. M.; Milroy, L. G.; de Greef, T. F.; Merkx, M.; Brunsveld, L. Supramolecular Control over Split-Luciferase Complementation. *Angewandte Chemie International Edition* **2016**, *55*, 8899-8903.

(25) Doolan, A. M.; Rennie, M. L.; Crowley, P. B. Protein recognition by functionalized sulfonatocalix [4] arenes. *Chemistry–A European Journal* **2018**, *24*, 984-991.

(26) de Vink, P. J.; Briels, J. M.; Schrader, T.; Milroy, L. G.; Brunsveld, L.; Ottmann, C. A Binary Bivalent Supramolecular Assembly Platform Based on Cucurbit [8] uril and Dimeric Adapter Protein 14-3-3. *Angewandte Chemie* **2017**, *129*, 9126-9130.

(27) Guagnini, F.; Antonik, P. M.; Rennie, M. L.; O'Byrne, P.; Khan, A. R.; Pinalli, R.; Dalcanale, E.; Crowley, P. B. Cucurbit [7] uril-Dimethyllysine Recognition in a Model Protein. *Angewandte Chemie International Edition* **2018**, *57*, 7126-7130.

(28) McGovern, R. E.; Fernandes, H.; Khan, A. R.; Power, N. P.; Crowley, P. B. Protein camouflage in cytochrome c–calixarene complexes. *Nature chemistry* **2012**, *4*, 527.

(29) Bier, D.; Rose, R.; Bravo-Rodriguez, K.; Bartel, M.; Ramirez-Anguita, J. M.; Dutt, S.; Wilch, C.; Klärner, F.-G.; Sanchez-Garcia, E.; Schrader, T. Molecular tweezers modulate 14-3-3 protein–protein interactions. *Nature chemistry* **2013**, *5*, 234.

(30) Chinai, J. M.; Taylor, A. B.; Ryno, L. M.; Hargreaves, N. D.; Morris, C. A.; Hart, P. J.; Urbach, A. R. Molecular recognition of insulin by a synthetic receptor. *Journal of the American Chemical Society* **2011**, *133*, 8810-8813.

(31) Bloch, W. M.; Champness, N. R.; Doonan, C. J. X-ray Crystallography in Open-Framework Materials. *Angewandte Chemie International Edition* **2015**, *54*, 12860-12867.

(32) Lee, J. W.; Shin, M. H.; Mobley, W.; Urbach, A. R.; Kim, H. I. Supramolecular Enhancement of Protein Analysis via the Recognition of Phenylalanine with Cucurbit [7] uril. *Journal of the American Chemical Society* **2015**, *137*, 15322-15329.

(33) Lee, J. W.; Heo, S. W.; Lee, S. J. C.; Ko, J. Y.; Kim, H.; Kim, H. I. Probing conformational changes of ubiquitin by host–guest chemistry using electrospray ionization mass spectrometry. *Journal of The American Society for Mass Spectrometry* **2013**, *24*, 21-29.

(34) Lomenick, B.; Hao, R.; Jonai, N.; Chin, R. M.; Aghajan, M.; Warburton, S.; Wang, J.; Wu, R. P.; Gomez, F.; Loo, J. A. Target identification using drug affinity responsive target stability (DARTS). *Proceedings of the National Academy of Sciences* **2009**, pnas. 0910040106.

(35) Smits, A. H.; Vermeulen, M. Characterizing protein–protein interactions using mass spectrometry: challenges and opportunities. *Trends in biotechnology* **2016**, *34*, 825-834.

(36) Walkey, C. D.; Olsen, J. B.; Song, F.; Liu, R.; Guo, H.; Olsen, D. W. H.; Cohen, Y.; Emili, A.; Chan, W. C. Protein corona fingerprinting predicts the cellular interaction of gold and silver nanoparticles. *ACS nano* **2014**, *8*, 2439-2455.

CHAPTER 6: Map Molecular Details of Protein Conformational Change in Biological Corona of Nanomaterials Using Limited Proteolysis

6.1: Introduction

Engineered nanomaterials (ENMs) have provided great opportunities to facilitate and accelerate the progress across a wide range of biomedical applications, especially as nanomedicine for cancer care.[1] Despite enormous developments, including numerous approvals for clinical usage, the nanomedicine community is facing many obstacles and challenges, including how to improve the specific targeting of tumors.[2-4] Since the approval of the first nanomedicine for cancer by the US FDA in 1995, the enhanced permeability and retention (EPR) effect of ENMs has been the foundation to advance the passive targeting of nanomedicine to solid tumors, based on the fact that tumor vessels are more permeable and can preferentially retain large molecules.[4,5] However, the heterogeneity of cancer tissues could hinder the uniform distribution of nanomedicines, and the unwanted uptake by normal tissues has also remained a problem.[1,2] Inspired by the success of antibody conjugated drugs,[6,7] ENMs functionalized with affinity proteins have attracted growing attention and have emerged as promising tools for active targeting.[8] One representative strategy for active targeting is to coat ENMs with plasma proteins (e.g. transferrin) or monoclonal antibodies, which could specifically be recognized by the corresponding targets (e.g. transferrin receptor, human epidermal growth factor receptor 2, or HER2, and epidermal growth factor receptor, or EGFR) overexpressed in the membrane of cancer cells or tumor-supporting cells.[2,8-10] Another strategy borrows the mechanisms of immuno-oncology, in which proteins associated with immune activations (e.g. CD86,

interferon gamma, or IFN-γ) are used to coat ENMs to stimulate the immune response to cancer cells.[2,11,12] However, there is very little improvement on the accumulations of ENMs in tumors for active targeting, compared to passive targeting.[1,13,14]

One possible reason for the failure of these strategies is that protein arrangement on ENMs is difficult to control, and the molecular details of the protein on ENM surface cannot be routinely mapped out, due to the lack of rapid and robust analytical tools.[15-19] The functional region of a protein, e.g. the epitope, can only be "seen" by the targets when exposed to the solvent. The binding of proteins on ENMs can result in situations in which the epitope can be blocked.[20] Moreover, proteins on ENMs can rearrange their structures or undergo conformational change to fit the surface of ENMs, which indicates that the function of those proteins, including molecular recognition or enzymatic abilities, would be compromised.[21] The conformational change of proteins could also expose cryptic epitopes that were buried inside, and change the fate of ENMs by facilitating the recognition by scavenger receptors and enhancing the uptake and clearance by the mononuclear phagocyte system (MPS).[16,18,22] Although the composition and kinetics of the protein corona have been widely studied,[21-24] more sophisticated molecular details of the proteins that compose the corona are still lacking due to the lack of robust analytical methods.[15]

To address the methodology discrepancy regarding what ENMs present and "what cells see",[20] immuno-probes have been applied to detect the orientation of protein epitopes on the surface of ENMs with standard proteins (e.g. transferrin) or in biological matrices (e.g. human serum). Results have shown that only a very small portion (~3.5%) of the

proteins adsorbed on the surface of ENMs can maintain the functional epitopes externally with the right orientation[15,19,25] However, this strategy cannot provide more molecular details including conformational changes, and the availability of immune-probes also restricts the method's applications. As a powerful tool for exploring protein structures, NMR spectroscopy has been shown to be very useful for study of protein-ENM interactions, with the help of isotope labeling and hydrogen/deuterium exchange (HDX).[26,27] However, the limitations of NMR, e.g. the size of protein cannot be too large, is still a big obstacle.[28] On the other hand, mass spectrometry (MS) coupled with cross-linking, isotope exchange, or chemical modification, has also been successfully used to study protein binding sites or conformational changes upon interaction with ENMs.[29-31] However, these chemical reagents are limited by the reactivity and detectability, which could complicate the identification of protein sequences. In contrast, proteases, with high reactivity and specificity, have proved themselves useful and robust at mapping conformational change of proteins.[32] The proteolysis method has already been used for proteins bound to NPs. For example, the binding sites of proteins (e.g. α-synuclein) on ENMs (e.g. gold nanoparticles) could also be explored by MS, after one step protease digestion and separation of adsorbed peptides.[33] However, the limitation of the method is also obvious: it is the peptide not the protein interacting with the ENMs after digestion, and the binding behaviors of peptides and proteins are not the same because protein motifs could compose peptides that are not sequentially linked.[34,35] To overcome this drawback, limited proteolysis could be used for this purpose. This method has been applied to probe

protein-protein interactions and protein-ligand interactions, but not yet on studying protein-ENMs interaction.[35,36]

In this study, we aim to obtain comprehensive molecular details of proteins on the surface of ENMs, including the binding motifs, conformational changes, and the orientation of functional epitopes, which could explain the biological outcomes of protein-ENM conjugates and contribute to the ongoing characterization and rational design of nanomedicines. Complementary limited proteolysis by both trypsin and chymotrypsin is used to cut the flexible and exposed regions of the protein, which could also be influenced by the binding of ENMs.[37] After being separated from the bulky parts by centrifugal filters, small peptide fragments could go through another step of proteolysis, resulting in proper sizes identifiable by LC-MS/MS. Comparing the results for proteins themselves (as the controls) and protein-ENM mixtures could show us the detailed molecular difference caused by binding to ENMs.

6.2 Materials and Methods

Chemicals and Biochemicals. Fluorescamine, Triton X-100, iodoacetamide (IAM) and all proteins including trypsin were obtained from Sigma-Aldrich (St. Louis, MO, USA). Hydrogen peroxide (H_2O_2), dimethyl sulfoxide (DMSO), Coomassie brilliant blue R-250, formic acid, acetonitrile (ACN), urea, dithiothreitol (DTT), anhydrous sodium phosphate dibasic, sodium chloride, ammonium bicarbonate, acrylamide, sodium dodecyl sulfate (SDS) and ammonium persulfate were purchased from Fisher Scientific (Waltham, MA, USA). 2× Laemmli sample buffer was obtained from Bio-Rad (Hercules, CA, USA). All proteins were purchased from Sigma-Aldrich (St. Louis, MO). Ultrapure water with

electric resistance larger than 18.2 MΩ was produced in-house, using the Milipore Milli-Q water purification system (Billerica, MA, USA).

Nanoparticles (NPs). Carboxylated polystyrene NPs with an average diameter of 48 nm (PS48) were purchased from Polysciences (Warrington, PA, USA). Silica nanoparticles with average diameter of 50 nm (Si50) were obtained from nanoComposix (San Diego, CA, USA) Iron oxide NPs of 10 nm was obtained from the HSPH-NIEHS Nanosafety Center of the NHIR consortium.

Trypsin Limited Digestion and In Gel Visualization. Proteins (0.2 mg/ml) were incubated with 10nM of PS48 or Si50 in PBS buffer (10 mM phosphate buffer at pH 7.4, with 137 mM NaCl and 2.7 mM KCl) for 1 hr at 37 ^{0}C. Then, trypsin was added in a 1:50 mass ratio. After 10 minutes incubation at 37 ^{0}C, 2× Laemmli sample buffer was added in 1:1 ratio to terminate the digestion. Proteins themselves in PBS were also processed in the same manner and used as controls. To visualize the digestion efficiency, 20 µL of each mixture were loaded and separated in SDS-PAGE (4% polyacrylamide stacking gel, 12% polyacrylamide separation gel, with 0.1% SDS). Coomassie brilliant blue R-250 was used to stain the gel, and the band intensities were quantified using ImageJ. [39] The band intensity of protein was normalized to that of the corresponding control, to represent the relative digestion efficiency changed by NPs.

Catalase Activity Assay. In a 16 x 125 mm borosilicate glass tube containing 100µl 0.5 mg/ml catalase in PBS buffer, 100 µl 1% Triton X-100 and 200 µl 30% H_2O_2 were added. After mixing, the height of foam generated by oxygen was recorded every 20 seconds. [40] The slope of the initial linear range of the plot of foam height against time

($V_{initial}$) would be used as the activity of catalase (EC_0, in mm/s). To measure the activity influenced by PS48 NPs, 0.5 mg/ml catalase was incubated with either 5 or 10 nM PS48 NPs for 1 hr at 37 ^{0}C, followed by the addition of Triton X-100 and H_2O_2. Then, foam heights were recorded, and the catalase activity (EC_{PS48}) calculated from $V_{initial}$ was normalized by dividing EC_0.

Complementary Limited Digestion. Trypsin and chymotrypsin were used to perform limited digestion, separately. The same procedure in the previous part was used, except for chymotrypsin in which 1 mM calcium was added. Following specific time of digestion (10 or 60 minutes), 8 M urea was added in 1:1 ratio to terminate reaction. Then, DTT was added with a final concentration of 5 mM, and the mixture was incubated at 56 ^{0}C for 30 minutes. After the mixture was cooled to room temperature, IAM was added to 10 mM, followed by a 20 minutes incubation in the dark. The whole solution was filtered through a 30 kDa cutoff Amino centrifugal filter in 14,000 ×g for 10 minutes, and PBS buffer (pH 7.4) was used to wash twice. All the filtrates were collected together, while the solution left in the filter (supernatant) was also collected through reverse centrifugation at 2000 ×g for 2 mins. Both filtrate and supernatant were diluted with 50 mM ammonium bicarbonate, and trypsin or chymotrypsin were added in 1:50 mass ratio. After the overnight digestion at 37 ^{0}C, all those solutions were lyophilized and desalted.

LC-MS/MS. Peptides generated in both filtrate and supernatant were subjected to liquid chromatographic separation with a CapLC system (Waters Corporation, Milford, MA, USA) that was connected to a Finnigan LTQ mass spectrometer (Fisher Scientific, Maltham, MA, USA) with an ESI nanospray source. The desalted and lyophilized samples

were dissolved in 0.1% formic acid (FA) in water and then loaded into a homemade reverse phase (RP) trap column (150 μm i.d × 2 cm) linked to a RP separation column (75 μm i.d. × 10 cm). Both columns were packed in home with 5 μm C18 silica beads (Dr. Maisch HPLC GmbH, Germany). For mobile phases, 0.1% FA in water (solution A) and 0.1% FA in CAN (solution B) were used. The separation was started at a flow rate of around 200 nL/min (after flow splitting) with 2% B for 20 minutes, followed by a linear gradient increase to 50% B during a period of 60 minutes. After that, 80% B for 20 minutes and 2% B for 30 minutes were used to wash and equilibrium the columns. The positive mode full MS scan was carried within a mass range of 300-2000 m/z and MS/MS was operated in a data-dependent mode for the 10 peaks with highest intensities. Collision induced dissociation (CID) with 35% collision energy was used as fragmentation activation.

Data Searching and Analysis. MSGF Plus was used to identify peptide sequences using the sequences of catalase (bovine), serum albumin (human), and transferrin (human) downloaded from UniProt. These sequences with reverse order were used as decoys. Iodoacetamide of cysteine was defined as constant modification, and oxidation of methionine was defined as dynamic modifications. The cutoff for parent ion tolerance was set to 3 Da, and the Q value that represent false discovery rate (FDR) was set to be smaller than $1E^{-4}$. The spectra counting (SC) of each peptide was used to quantify the cutting sites to generate this peptide. The relative abundance (RA) of the cutting site i was calculated by Equation 1:

$$\%RA = \frac{SC_i}{\sum SC} \times 100\% \quad (1)$$

Measurement of NPs Size and Concentration Change. Ten nanomolar of the PS48 NPs were incubated with different concentrations of proteins (0-8 μM) in 10 mM PBS buffer (pH 7.4) at 37 ^{0}C for 1 hr. After proper dilution by the PBS buffer (100-2000 fold), nanoparticle tracking analysis (NTA) was used to measure the hydrodynamic diameter and concentration of the PS48 NPs. For each sample, the maximum count was normalized to 1. On the other hand, the same proteins (0-8μM) were digested at first with trypsin (1:50 mass ratio) for overnight at 37 ^{0}C, and the digested peptides were then incubated with 10 nM PS48 NPs in PBS buffer (pH 7.4) for 1 hr at 37 ^{0}C. Afterwards, the hydrodynamic diameters and concentrations of PS48 NPs in these mixtures were measured by NTA.

6.3: Results and Discussion

Limited trypsin proteolysis visualized by SDS-PAGE. In our previous work, fluorescamine labeling had been developed to quantify the interactions between proteins and ENMs. Based on the screening results, proteins, including human transferrin (Tf) and bovine catalase (CAT), demonstrated obvious unfolding and strong binding with polystyrene (PS) NPs, and were used as positive hits for this study.[38,39] Besides Tf and CAT, human serum albumin (HSA) was chosen as the control, because of the relatively lower fluorescence enhancement shown in previous fluorescamine labeling screening. Fluorescamine mainly labels lysine residues on protein surfaces, and the protein surface blockage induced by ENMs could change the number of surface lysine residues, which results in different fluorescence profiles. Since lysine residues are also the cutting sites for

trypsin, it is expected that the digestion efficiency of trypsin on transferrin and catalase with PS NPs would also change.

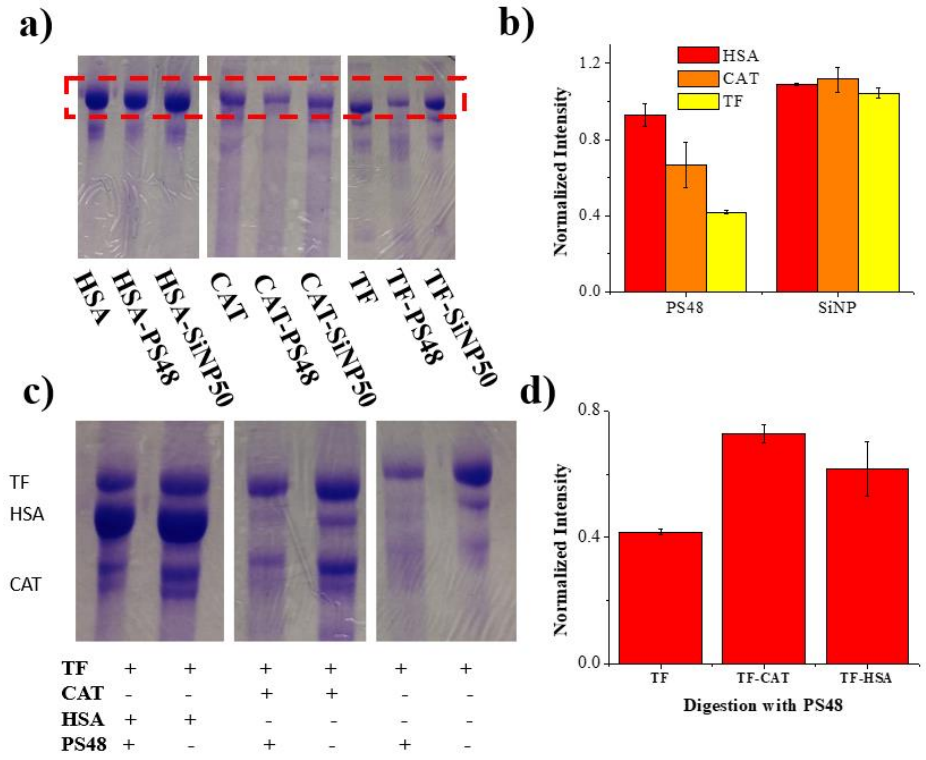


Figure 6.1. (a) HSA, CAT, and Tf were digested by trypsin in 1:50 mass ratio to proteins for 10 mins under room temperature, and were visualized by SDS-PAGE with Coomassie blue staining. 4μM proteins were also incubated with 10nM PS48 or SiNP50 first, then were digested and visualized. (b) The protein band intensities for PS48 or SiNP50 mixtures were quantified by ImageJ and normalized to protein controls. (c) Tf and another protein (either HSA or CAT) were incubated with PS48 together, and the mixture was digested by trypsin for 10 minutes. Coomassie blue staining was used, and (d) the protein band intensities of Tf were quantified by ImageJ.

To prove this assumption, SDS-PAGE was used to visualize the trypsin digestion efficiency. Trypsin in a 1:50 mass ratio to proteins was added for digestion. After 10 minutess, 2× Laemmli sample buffer was added in a 1:1 volume ratio to terminate the reaction. The digested peptides would be separated from the intact proteins in SDS-PAGE,

and Coomassie blue was used for staining. For Tf and CAT, it was evident to see less intact proteins remained in the mixture containing the PS48 NPs (Figure 6.1 a). The band intensities of the intact proteins were also quantified by ImageJ (Figure 6.1 b), and it was verified that the decreases of Tf (0.4) and CAT (0.65) were much more obvious than HSA (0.95). This indicated that the trypsin digestion efficiency was enhanced for the mixture of the protein and the PS48 NPs. The enhancement was not due to the change of trypsin activity by PS48 NPs, because there was no obvious change for HSA incubated with PS48 NPs at the same condition (Figure 6.1 a). Beside the intact proteins, the digested peptides were also stained, as shown in the multiple bands at the lower part of the gel. Changes were also observed for the intensities of these peptides. For Tf, there was an increase for the band just below the intact protein, which indicated that the enhanced digestion of Tf could generate more long peptides (Figure 6.1a). For CAT, the band next to the intact protein showed decreased intensity, which indicated the long peptide fragments of CTA could also be digested faster with the PS48 NPs (Figure 6.1a). Moreover, Si NP, which showed relatively small changes with all these three proteins used in the previous fluorescamine screening, was used as another control. As expected, no change was observed in the trypsin digestion efficiency of all three proteins incubated with Si NPs. These results verified the assumption that the unfolding of Tf and CAT induced by binding to PS48 NPs could show differences in limited proteolysis.

Furthermore, the influence of PS NPs on trypsin digestion efficiency for the mixture of two proteins was also tested. Tf was mixed with HSA or CAT in a 1:1 molar ratio, and the mixture was incubated with PS NPs for 1hr at 37 ^{0}C. The trypsin digestion and SDS-

PAGE followed the same procedure used for the single protein, which was mentioned previously. Compared to Tf itself incubated with PS NPs, HSA and CAT mixed with Tf alleviated the enhancement of PS NPs on the trypsin digestion efficiency of Tf, i.e. there was less decrease for the band intensity of intact Tf (Figure 6.1 c-d). One reason for this phenomenon could be that, less Tf proteins interacted with PS NPs, due to the competitive binding of HSA and CAT toward the limited surface of PS NPs. Moreover, CAT showed better protection of Tf from PS NPs, compared to HSA. This was also consistent with the previous result that the binding affinity of CAT to PS NPs was higher than that of HSA.

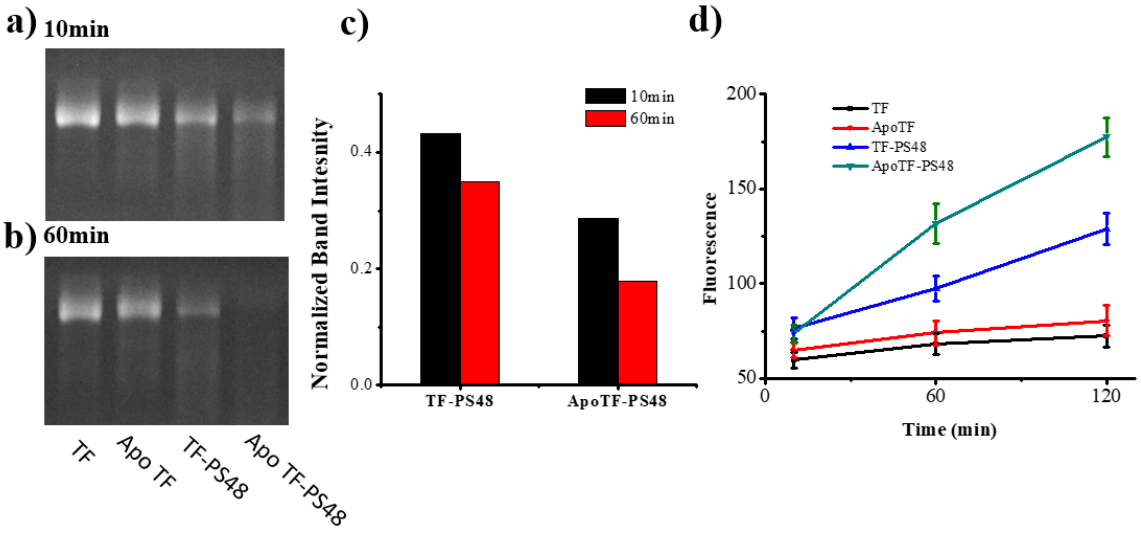


Figure 6.2. SDS-PAGE visualization of Tf and apo Tf, with or without PS48, after (a) 10 minutes or (b) 60 minutes digestion by trypsin. (c) The band intensity of TF or apo TF with PS48 was quantified by ImageJ and was divided by that of protein only to get a percentage. (d) Different digestion efficiencies were monitored using fluorescamine labeling. TF and apo TF, with or without PS48, were digested by trypsin for 10, 60, or 120 minutes. To terminate the reaction, SDS was added into the solution to a final concentration of 1%, and the mixture was then heated to boiling. Fluorescamine was added to 10 mM to label primary amines, and fluorescence was measured at 490 nm.

Limited proteolysis for Tf and apo-Tf with PS NPs. Although limited proteolysis had been able to differentiate distinct proteins (e.g. Tf and CAT) that interacted with PS NPs, there was still a concern regarding the precision of limited proteolysis – specifically the ability to was differentiate the small conformational changes of the same protein. To address this concern, two different forms of transferrin, holo- and apo-Tf, were used. Holo-Tf is saturated with iron ions, while apo-Tf is free of iron ions. To avoid the influence of excess iron ions on PS NPs, Tf with 80% iron bound was used as holo-Tf, and all Tf in this dissertation refer to it. The binding of iron ions to Tf would change the protein conformation, despite it being a relatively small change. As a result, the Tf and apo-Tf pair could be a good model to test the performance of this limited proteolysis on proteins with small conformational differences.

The SDS-PAGE results were shown in Figure 6.2 a-b. After 10 minutes of trypsin digestion, it was observed that the digestion efficiency for both Tf and apo-Tf could be enhanced by PS48 NPs, and there was more enhancement for apo-Tf with PS48 NPs, compared to Tf. After 60 minutes of trypsin digestion, the difference between Tf and apo-Tf with PS48 NPs were more obvious. Based on the quantification results by ImageJ, apo-Tf mixed with PS48 NPs only had 20% intact protein left after 60 minutes, but Tf mixed with PS48 NPs still had 30% intact protein left (Figure 6.2 c). The digestion process was also monitored using fluorescamine labeling. As expected, the native forms of Tf and apo-Tf were quite resistant to trypsin digestion, but Tf and apo-Tf bound to PS48 NPs showed much higher digestion efficiencies (figure 6.2 d). Collectively, these results indicate that

limited proteolysis could be sensitive enough to differentiate proteins with small conformational changes when proteins interact with PS NPs.

Molecular details of Tf/apo-Tf with PS NPs revealed by limited proteolysis. With the confidence that limited proteolysis could reveal the changes of Tf and CAT induced by PS NPs, the next step was to obtain more molecular details by identifying the peptides generated in the limited proteolysis by LC-MS/MS. At first, these peptides were separated from the other parts of the protein that were still adsorbed on the PS NPs, using Amicon ultra centrifugal filters with a molecular-weight-cut-off value (MWCO) of 30 kDa. Afterwards, an additional step of trypsin digestion was performed on these peptides to produce suitable lengths for LC-MS/MS. The spectral counts (SC) of each identified peptide was also used as the quantification of the cutting sites. The relative abundance (RA) of the cutting site, which could represent the digestion efficiency, was calculated by Equation 1. After comparing to the control, which was the protein alone, the RA changes for each cutting site could be used to map out the molecular details about the orientation and unfolding of proteins on PS NPs. Since the solvent accessibility of the cutting site was the determinant factor of the digestion efficiency, its change would be reflected as a RA change. As a result, the surface blockage caused by binding between the PS NPs and the proteins would result in a decrease of RA, while the unfolding could result in an increase of RA.

We studied the binding of Tf or apo-Tf to the PS NPs with limited proteolysis. As shown in Figure 6.3 a, there was little difference between the RA plots of Tf (black line) and apo-Tf (red line). This indicated that the small conformational change of the native

protein could not be mapped out by limited proteolysis. However, the differences of RA plots became more obvious after Tf/apo-Tf interacted with the PS48 NPs (Figure 6.3 a). Both Tf and apo-Tf showed dramatically decreased RA values in several regions (highlighted in red), and increased RA values in some other regions (highlighted in blue). There were also differences between Tf and apo-Tf on these changes of RA values induced by PS48 NPs, e.g. apo-Tf with PS48 NPs showed less decrease on RA values at the C-terminal, and the increase on RA values was more obvious for apo-Tf with PS48 NPs at the region of # 400-440 and # 550-600, compared to Tf with PS48 NPs. These results indicated the small conformational differences between Tf and apo-Tf would result in different digestion efficiencies at different regions when interacting with PS NPs, and these differences could be revealed by limited proteolysis.

To verify these results, another limited proteolysis by chymotrypsin was performed. Different from positively charged residues cut by trypsin, aromatic residues would be the cutting sites for chymotrypsin. Thus, more complementary information would be obtained by combining both trypsin and chymotrypsin limited digestion results,[40] despite the generation of different peptides by the chymotrypsin limited digestion. As shown in Figure 6.3 b, similar changes on RA plots induced by PS48 NPs, compared to the plots in figure 6.3 a, were observed: the region of # 100-140, # 460-500 and the C-terminus exhibited decreased RA values. To better visualize the areas influenced by PS48 NPs, the regions that showed RA changes were highlighted in the crystal structure of Tf: regions with RA values reduced by PS48 NPs were colored in red, while regions with RA values amplified by PS48 NPs were in colored in blue (Figure 6.3 c). Based on the previous assumption, the

reduction of RA values was induced by the protein surface blockage, indicating that these red regions should be close to the binding interface of Tf/apo-Tf with PS48 NPs. On the other hand, there might be unfolding happened at the blue regions, which could expose previously buried residues to the solvent, as demonstrated in the fluorescamine labeling. Moreover, it was interesting to see that these regions, including the blue and red ones, were located at the two far ends of the protein, suggesting that there can be more than one binding site for Tf/apo-Tf to interact with PS48 NPs.

These details of the binding sites and unfolding areas could provide new insights for the fate of ENMs coated with Tf, since Tf has been reported to functionalize ENMs to facilitate their targeting accumulation in prostate cancer cells.[41,42] The targeting mechanism was based on the recognition of Tf receptor (TfR) to Tf with iron ions (not apo-Tf); and the binding motifs of Tf to TfR include regions of # 68-75, # 142-145, # 166-167, #349-360, and # 641-651. Besides, the conformational changes of Tf after binding with iron ions, are primarily in the region of # 100-230, # 350-380, and # 490-510, and # 610-620 (Figure 6.3 d). Based on our results, several of these regions were assumed to be very close to the binding interface of Tf with PS NPs, as the RA values decreased with PS48 NPs, which suggested that the interactions with PS48 NPs could interfere with the binding of Tf by TfR. Thus, it is not very likely that PS48 NPs coated with Tf could show any targeting ability toward cancer cells with TfR overexpress.

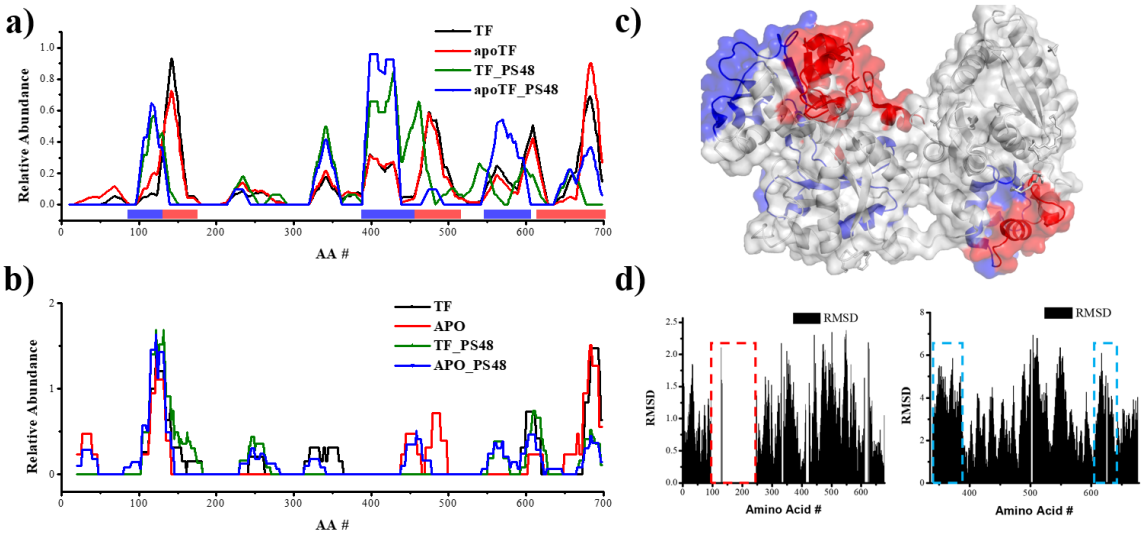


Figure 6.3. The relative abundance of cutting sites of (a) trypsin or (b) chymotrypsin along the sequence of Tf. These curves are the moving average of RA over 20 periods. (c) The regions with changed RA by PS48 NPs are shown on the crystal structure of Tf: the ones with higher RA for free protein are colored red, and the ones with higher RA for protein-PS NPs are colored blue. (d) RMSD plot for Tf with or without iron bound. A red box is used to highlight the difference for N-lobe, and blue boxes show the differences on C-lobe.

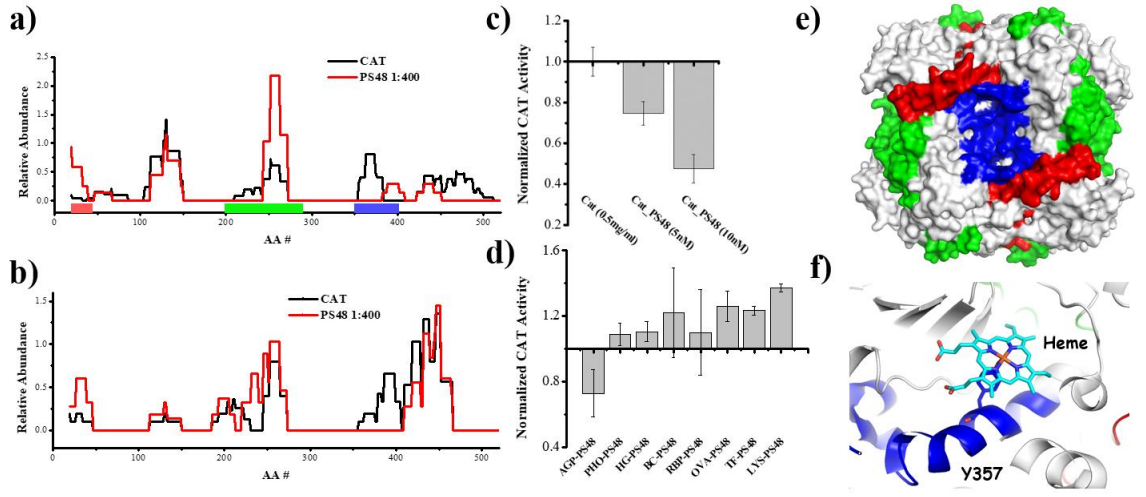


Figure 6.4. The relative abundance of cutting sites of (a) trypsin or (b) chymotrypsin along the sequence of catalase. These curves are the moving average of RA over 20 residues. (c) Normalized catalase enzymatic activity with different amount of PS48 NPs. The values are normalized to the signal of catalase itself. (d) Normalized enzymatic activity of catalase that mixed with another

protein and PS48 NPs. Normalization is based on the value of catalase with PS48 NPs. (e) The digestion regions altered by PS48 NPs are shown on crystal structure of Tf: the ones with higher RA for free protein are colored blue, and the ones with higher RA for protein-PS NPs are colored as red and green. (f) A closeup look of the heme group. Regions are colored in the same manner as in (e). The heme group is colored in cyan.

Molecular details of CAT with PS NPs revealed by limited proteolysis. As another positive hit obtained in previous fluorescamine labeling screening, catalase (CAT) would be explored for the molecular detail of its interaction with PS NPs in this section. The limited proteolysis by both trypsin and chymotrypsin followed the same procedure used for Tf in the previous section. As shown in Figure 6.4 a-b, the changes of RA plots induced by PS48 NPs were quite similar for both proteases. There was one region containing residues # 350-400 (highlighted in blue) that displayed a decreased RA value, indicating that this region should be close to the binding interface and was blocked by PS48 NPs. In contrast, two regions, including the N-terminus (highlighted in red) and residues 240-270 (highlighted in green), showed increased RA values, and it suggested that these two regions underwent unfolding after interacting with PS48 NPs. All the three regions were also highlighted in the crystal structure of CAT (Figure 6.4 e). CAT is a heme group containing protein, and it could also form a tetramer of four identical proteins in solution. Both the heme group and the tetramer structure are critical for the enzymatic activity of CAT. As one important enzyme and scavenger for reactive oxygen species (ROS), CAT could protect cells from oxidative damage, and it has also been conjugated to ENMs to inhibit ROS-mediated tumor metastasis.[43] It is imperative to conserve the enzymatic activity of CAT after it binds to ENMs; and the tetramer structure and the heme group should not be obstructed by ENMs. However, binding with PS48 NPs could interfere with

both. The N-terminus contributes to half of the interactions between subunits and is important for the formation of the tetramer (Figure 6.4 e). Since the RA values at the N-terminal increased when CAT interacted with PS48 NPs, it indicated that the N-terminal was no longer buried inside and PS48 NPs could disrupt the tetramer structure of CAT. Moreover, the residue (Tyr357) directly linked to the heme group is in the region that was assumed to be the binding site (Figure 6.4 f), indicating that the interaction between CAT and PS48 NPs could block the heme group.

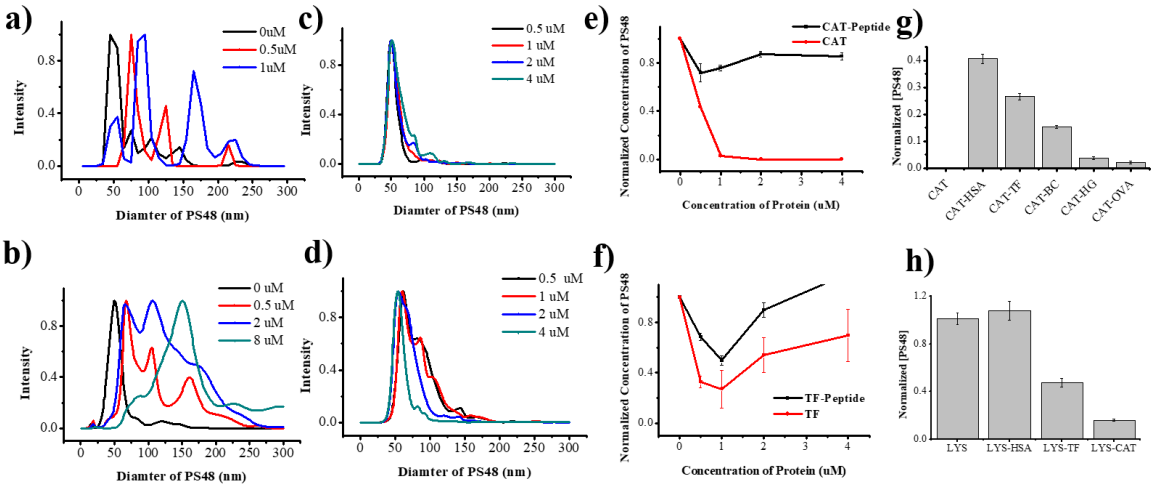


Figure 6.5. Size distributions of PS48 NPs incubated with different amounts of (a) catalase, (b) transferrin, (c) digested catalase, (d) digested transferrin. Intensity is normalized to the highest value of each run. Concentrations of PS48 NPs incubated with different amounts of (a) catalase and (b) transferrin. Normalization was based on the concentration of PS48 NPs alone. (g) Normalized concentration of PS NPs incubated with the mixture of catalase and another protein. (h) Normalized concentration of PS NPs incubated with the mixture of lysozyme and another protein.

The enzymatic activity of CAT was hypothesized to be compromised by PS48 NPs, and it was verified with the activity test (Figure 6.4 c). With 5 nM of PS48 NPs, the activity of CAT decreased by 20% compared to the control, whereas with 10 nM PS48 NPs, the

decrease was more obvious (by 50%). If another protein (e.g. lysozyme, Tf, etc.) was added to the mixture of CAT and PS48 NPs in the same molar ratio to CAT, the activity of CAT could be restored in different degrees (10-40%), due to the competitive binding of these proteins to the limited surface of PS48 NPs. In summary, these results suggested that PS48 NPs could induce binding to and the conformational change of catalase resulting in its functional loss.

Protein binding and NPs aggregation. To further investigate the interaction between proteins and PS nanoparticles, we utilized nanoparticle tracking analysis (NTA) to measure the size and concentration change of PS48 NPs after incubating with increasing protein concentrations. The results were very interesting. The concentration of PS48 NPs decreased and their sizes increased after adding CAT or Tf (figure 6.5 a-b, e-f), indicating the adsorption of CAT or Tf could induce the aggregation of PS48 NPs. Between these two, CAT showed a more severe aggregation with PS48 NPs, even at a lower concentration (1 nM). To verify it, the mixture of CTA with another protein (e.g. HSA, Tf, beta-casein, hemoglobin, and ovalbumin) was incubated with PS48 NPs. HSA and Tf could obviously alleviate the aggregation, due to the competitive binding to PS48 NPs (Figure 6.5 g). Moreover, CAT and Tf were also added to the mixture of lysozyme and PS48 NPs. Lysozyme itself didn't aggregate PS48 NPs. However, adding CAT and Tf could trigger the aggregation of PS48 NPs (Figure 6.5 h). Nevertheless, there would not be aggregation of PS48 NPs incubated with peptides of CAT or Tf (Figure 6.5 c-d, e-d). These data suggest that the protein 3D structure was also important to induce the aggregation of PS48 NPs. As shown in previous sections, there can be multiple binding sites on either Tf or CAT to bind

to PS48 NPs, and these binding sites could bridge PS48 NPs to form aggregates, especially at lower protein concentrations.

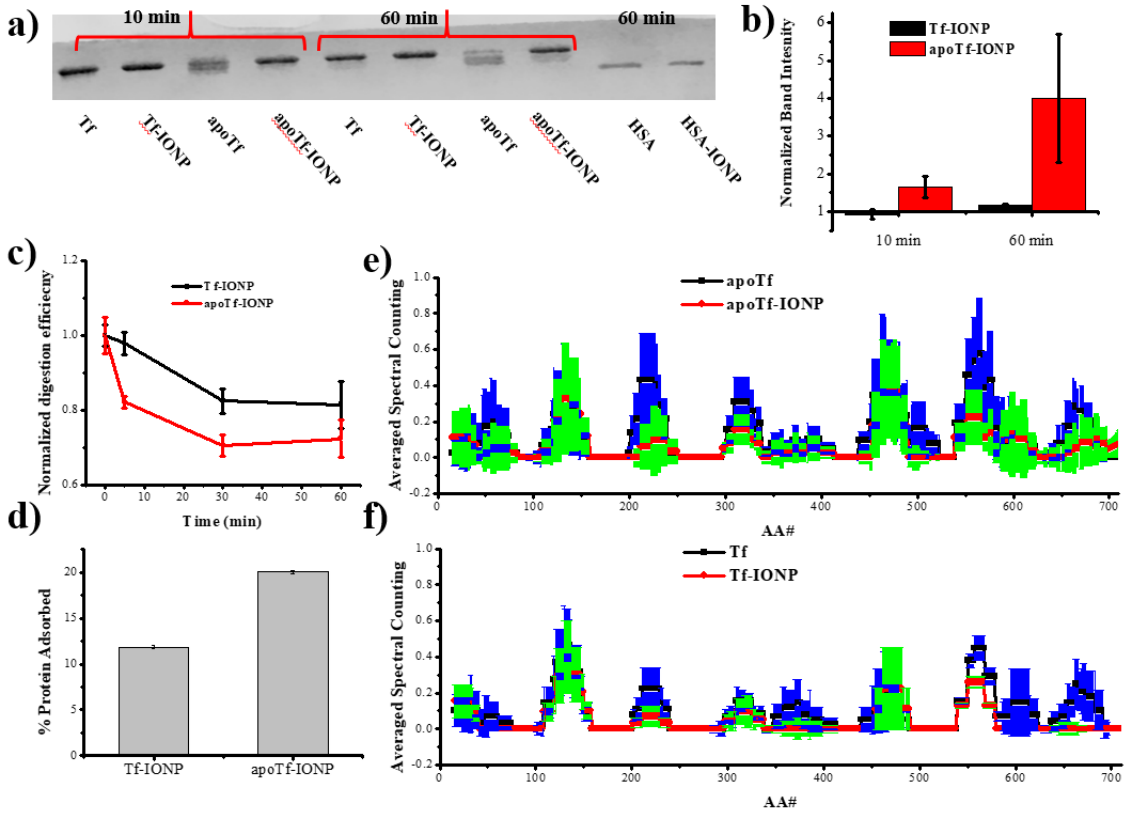


Figure 6.6. (a) SDS-PAGE visualization of trypsin digestion of Tf and apo-Tf with and without iron oxide NPs, after 10 or 60 minutes. HSA with iron oxide NPs is used as the control. Trypsin is added in a 1:50 mass ratio. Incubation was under 60 ^{0}C. (b) Intensity of intact protein band is quantified by ImageJ, and the intensity of protein/NPs mixture is normalized to the value of protein itself. (c) Fluorescamine labeling to monitor the digestion process. 1:50 trypsin is added, and incubation iss at 60 ^{0}C, for 0. 5, 30, and 60 minutes. Fluorescamine was added to the final concentration of 1 mM, to quantify the number of N-terminus generated by digestion. Fluorescence signal is normalized to the value at 0 min. (d) The percentage of protein adsorbed on iron oxide NPs after 1 hr incubation. Protein in 0.2 mg/ml is mixed with 0.1 mg/ml of NPs. Centrifugation at 15,000 ×g for 10 minutes is employed to pellet NPs, and the proteins adsorbed on NPs are quantified by BCA assay. Averaged SC for each cutting site by a 30 minutes trypsin digestion is

plotted for (e) transferrin and (f) apo-transferrin. An average is calculated for every 20 closest residues. Error bars are based on four LC-MS/MS runs.

Limited proteolysis for Tf/apoTf with iron oxide NP. All previous studies were on PS NPs. To verify the robustness of our method, another type of NP, iron oxide NP with 10 nm size, was tested in this section. In the previous fluorescamine screening, iron oxide NP was found to diminish the fluorescence signal for Tf/apo-Tf. This implied that the binding of iron oxide NP could block the surface of Tf/apo-Tf but would not induce obvious unfolding, which was different from the PS NPs. The protein surface blockage was not advantageous for this limited proteolysis study, on account that very small changes would be observed. To enlarge the difference, heating pressure was applied on proteins to induce unfolding and accelerate digestion rate. At 60 ^{0}C, Tf/apo-Tf would not aggregate, but they could be digested much faster by trypsin, as visualized by SDS-PAGE (Figure 6.6 a). The intensity of the intact protein was also quantified by ImageJ, and the result showed that the impact of iron oxide NP on apo-Tf was more obvious than that on Tf (Figure 6.6 b). The decreased digestion efficiency of Tf/apo-Tf induced by iron oxide NP was not caused by the activity loss of trypsin, since no change was observed in HSA incubated with iron oxide NP (Figure 6.6 a). Furthermore, the digestion process was monitored by fluorescamine labeling, which could label and quantify the number of N-terminus generated in the digestion process (Figure 6.6 c). The different decreases in the digestion efficiency of Tf/apoTf induced by iron oxide NP was observed again. Moreover, the amount of Tf/apo-Tf strongly adsorbed on the iron oxide NP was quantified by BCA kit, after NP-protein complex was centrifuged down. It was interesting to see that more apo-Tf could be adsorbed on iron oxide NP compared to Tf (Figure 6.6 d). This indicated that the

binding affinities of Tf/apo-Tf toward iron oxide NP were different, which could also be one of the reasons for the different impacts of iron oxide NP on the digestion efficiency of these two proteins.

To acquire the molecular details about the binding behaviors of Tf/apo-Tf to iron oxide NP, LC-MS/MS was used to identify the cutting sites generated in the limited proteolysis step. Since the complete digestion of the protein occurs in 30 minutes at 60 ^{0}C (Figure 6.6 c), the time allotted for the limited digestion was 10 minutes. Amico filer with 30 kDa cutoffs was used to extract peptides generated in the limited digestion. For better visualization, the SC for each cutting site was smoothed by calculating the averaged SC (ACS) of the twenty closest residues. The standard deviation of four repeats was used for the error bar. From the results shown in Figure 6.6 e-f, it was clear to see the decrease of ASC for both Tf and apo-Tf induced by iron oxide NP, Regions with decreased ASC induced by iron oxide NP included residues 200-240, as well as residues 550-580. These sections were different from the regions identified as the binding sties of Tf/apo-Tf to PS48 NP, suggesting that the binding mechanisms of Tf/apo-Tf to these two NPs were different.

6.4: Conclusion

Limited proteolysis coupled with LC-MS/MS has proven to be a powerful tool to explore the molecular details regarding the orientation and conformational changes of proteins adsorbed on ENMs. Two different outcomes induced by interactions with ENMs, including increased and decreased digestion efficiency, could be revealed in the sub-molecular level by this method. The regions with decreased digestion efficiency are more likely to be in the binding interface of the protein with ENMs, while an increased digestion

efficiency usually signifies that the region has undergone unfolding. Although the precision of this method is not enough to unveil the arrangement of proteins on ENMs on the atomic level, one of the significances of this method is that it provides new insights on the function of proteins on ENMs, by evaluating the orientation and conformation of protein. Moreover, it is interesting to identify multiple binding sites on one protein, which could result in the aggregation of NPs through a bridging effect. These results can provide useful information to facilitate the design of ENMs, to increase the stability, achieve active targeting, and mitigate toxicity.

References

(1) Björnmalm, M.; Thurecht, K. J.; Michael, M.; Scott, A. M.; Caruso, F. Bridging Bio–Nano Science and Cancer Nanomedicine. *ACS nano* **2017**, *11*, 9594-9613.

(2) Jiang, W.; Von Roemeling, C. A.; Chen, Y.; Qie, Y.; Liu, X.; Chen, J.; Kim, B. Y. Designing nanomedicine for immuno-oncology. *Nature Biomedical Engineering* **2017**, *1*, 0029.

(3) von Roemeling, C.; Jiang, W.; Chan, C. K.; Weissman, I. L.; Kim, B. Y. Breaking down the barriers to precision cancer nanomedicine. *Trends in biotechnology* **2017**, *35*, 159-171.

(4) Shi, J.; Kantoff, P. W.; Wooster, R.; Farokhzad, O. C. Cancer nanomedicine: progress, challenges and opportunities. *Nature Reviews Cancer* **2017**, *17*, 20.

(5) Senders, M. L.; Fayad, Z. A.; Reiner, T.; Mulder, W. J.; Pérez-Medina, C.; Future Medicine, 2017.

(6) Thomas, A.; Teicher, B. A.; Hassan, R. Antibody–drug conjugates for cancer therapy. *The Lancet Oncology* **2016**, *17*, e254-e262.

(7) Chudasama, V.; Maruani, A.; Caddick, S. Recent advances in the construction of antibody–drug conjugates. *Nature chemistry* **2016**, *8*, 114.

(8) Bertrand, N.; Wu, J.; Xu, X.; Kamaly, N.; Farokhzad, O. C. Cancer nanotechnology: the impact of passive and active targeting in the era of modern cancer biology. *Advanced drug delivery reviews* **2014**, *66*, 2-25.

(9) Wang, S.; Huang, P.; Chen, X. Stimuli-responsive programmed specific targeting in nanomedicine. *Acs Nano* **2016**, *10*, 2991-2994.

(10) Chan, A. C.; Carter, P. J. Therapeutic antibodies for autoimmunity and inflammation. *Nature Reviews Immunology* **2010**, *10*, 301.

(11) Besse, B.; Charrier, M.; Lapierre, V.; Dansin, E.; Lantz, O.; Planchard, D.; Le Chevalier, T.; Livartoski, A.; Barlesi, F.; Laplanche, A. Dendritic cell-derived exosomes as maintenance immunotherapy after first line chemotherapy in NSCLC. *Oncoimmunology* **2016**, *5*, e1071008.

(12) Utsugi-Kobukai, S.; Fujimaki, H.; Hotta, C.; Nakazawa, M.; Minami, M. MHC class I-mediated exogenous antigen presentation by exosomes secreted from immature and mature bone marrow derived dendritic cells. *Immunology letters* **2003**, *89*, 125-131.

(13) Wilhelm, S.; Tavares, A. J.; Dai, Q.; Ohta, S.; Audet, J.; Dvorak, H. F.; Chan, W. C. Analysis of nanoparticle delivery to tumours. *Nature reviews materials* **2016**, *1*, 16014.

(14) Schöttler, S.; Becker, G.; Winzen, S.; Steinbach, T.; Mohr, K.; Landfester, K.; Mailänder, V.; Wurm, F. R. Protein adsorption is required for stealth effect of poly (ethylene glycol)-and poly (phosphoester)-coated nanocarriers. *Nature nanotechnology* **2016**, *11*, 372.

(15) Herda, L. M.; Hristov, D. R.; Lo Giudice, M. C.; Polo, E.; Dawson, K. A. Mapping of molecular structure of the nanoscale surface in bionanoparticles. *Journal of the American Chemical Society* **2016**, *139*, 111-114.

(16) Mortimer, G. M.; Butcher, N. J.; Musumeci, A. W.; Deng, Z. J.; Martin, D. J.; Minchin, R. F. Cryptic epitopes of albumin determine mononuclear phagocyte system clearance of nanomaterials. *ACS nano* **2014**, *8*, 3357-3366.

(17) Gebauer, J. S.; Malissek, M.; Simon, S.; Knauer, S. K.; Maskos, M.; Stauber, R. H.; Peukert, W.; Treuel, L. Impact of the nanoparticle–protein corona on colloidal stability and protein structure. *Langmuir* **2012**, *28*, 9673-9679.

(18) Lynch, I.; Dawson, K. A.; Linse, S. Detecting cryptic epitopes created by nanoparticles. *Sci. Stke* **2006**, *2006*, pe14-pe14.

(19) Lara, S.; Alnasser, F.; Polo, E.; Garry, D.; Lo Giudice, M. C.; Hristov, D. R.; Rocks, L.; Salvati, A.; Yan, Y.; Dawson, K. A. Identification of receptor binding to the biomolecular corona of nanoparticles. *ACS nano* **2017**, *11*, 1884-1893.

(20) Walczyk, D.; Bombelli, F. B.; Monopoli, M. P.; Lynch, I.; Dawson, K. A. What the cell "sees" in bionanoscience. *J Am Chem Soc* **2010**, *132*, 5761-5768.

(21) Van Hong Nguyen, B.-J. L. Protein corona: a new approach for nanomedicine design. *International journal of nanomedicine* **2017**, *12*, 3137.

(22) Corbo, C.; Molinaro, R.; Tabatabaei, M.; Farokhzad, O. C.; Mahmoudi, M. Personalized protein corona on nanoparticles and its clinical implications. *Biomaterials science* **2017**, *5*, 378-387.

(23) Lundqvist, M.; Stigler, J.; Elia, G.; Lynch, I.; Cedervall, T.; Dawson, K. A. Nanoparticle size and surface properties determine the protein corona with possible implications for biological impacts. *Proceedings of the National Academy of Sciences* **2008**.

(24) Lundqvist, M.; Stigler, J.; Cedervall, T.; Berggård, T.; Flanagan, M. B.; Lynch, I.; Elia, G.; Dawson, K. The evolution of the protein corona around nanoparticles: a test study. *ACS nano* **2011**, *5*, 7503-7509.

(25) Kelly, P. M.; Åberg, C.; Polo, E.; O'connell, A.; Cookman, J.; Fallon, J.; Krpetić, Ž.; Dawson, K. A. Mapping protein binding sites on the biomolecular corona of nanoparticles. *Nature nanotechnology* **2015**, *10*, 472.

(26) Assfalg, M.; Ragona, L.; Pagano, K.; D'Onofrio, M.; Zanzoni, S.; Tomaselli, S.; Molinari, H. The study of transient protein-nanoparticle interactions by solution NMR spectroscopy. *Biochimica et biophysica acta* **2016**, *1864*, 102-114.

(27) Xu, Y.; Matthews, S. TROSY NMR spectroscopy of large soluble proteins. *Topics in current chemistry* **2013**, *335*, 97-119.

(28) Frueh, D. P.; Goodrich, A. C.; Mishra, S. H.; Nichols, S. R. NMR methods for structural studies of large monomeric and multimeric proteins. *Current opinion in structural biology* **2013**, *23*, 734-739.

(29) Shrivastava, S.; Nuffer, J. H.; Siegel, R. W.; Dordick, J. S. Position-specific chemical modification and quantitative proteomics disclose protein orientation adsorbed on silica nanoparticles. *Nano letters* **2012**, *12*, 1583-1587.

(30) Zeng, S.; Yu-ming, M. H.; Chia-en, A. C.; Zhong, W. Protein binding for detection of small changes on a nanoparticle surface. *Analyst* **2014**, *139*, 1364-1371.

(31) Srinivasu, B. Y.; Bose, B.; Mitra, G.; Kurpad, A. V.; Mandal, A. K. Adsorption induced changes of human hemoglobin on ferric pyrophosphate nanoparticle surface probed by isotope exchange mass spectrometry: An implication on structure–function correlation. *Langmuir* **2017**, *33*, 8032-8042.

(32) Jemmerson, R.; Paterson, Y. Mapping epitopes on a protein antigen by the proteolysis of antigen-antibody complexes. *Science* **1986**, *232*, 1001-1004.

(33) Yang, J. A.; Lin, W.; Woods, W. S.; George, J. M.; Murphy, C. J. alpha-Synuclein's adsorption, conformation, and orientation on cationic gold nanoparticle surfaces seeds global conformation change. *The journal of physical chemistry. B* **2014**, *118*, 3559-3571.

(34) Laurie, A. T.; Jackson, R. M. Q-SiteFinder: an energy-based method for the prediction of protein–ligand binding sites. *Bioinformatics* **2005**, *21*, 1908-1916.

(35) Keskin, O.; Gursoy, A.; Ma, B.; Nussinov, R. Principles of protein− protein interactions: What are the preferred ways for proteins to interact? *Chemical reviews* **2008**, *108*, 1225-1244.

(36) Lomenick, B.; Hao, R.; Jonai, N.; Chin, R. M.; Aghajan, M.; Warburton, S.; Wang, J.; Wu, R. P.; Gomez, F.; Loo, J. A. Target identification using drug affinity responsive target stability (DARTS). *Proceedings of the National Academy of Sciences* **2009**, pnas. 0910040106.

(37) Receveur-Brechot, V.; Bourhis, J. M.; Uversky, V. N.; Canard, B.; Longhi, S. Assessing protein disorder and induced folding. *Proteins* **2006**, *62*, 24-45.

(38) Duan, Y.; Liu, Y.; Shen, W.; Zhong, W. Fluorescamine Labeling for Assessment of Protein Conformational Change and Binding Affinity in Protein–Nanoparticle Interaction. *Analytical chemistry* **2017**, *89*, 12160-12167.

(39) Ashby, J.; Duan, Y.; Ligans, E.; Tamsi, M.; Zhong, W. High-Throughput Profiling of Nanoparticle–Protein Interactions by Fluorescamine Labeling. *Analytical chemistry* **2015**, *87*, 2213-2219.

(40) Golghalyani, V.; Neupartl, M.; Wittig, I.; Bahr, U.; Karas, M. ArgC-Like Digestion: Complementary or Alternative to Tryptic Digestion? *Journal of proteome research* **2017**, *16*, 978-987.

(41) Yhee, J. Y.; Lee, S. J.; Lee, S.; Song, S.; Min, H. S.; Kang, S. W.; Son, S.; Jeong, S. Y.; Kwon, I. C.; Kim, S. H.; Kim, K. Tumor-targeting transferrin nanoparticles

for systemic polymerized siRNA delivery in tumor-bearing mice. *Bioconjugate chemistry* **2013**, *24*, 1850-1860.

(42) Singh, R.; Norret, M.; House, M. J.; Galabura, Y.; Bradshaw, M.; Ho, D.; Woodward, R. C.; St Pierre, T. G.; Luzinov, I.; Smith, N. M.; Lim, L. Y.; Iyer, K. S. Dose-Dependent Therapeutic Distinction between Active and Passive Targeting Revealed Using Transferrin-Coated PGMA Nanoparticles. *Small* **2016**, *12*, 351-359.

(43) Singhal, A.; Morris, V. B.; Labhasetwar, V.; Ghorpade, A. Nanoparticle-mediated catalase delivery protects human neurons from oxidative stress. *Cell death & disease* **2013**, *4*, e903.

CHAPTER 7: Conclusion and Future Outlook

This dissertation has focused on the method development for two problems in the study of protein corona: (1) the prediction of protein corona compositions, and (2) the molecular details of each protein in the corona. For the first problem, a high throughput screening method based on fluorescamine labeling has been demonstrated useful to characterize the interaction between NMs and proteins. Using the obtained fluorescence profiles as descriptors for NMs, a robust SAR model could be built to predict the serum protein corona composition. The data required for this prediction model (e.g. fluorescamine labeling profile) is more accessible than the protein corona compositions identified by traditional methods (e.g. LC-MS/MS). Thus, this prediction model could be applied to quickly and routinely evaluate the protein corona composition for new NMs, facilitating their design and optimization. As for the second problem, an additional method based on limited proteolysis has also proven able to map out the molecular details (e.g. orientation and unfolding) of the protein in the corona of NMs. There is great significance to evaluate the molecular details of proteins in corona. Since protein corona formed around NMs will provide a new biological identify for NMs, it has been utilized for the rational design of ENMs to increase the targeting ability of NMs and create new functions for NMs.[1-3] One importance factor constraining the success of this strategy is the lack of tools to rapidly evaluate the orientation and unfolding of the protein in corona.[4,5] The limited proteolysis method discussed in this thesis can make up for this deficiency and contribute to the tool repertoire for the protein corona community.

Despite the robustness of these methods verified in this thesis, there is still room to improve their performances. For the prediction model of protein corona compositions, more descriptors for NMs can be incorporated to enhance its accuracy and precision. Because lipids, nucleic acids, and metabolites are also the importance and abundant components in biological matrix,[1,6] they will compete with proteins for the limited surface of NMs and interfere with the protein corona compositions. Consequently, a quick method to characterize these influences would provide more comprehensive descriptors for NMs, and the performance of the prediction model will benefit from the new descriptors.

For the limited proteolysis method, the dynamic exchange of proteins in the corona may limit its performance, due to that the generated peptides would replace the protein in a very short time.[7] To solve this issue, a fixation method (e.g. formaldehyde crosslinking)[8] can be used to conserve the conformation and orientation of proteins in corona; and the resulted rigid protein structure will not be altered by the fast dissociation and greatly improve the performance of the limited digestion.

There are also more potential applications of our methods. NMs are commonly used in food and pharmaceutical industries, and many NMs can also be generated in daily life, e.g. cooking.[9] Thus, there is a big chance for NMs to enter human bodies through the lung, skin, and other mucosa tissues. The biological fluids (e.g. lachrymal fluid and bronchoalveolar lavage fluid)[10,11] in these organs could be explored and tested, to reveal the impacts of NMs on environmental health. The protein corona of NMs formed in tear or lung fluid is very important for the biological outcomes of NMs in human bodies.[12] As a result, the quick prediction of the protein corona compositions could be helpful to evaluate

the nanotoxicity of NMs. Moreover, the composition and concentration of proteins in different matrices could be very heterogenous.[13] It will be challenging but interesting to build a robust SAR model to predict the protein corona of ENMs in different matrices.

References

(1) Walczyk, D.; Bombelli, F. B.; Monopoli, M. P.; Lynch, I.; Dawson, K. A. What the cell "sees" in bionanoscience. *J Am Chem Soc* **2010**, *132*, 5761-5768.

(2) Hu, W.; Peng, C.; Lv, M.; Li, X.; Zhang, Y.; Chen, N.; Fan, C.; Huang, Q. Protein corona-mediated mitigation of cytotoxicity of graphene oxide. *ACS nano* **2011**, *5*, 3693-3700.

(3) Van Hong Nguyen, B.-J. L. Protein corona: a new approach for nanomedicine design. *International journal of nanomedicine* **2017**, *12*, 3137.

(4) Kelly, P. M.; Åberg, C.; Polo, E.; O'connell, A.; Cookman, J.; Fallon, J.; Krpetić, Ž.; Dawson, K. A. Mapping protein binding sites on the biomolecular corona of nanoparticles. *Nature nanotechnology* **2015**, *10*, 472.

(5) Lara, S.; Perez-Potti, A.; Herda, L. M.; Adumeau, L.; Dawson, K. A.; Yan, Y. Differential Recognition of Nanoparticle Protein Corona and Modified Low-Density Lipoprotein by Macrophage Receptor with Collagenous Structure. *ACS nano* **2018**, *12*, 4930-4937.

(6) Monopoli, M. P.; Åberg, C.; Salvati, A.; Dawson, K. A. Biomolecular coronas provide the biological identity of nanosized materials. *Nature nanotechnology* **2012**, *7*, 779.

(7) Tenzer, S.; Docter, D.; Kuharev, J.; Musyanovych, A.; Fetz, V.; Hecht, R.; Schlenk, F.; Fischer, D.; Kiouptsi, K.; Reinhardt, C. Rapid formation of plasma protein corona critically affects nanoparticle pathophysiology. *Nature nanotechnology* **2013**, *8*, 772.

(8) Mohammed, H.; Taylor, C.; Brown, G. D.; Papachristou, E. K.; Carroll, J. S.; D'santos, C. S. Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. *Nature protocols* **2016**, *11*, 316.

(9) Fischer, H. C.; Chan, W. C. Nanotoxicity: the growing need for in vivo study. *Current opinion in biotechnology* **2007**, *18*, 565-571.

(10) Bhatta, R.; Chandasana, H.; Chhonker, Y.; Rathi, C.; Kumar, D.; Mitra, K.; Shukla, P. Mucoadhesive nanoparticles for prolonged ocular delivery of natamycin: in vitro and pharmacokinetics studies. *International journal of pharmaceutics* **2012**, *432*, 105-112.

(11) Cho, W. S.; Duffin, R.; Poland, C. A.; Howie, S. E. M.; Macnee, W.; Bradley, M.; Megson, I. L.; Donaldson, K. Metal Oxide Nanoparticles Induce Unique Inflammatory Footprints in the Lung: Important Implications for Nanoparticle Testing. *Environ Health Perspect* **2010**, *118*, 1699-1706.

(12) Sapsford, K. E.; Algar, W. R.; Berti, L.; Gemmill, K. B.; Casey, B. J.; Oh, E.; Stewart, M. H.; Medintz, I. L. Functionalizing nanoparticles with biological molecules: developing chemistries that facilitate nanotechnology. *Chemical reviews* **2013**, *113*, 1904-2074.

(13) Tejler, L.; Grubb, A. O. A complex forming protein heterogeneous in charge and present in human plasma, urine, and cerebrospinal fluid. *Biochimica et biophysica acta* **1976**, *439*, 82-94.