Insights into the Roles of DOPA – Vanadium Complexes in the Formation and Function of the Blue Mussel Byssus Cuticle

Mihai Mircea Mesko

Department of Chemistry

McGill University

Montreal, Quebec, Canada

H3A 0B8

April 2021

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

Abstract

Mytilus edulis is a species of marine mussel that spends its life anchored to substrates in rocky seashore habitats worldwide via its byssus (or beard), a collection of adhesive protein-based fibers. The surface of byssal threads are coated in a thin protein layer called the cuticle, which has hardness and stiffness values rivaling certain engineering epoxies, while also being highly extensible, biodegradable and possessing self-healing behavior. The principal protein in the cuticle (mfp-1) contains a post translational modification of tyrosine, known as 3,4-dihydroxyphenylalanine (DOPA). Mfp-1 contributes to the remarkable properties of the cuticle via formation of coordination bonds between DOPA residues and transition metal ions accumulated from seawater. Mfp-1 has become a source inspiration for researchers developing bio-inspired polymeric materials. However, a more thorough understanding of the byssal cuticle structure-property relationships and fabrication process is required. In particular, significant open questions concern: 1) the role of proteins other than mfp-1 in the cuticle formation and function. 2) The role of different metal ions (Fe and V) in cuticle properties.

In this thesis, the structure-property relationships at play in the function and formation of the cuticle was probed taking a two-pronged approach to address these open questions. In the first part of the thesis, tip-enhanced Raman spectroscopy (TERS) was harnessed to probe the composition of cuticle sections at the nanoscale. I applied different statistical analyses in order to extract meaningful data. It was determined, however, that the highly complex composition of the cuticle and certain limitations of the TERS technique made this challenging.

In the second part, I recombinantly expressed a truncated form of mfp-1 to probe the metal-binding properties of the cuticle *in vitro*. Experimentation found increased mechanical cohesion forces in the presence of vanadium ions versus iron ions. To better understand the molecular cause of this mechanical difference, I utilized a range of spectroscopic techniques including UV-vis, Raman and EPR to characterize the nature of the metal coordination complexes under different conditions, in comparison with simple catechols. These studies present strong evidence that V coordination provides a functional advantage by creating stronger, more stable cross-links. These findings provide important insights into our understanding of this remarkable biological material, relevant to the rapidly growing field of mussel-inspired metallopolymers.

Résumé

Mytilus edulis est une espèce de moule marine qui passe sa vie ancrée à des substrats dans des habitats côtiers rocheux du monde entier avec son byssus (ou barbe), un ensemble de fibres adhésives à base de protéines. La surface des fibres est recouverte d'une fine couche de protéines appelée la cuticule, qui a une dureté et rigidité rivalisant certains époxydes d'ingénierie, tout en étant remarquablement extensible, biodégradable et possédant un comportement d'auto-guérison. La principale protéine de la cuticule (mfp-1) contient une modification post-traductionnelle de la tyrosine, connue sous le nom de 3,4-dihydroxyphénylalanine (DOPA). Mfp-1 contribue aux propriétés remarquables de la cuticule via la formation de liaisons de coordination entre les résidus DOPA et les ions de métaux de transition accumulées de l'eau de mer. Mfp-1 est devenu une source d'inspiration pour les chercheurs développant des matériaux polymères bio-inspirés. Cependant, une compréhension plus approfondie des relations structure-propriété de la cuticule byssale et des processus de fabrication est nécessaire. En particulier, des questions ouvertes concernent : 1) le rôle des protéines autres que la mfp-1 dans la formation et la fonction de la cuticule. 2) Le rôle des différents ions métalliques (Fe et V) sur les propriétés de la cuticule.

Dans cette thèse, les relations structure-propriété en jeu dans la fonction et la formation de la cuticule ont été explorées par deux approches. Dans la première partie de la thèse, la « Tip-Enhanced Raman Spectroscopy » (TERS) a été exploitée pour sonder la composition des coupes de cuticule à l'échelle nanométrique. En utilisant des analyses statistiques différentes, il a été déterminé que la composition très complexe de la cuticule et certaines limites de la technique TERS rendent cette tâche difficile.

Dans la deuxième partie, j'ai exprimé de manière recombinante une forme tronquée de mfp-1 pour sonder les propriétés des liaisons aux métaux de la cuticule *in vitro*. L'expérimentation a révélé des forces de cohésion augmentées en présence d'ions vanadium par rapport aux ions fer. Pour mieux comprendre la cause moléculaire de cette différence mécanique, j'ai utilisé une gamme de techniques spectroscopiques comprenant l'UV-VIS, le Raman et l'EPR pour caractériser la nature des complexes de coordination des métaux sous différentes conditions, en comparaison avec de simples catéchols. Ces études présentent des

preuves rigoureuses que la coordination avec V fournit un avantage fonctionnel en créant des liens croisés plus forts et plus stables. Ces découvertes fournissent des informations importantes sur notre compréhension de ce matériel biologique remarquable, pertinent dans le domaine des métallopolymères inspirés des moules.

Acknowledgements

I would like to begin by thanking my supervisor Dr. Matt Harrington, who kindly accepted me into his lab at a time when I needed one. He went above and beyond, patiently teaching me, inspiring me to feed my curiosity and encouraging me to follow my dreams. I never envisioned myself researching marine mussels and loving every moment of it, yet here we are at the end of this chapter, and my heart weighs heavy having to turn the page to the next. I will forever be grateful for everything he has done for me.

Thanks are also in order for a host of other people. Thank you to my parents, Silvana and Dezso, for all the sacrifices they have made for me, for supporting me unconditionally and trying their best to understand the weird science I do. Thank you to the entire Harrington Lab (Tobias Priemel, Mimi Simmons, Max Renner-Rao, Mostafa Rammal, Lucia Youssef, Stephen George and Franzi Jehle), for the science advice, office distractions, lab shenanigans, bar outings, and group meeting snacks. Thank you to all my friends (including, but not limited to Jake Pierscianowski, Jath Asohan, Michelle Ting, Mimi Simmons, Yiram Kim, Shayne Gracious, Alex Hebert, Cameron Reid, Tony Jin, Kevin Uggowitzer, Kaitlin Nuechterlein, Matt Renaud, Kelly McManus), who saw me through my highs and lows, and made these last few years an experience I hope to never forget. I will still get to see most of you in Montreal, but for the ones that are moving, I will miss you dearly!

Lastly, I must thank the lab groups, staff members and organizations without which this thesis would not have been possible. Thank you to Thibodeaux, Auclair, Mittermaier and Sleiman Groups for welcoming me into your labs to use your equipment. Thank you to Li Xiang, who performed my SFA experiments at University of Alberta as my trip to Edmonton was cancelled due to COVID-19. Thank you to Dr. Alex Wahba for all the help with my mass spectrometry of my proteins. Thank you to Dr. Robin Stein, who introduced me to EPR and helped me acquire my spectra. Thank you to NSERC and McGill University for seeing merit in my research abilities and supplying funding throughout this program.

Table of Contents

List of Figures

Contribution of Authors

This thesis is written as a singular chapter.

SFA data was collected and analyzed by Li Xiang (Department of Chemical and Materials Engineering, University of Alberta).

TERS samples were prepared by Clemens Schmitt (Biomaterials Group, Max Planck Institute of Colloids and Interfaces), and TERS data was collected by Tanja Deckert-Gaudig (Faculty of Chemistry and Earth Sciences, University of Jena).

4-arm PEG-DOPA was prepared by Mostafa Rammal (Department of Chemistry, McGill University).

All remaining sample preparation, experimentation and data analysis was conducted solely by the Author.

1. Introduction

Biological organisms have evolved vast arrays of materials to meet the challenge of survival within their ecological niches. Studying and mimicking the structure-function relationships and self-assembly processes of these materials allows humans to achieve novel properties in our own materials[1]; this brings previously unattainable combinations of physical properties, improved sustainability and increased economy in the manufacturing process within reach simultaneously. The drivers of material advances in human society are very often linked to the specific pressures that threaten our existence. Currently, a great deal of research into biological materials hopes to provide some inspiration for handling one of humanity's greatest obstacles to date: finding equally high performing, yet environmentally friendly alternatives to single-use plastics that are piling up across the globe. In this thesis, I focus on improving our understanding of the hard, yet extensible cuticle coating of the byssus, a collection of proteinaceous threads utilized by mussels to anchor themselves to inhospitable aquatic environments the world over.

1.1. The Evolution of Humans and Their Materials

The tools that humans have developed throughout their history (and pre-history) have made an indelible impact on our civilizations and our planet. Inseparable from the development of tools and devices ranging from stone axes and iron swords to plastic bottles and smart phones is the development of new materials and novel ways of processing and fabricating them. Currently, we find ourselves at another junction in which the fate of the planet depends on our ability to adapt our means of materials production and consumption to be more sustainable.

Humans have a long history of tool use and with the materials processing developments that have enabled the production of new tools. About 7 million years ago, the Hominiae Subfamily (often referred to as "African Hominids" or "African Apes") experienced a divergence from a common ancestor.[2] One half of this split eventually resolved as the Gorilla genus, which still exists today. The other underwent several more divisions, yielding genera such as Pan, comprising of chimpanzees and bonobos, and eventually Homo, of which the last extant species

is *Homo sapiens.* These divergences that eventually culminated in *Homo sapiens* brought about not only distinct physiological changes, but also changes in societal structure, and the creation of sophisticated tools and materials (**Figure 1**).

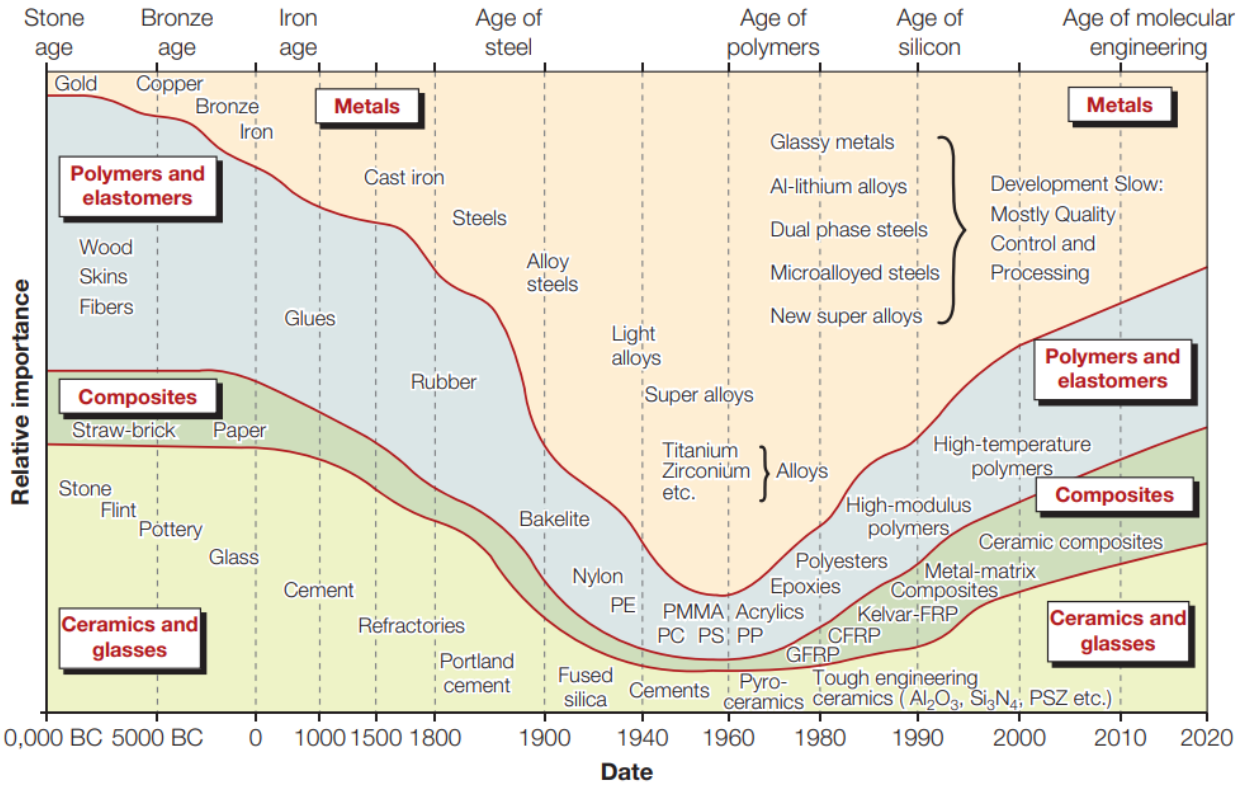

Figure 1: The evolution of engineering materials with time. Relative importance is derived from historical data, and since 1960, the teaching hours allocated to each material class in U.K. and U.S. universities. It is important to note the timescale is not linear. Reproduced with permission from Materials Selection in Mechanical Design (4th Ed.) by M. Ashby.[3]

Tool-making is not a trait exclusive to humans, although the opposite was widely believed to be the criterion separating humans from other animals until Jane Goodall's observations of chimpanzees using long grass stalks as a tool for collecting termites from their mounds.[4] Since that initial observation, wild chimpanzees have been observed fashioning "spears" to hunt lesser bushbabies,[5] while orangutans frequently utilize specialized sticks to fish termites out of tree hollows.[6] Beyond primates, tools are frequently employed by birds,[7] large mammals such as elephants,[7] and cetaceans.[8] Despite the widespread usage of tools in the animal kingdom, humans distinguished themselves by progressively introducing new tools, and improving existing ones using novel materials.

Well-agreed upon timelines exist for the progression of materials used by humans.[9] Around 2 million years ago, stone cutting implements made from stone allowed early humans to butcher wild game that they scavenged. Only 400 000 years ago, wooden spears became commonplace and hunting developed into a regular method of obtaining sustenance. 100 000 years back, intricate multi-part tools comprised of biologically derived materials such as bone, ivory or antler were first created. Around 25 millennia ago, plant fibers began to be woven together, allowing the creation of baskets, sacks and clothing. Clay began to be shaped and cured into pottery only around 5000 BC. During this same time period, humans began to focus their material design efforts to metalworking. A likely candidate for the first metal used was gold; albeit rare, it was readily found in metallic form and was more easily manipulated using primitive tools.[10] Copper followed soon after, found as native metal, but also as ores that were easy to smelt.[11] Over time, soft copper gave way to harder and stiffer alloys such as bronze,[12] and in face of rising prices and supply shortages of the critical alloying element tin, bronze slowly got replaced by the cheaper and more readily available iron.[13] Although iron was mechanically inferior to bronze, more sophisticated smelting methods allowed for precise control of carbon content in the iron, fueling the transition from ornamental iron to utilitarian steel.[14] Advances such as quenching helped avoid undesirable low-temperature processes such as phase transitions, and so on.[15]

Perhaps one of the most remarkable material advances in human history was the creation of synthetic plastics. Initially, plastics derived from natural products started appearing in the mid 1850s, with prominent examples such as polystyrene derived from an extract of the Oriental sweetgum tree (*Liquidambar orientalis),*[16] and celluloid, a mix consisting predominantly of nitrated cellulose and camphor from camphor laurel (*Cinnamomum camphora*).[17] This time period also saw a boom in the use of natural rubber, and the development of vulcanization (the process of heating rubber with sulfur to create materials of varying hardness, elasticity and durability) by Charles Goodyear, paving the way for car tires and waterproof raincoats.[18] The 1900s yielded some of the first fully synthetic plastics, such as Bakelite (deriving from phenol and formaldehyde),[19] Nylon 66 (produced by condensing diamines with carboxylic acids),[20] and plasticized polyvinyl chloride (PVC, created by radical chain reactions between vinyl chloride

monomers).[21] Today, plastics are omnipresent features of our culture and society - they preserve our foods in shelf-stable environments, adorn our bodies in the form of clothing, and are integral for medicine and electronics. They have undoubtedly become one of the most critical parts of our modern existence, so much so that a world without them would likely be unrecognizable. However, because common plastics are completely insusceptible to biodegradation, they also present an enormous environmental problem, as plastic waste continuously accumulates in our oceans and landfills (**Figure 2**).[22]

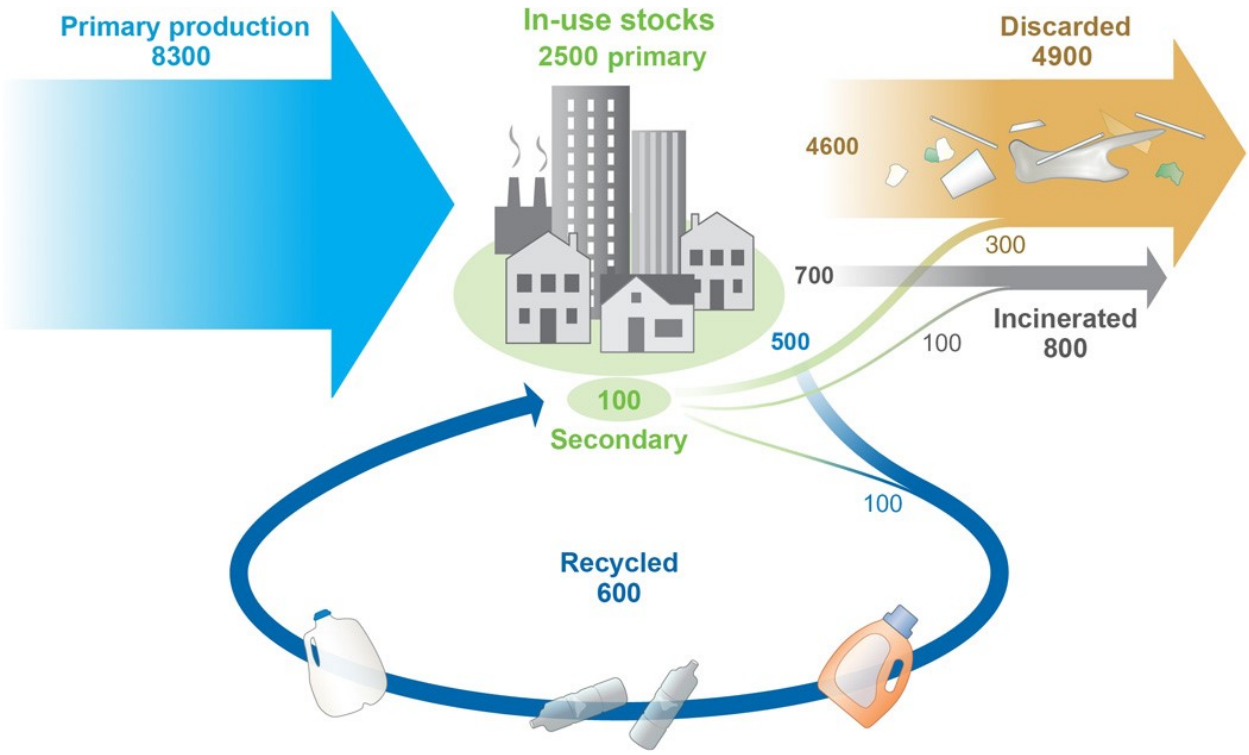

Figure 2: The global use, and ultimate fate of the polymer resins, synthetic fibers, and additives produced between 1950 to 2015 , in millions of metric tons. All 8300 million metric tons of primary production exists today as primary or secondary in-use stocks, or has been discarded or incinerated. Almost 75% of primary production has been discarded. Reproduced from Geyer et al. (2017) under a CC BY-NC 4.0 license.[22]

Throughout our speciation into today's *Homo sapiens*, we have discovered and developed innumerable materials, permitting the constructions ranging from gargantuan skyscrapers to minuscule electronics. To keep track of all these materials, their properties and what applications they are well suited for, human chemists and engineers often employ Ashby plots (also known as materials selection charts). Named after British metallurgical engineer Dr. Michael Ashby, Ashby

plots are typically 2-dimensional scatter plots of 2 different material properties (e.g. Young's modulus vs. density, **Figure 3**).[23] Essentially, one can plot whichever properties give the most relevant information to their field or material applications. No matter which properties are compared, a consistent and glaring observation is that Ashby plots contain a large amount of empty space. Simply put, there are many highly desirable material property combinations which cannot be satisfied by any known material.

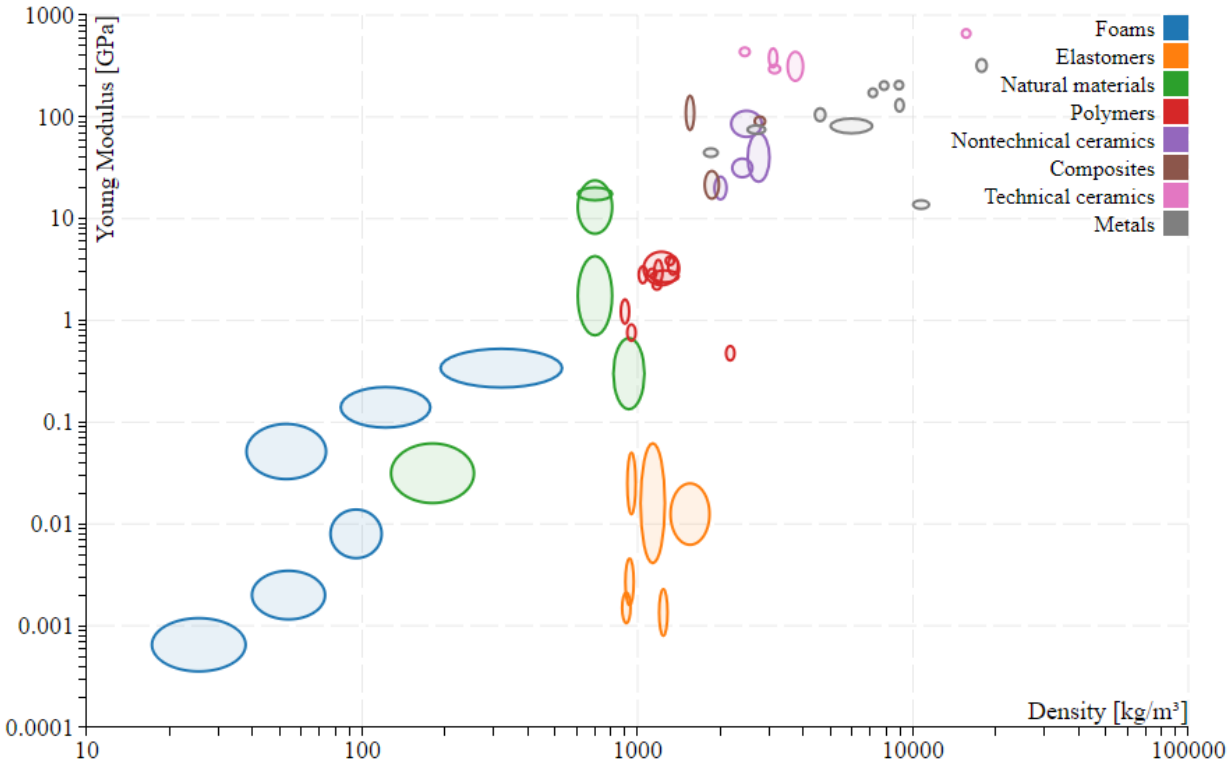

Figure 3: An Ashby plot comparing the Young's Modulus and the density of various material classes. The size of the ellipses reflect the inherent material variations present within classes. "Plot of Young's modulus vs density with log-log scaling" by Nicoguaro is licensed under CC BY 4.0.

What happens when you need a material with the high stiffness and chemical resistance of ceramic, but the material must be lightweight and resistant to catastrophic modes of failure? Alternatively, what can you use when you require the high extensibility and elasticity of rubber, but you require your material to be stiff or hard? Finding a suitable material candidate for a project becomes increasingly challenging when there are multiple competing design considerations (e.g. multiple properties that the material must excel in that do not typically track

together, such as hardness and extensibility). Moreover, properties such as bio-renewability and recyclability are becoming more important as design considerations. As such, our pursuit of better materials has become more intense than ever before. While we await new material advances that satisfy such properties, engineers must unfortunately resign themselves to curbing their ambition and utilizing existing materials that are less suited to their application.

Co-existing with the hunt for better materials (in a mechanical sense) is another pursuit for sustainable materials that derive from renewable feedstocks and possess qualities such as biodegradability or biocompatibility. The reason for this is that whilst some materials present excellent properties, they have significant deleterious effects on many of the Earth's ecosystems during their manufacturing or via improper disposal after use. Plastics are a good example of such a material; they are highly mouldable, lightweight, easy to sterilize, durable, recyclable and inexpensive.[24] These amazing properties unfortunately come with a lack of sustainability and a dramatic environmental cost: plastic production consumes around 8% of extracted crude oil annually,[25] and improper disposal into landfills or bodies of water results in the generation of harmful microplastics and the leaching of toxic phenolic additives that have powerful disruptive effects on aquatic ecology.[26-27] As the world population continually expands, the use of plastics (and many other materials that require harmful extraction, manufacture or disposal processes) continues to rise in tandem with their environmental costs and the pressure to find sustainable alternatives.

In the past few millennia, incredible advances in the field of materials chemistry have occurred without much attention being given to the potential harms of using them. The materials that will shape materials science and engineering in the future must excel not only in their material properties, but in their overall sustainability as well. Merging these properties in material design is a monumental challenge facing scientists, but despite this, there is more than a glimmer of hope available; inspiration can readily be found in nature, a source that has evolved remarkable materials that are readily assembled and broken down or recycled with minimal ruinous effects on their environment.

1.2. Creating Bio-Inspired Materials

Before we dive into the topic of bio-inspiration, it is crucial to set parameters for what it truly means. Bio-inspiration is widely regarded as the process of creating novel materials or improving existing materials by using evolved solutions found in biological systems as a role model.[28] To do this effectively, one must first study the structure-function relationships inherent to biological materials, as well as the bio-fabrication processes by which they are fashioned. Obtaining this information grants one the power to abstract desirable properties and design principles from the studied system and use them in entirely different application contexts.[29] However, effectively transferring properties from one system to another is an incredibly difficult task if one does not intimately understand the properties of the material they are working with. Out of 8.7 million species estimated to comprise the Earth's biosphere, less than 20% have been identified; from this 20%, only a truly minute fraction has been thoroughly studied. Thus, obtaining data to better understand how the native materials of these species function and are created are crucial for achieving these bio-inspiration goals.[30]

An excellent example of a property that has been repurposed by scientists is the lotus effect, referring to the ultrahydrophobicity (and the resulting self-cleaning behavior) of the lotus flower.[31] The water repellency is derived from the spiky micro- and nano-structure of the leaf surface and the subsequent waxes by which they are coated, dramatically reducing the adhesion forces of water droplets to the surface (**Figure 4**).[32] Artificially generating similar hierarchically structured surfaces on textile and metal surfaces grants them ultrahydrophobic properties, thereby unlocking a previously unattainable range of applications for these materials.[33] This bio-inspired material design process has been repeated countless times, whether it has been drawing influence from Morpho butterfly wings to make highly coloured iridescent films without the use of pigments,[34] or spiders and their silks to produce super-tough artificial tendons capable of conducting electronic impulses.[35] The numerous applications yielded by understanding the chemistry and material structure of such organisms is a testament to our thorough understanding of how their biological materials are fabricated, and how they work. Nevertheless, such information is not readily available all the time, and an immense number of organisms and their structural biology have been only superficially, or not-at-all studied. This hampers our capacity to recreate their properties and find relevant applications for them. One such system that still

contains a multitude secrets to be unveiled is the marine mussel and its remarkable capacity for underwater adhesion.

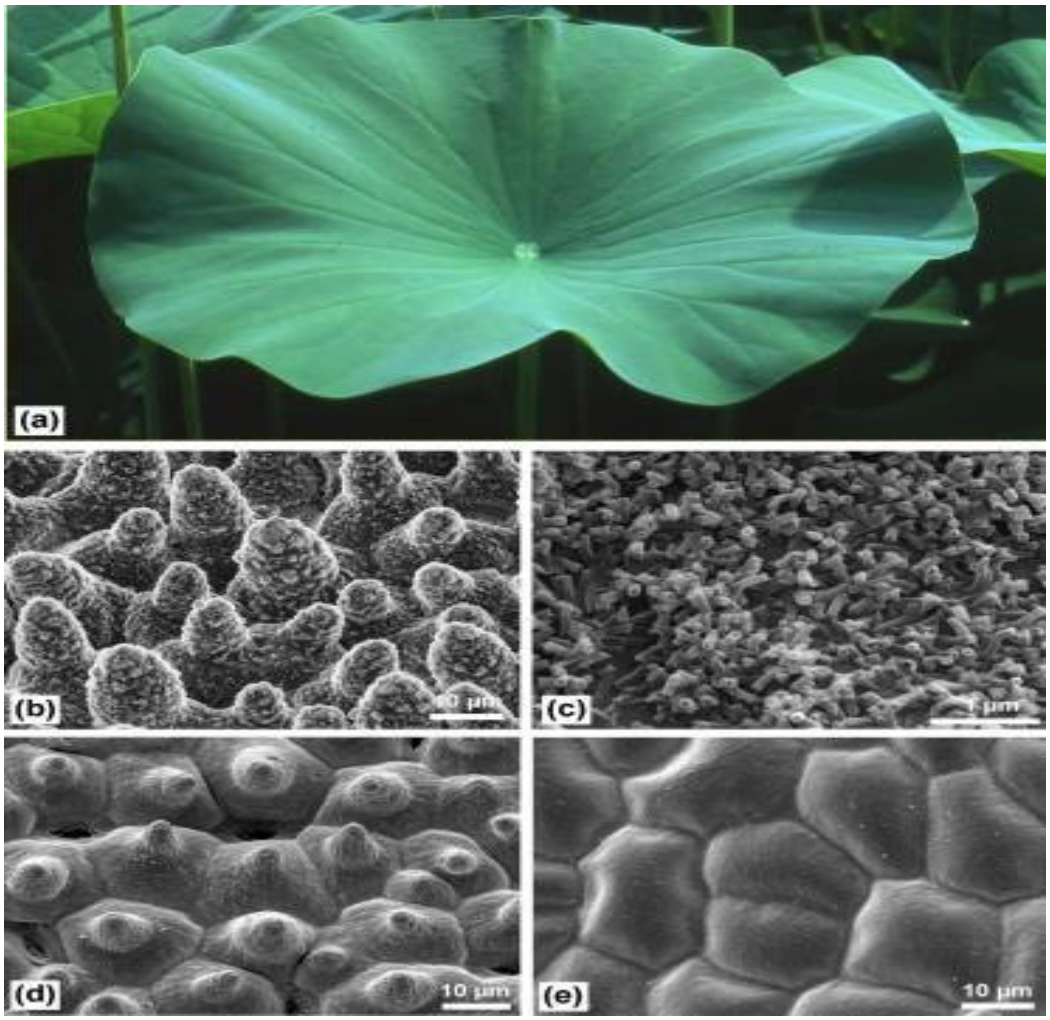

Figure 4: Bio-inspiration from lotus leaves. a) A lotus leaf that exhibits remarkable water repellency on its upper side. b) SEM of the upper leaf surface highlighting papillae covered in waxes. c) A closer look at the wax tubules formed by the leaf. d) Upper leaf and removal of waxes via critical-point drying. e) Leaf underside, bearing convex cells with completely different morphology. Reproduced from Ensikat et al. (2011) under a CC BY 2.0 license.[36]

1.3. Marine Mussels: A Bio-Inspired Role Model for Adhesion

Mussels form a large taxonomical class of bivalve molluscs that are distributed across aquatic environments worldwide. Perhaps the most widely recognized mussels are those comprising the family *Mytilidae* (due to their extensive cultivation and consumption as a food source around the globe[37]), which live attached to firm substrate in the intertidal zones of the

world's oceans and large seas.[38] To achieve this attachment in such inhospitable areas of the intertidal zone where the forces from crashing waves provide a constant threat to survival, *Mytilids* employ the use of a byssus (commonly referred to as a beard), which is a collection of thread-like filaments that allow secure anchoring to a multitude of aquatic surfaces (**Figure 5**).[39]

Figure 5: A mussel anchoring itself to substrate via its byssal threads. Collectively, these threads are referred to as a "beard" or "byssus". "Mytilus with byssus" by Broken Inaglory is licensed under CC-BY-SA 3.0.

These threads and the remarkable adhesion that they enable has been of interest to scientists for over half a century. In recent decades, the bio-inspired materials community has taken interest in studying their chemical composition, their structural hierarchy, and their bio-

fabrication processes. Although hundreds of species comprise the *Mytilidae* family, most of the studies on mussel adhesion has centered around the species of the *Mytilus* genus, such as *M. edulis* (blue mussel), *M. galloprovincialis* (Mediterranean mussel) and *M. californianus* (Californian mussel).

1.4. Morphology of Mussel Byssal Threads

The byssal threads are produced by a highly pigmented, muscular organ in the mussel known as the foot that can stretch up to 4 times its length .[40] More specifically, the foot contains a groove that begins close to the base, and ends in a wider cavity called the distal depression; it is in this groove and depression that the threads are synthesized.[40] As they are produced, these threads are attached to the byssal stem, which is anchored to the base of the foot. As one moves away from the stem along the thread, different morphologies are observed (**Figure 6**). Proximal to the stem, the emanating thread is rather elastic, becoming increasingly stiff as one moves distally along it.[41] At the substrate interface, the thread undergoes a structural transition and forms a sticky, disc-shaped plaque that anchors it to the surface. This entire thread structure is coated in a thin protective layer called the cuticle; due to this covering, the internal part of the thread is referred to as the "core".[42] Typically, dozens of threads get produced to form the entire byssal assembly, which allow the mussel to safely immobilize itself on its substrate.[40] All three regions of the thread (the core, plaque and cuticle) have further substructure which will explored further in the following sections.

1.5. Chemical Composition of Mussel Byssal Threads

Byssal threads are almost exclusively comprised of protein substituents.[40] The proteins comprising the thread proved to be rather difficult to study, due to their intrinsic propensity to cross-link with each other. This prevents the extraction of whole protein sequences for study, resigning one to the use of harsh hydrolytic methods that succeed at extracting peptides and amino acids, but consequently destroy a lot of information on the purpose of proteins and how they are organized.[43]

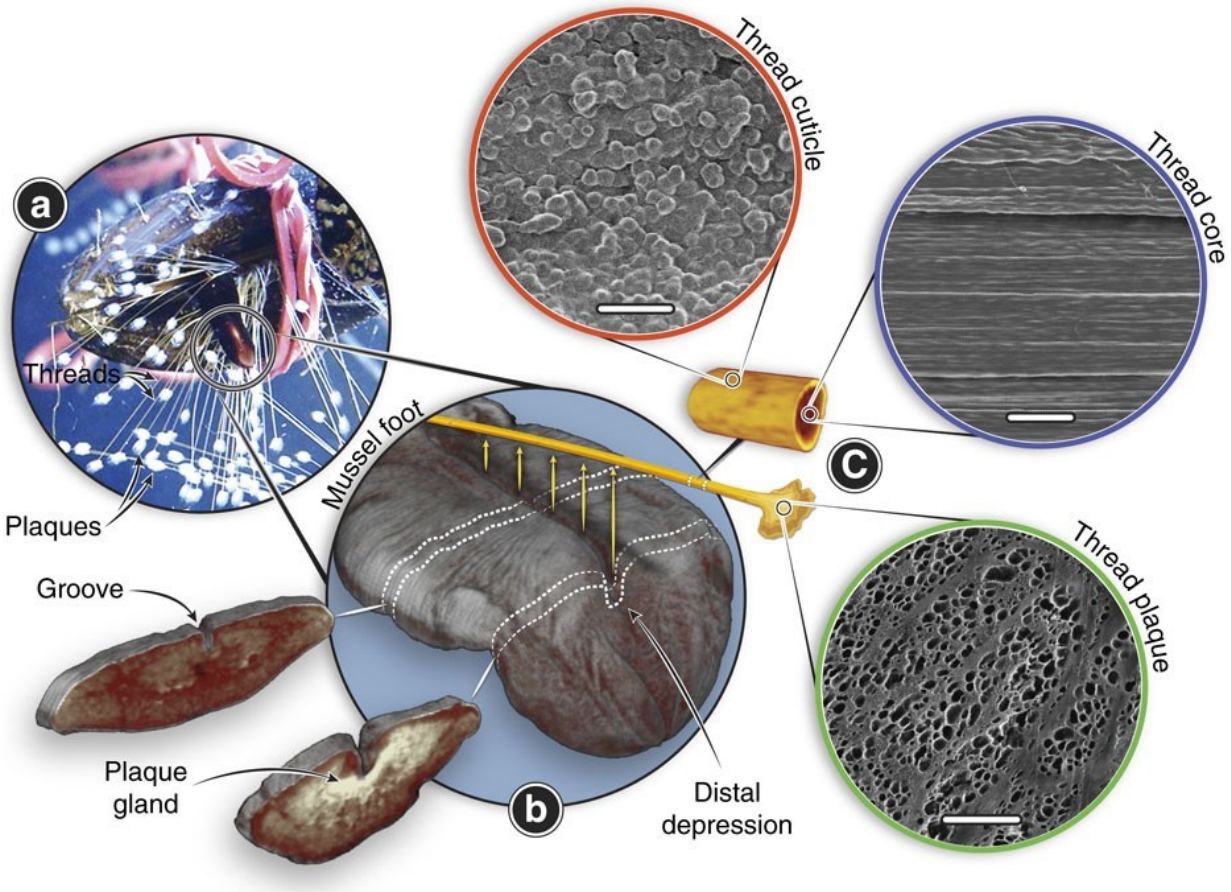

Figure 6: An overview of byssal thread morphology. a) An image of a mussel in the process of forming a thread on plexiglass. b) A μ-CT image of the mussel foot, highlighting the ventral groove and distal depression. c) Schematic of a singular byssal thread, showing the varying microstructure of the core (scale bar 5 μm), cuticle (3 μm) and plaque (50 μm) via SEM images. Reproduced from Priemel et al. (2017) under a CC BY 4.0 license.[44]

Nevertheless, the hydrolytic methods primarily permit the quantitation of amino acids comprising the proteinaceous structures of various parts of the thread. The thread contains large amounts of glycine, proline and alanine residues.[43] Meanwhile, the plaque exhibits a marked decrease in the abundance of these amino acids, and instead sees increases in tyrosine, cysteine and phenylalanine.[43] This is indicative of different protein substituents in different parts of the byssal thread. Additional hints regarding the purpose of these proteins is contained in the presence of unnatural amino acids; the thread is enriched in hydroxyproline residues and contains small amounts of 3,4-dihydroxyphenylalanine (L-DOPA), whereas the plaque contains large amounts of L-DOPA and significantly less hydroxyproline.[43] While the overall DOPA concentration in byssal threads remains fairly low, certain thread components can contain up to

30 mol% DOPA.[45] Due to its integral roles in thread adhesion and cohesion, this unnatural amino acid will receive additional discussion in a later section.

Alongside amino acid analysis of hydrolysates, identification of precursors was mainly performed through two other means. When larger, non-crosslinked peptidic fragments could be extracted from byssal fibers, sequencing via Edman degradation was initially used to identify fiber constituents.[46] Alternatively, analysis via cDNA libraries of the foot tissue proved successful by detecting the complete sequences of proteins like preCOL-P and preCOL-D using DNA probes based on partial sequences from Edman degradation.[47-48] More recently, transcriptomics of the mRNA in the mussel foot glands has gained popularity, confirming the presence of known proteins of the byssus, and suggesting the existence of "new" proteins that have eluded detection for decades.[49] Collectively, many studies that followed these general techniques provided sequence and structure information for several protein constituents of the core, plaque and cuticle.

The core of the threads contains several collagen-like proteins.[50] The primary components are the three types of pre-pepsinized collagens (preCOLs) termed distal (preCOL-D), proximal (preCOL-P) and non-gradient (preCOL-NG). These preCOLs alone can account for up to 90% of the distal thread region's dry mass.[51] While the central domains of these preCOLs are collagen-like, they are flanked on both termini by "stiff", spider dragline silk-like domains (in preCOL-D), or "compliant" elastin-like domains (in PreCOL-P).[52] Lastly, preCOL proteins are capped on both ends with "sticky", histidine rich domains that bridge interactions between preCOL termini via formation of Histidine-zinc metal coordination bonds.[53] Also found in the core are thread matrix proteins (TMPs), a family of proteins rich in glycine, asparagine and tyrosine, that eventually colocalize with the preCOLs in the thread core.[54]

The adhesive plaque structure is predominantly composed of a family of proteins called mussel foot proteins (Mfps). Mfp-2 through Mfp-6 all play large roles in forming the plaque and adhering to surfaces.[55] Mfp-3 and Mfp-5 are both highly enriched with L-DOPA residues (up to 30%), and are localized at the interface between plaque and substrate.[56] Comingling with these two proteins is Mfp-6, containing significant amounts of cysteines (11 mol%, primarily in thiolate

form), as well as a modest amount of L-DOPA.[45] As one migrates away from the interface, Mfp-2, a protein with modest cysteine content (6-7 mol%) becomes highly abundant, comprising 25% of the total plaque protein mass.[57] Beyond this one can find Mfp-4, which forms an interface between Mfp-2 and preCOL fibers extending down from the core of the thread. Mfp-4 contains a repeating histidine-rich decapeptide, and is only sparingly modified with L-DOPA.[58] Although these are the main protein components that have been isolated and studied to date, the entire composition of the plaque remains incomplete; recent transcriptomic analysis of the phenol gland has detected mRNAs for 9 other proteins, termed Mfp-7 through 15, which have yet to be isolated and studied from glandular tissue.[49]

The last structure, which is central to this thesis and will receive increased attention compared with the core and plaque in coming sections is the protective cuticle that sheathes the core and the plaque. Mfp-1 is a principal protein component of the cuticle and is composed primarily of hexapeptide and decapeptide repeats containing large amounts of proline, lysine (~20 mol%) and DOPA (~ 15 mol%).[59] Despite the unlikely prospect that the cuticle is composed solely of Mfp-1, no other protein precursors have been isolated from the cuticle to date. Recently, transcriptomic analysis has yielded evidence of mRNAs for cysteine-rich proteins in the accessory gland,[49] while fluorescence labelling and TEM studies have detected presence of thiols in the final cuticle structure.[60-61]

1.6. Structure and Function of Proteins Comprising the Mussel Byssus.

Before the structure-function relationship of these proteins can be adequately understood, it is necessary to specifically discuss the chemistry of DOPA considering the crucial role it plays in the properties of byssal threads. DOPA is the shortened name of the amino acid 3,4-dihydroxyphenylalanine, and is generally formed via oxidation of its metabolic precursor tyrosine.[62] This modification is enzymatic, performed by enzymes known as tyrosinases, which can be found in a multitude of plants, animals and fungi.[63] Despite widespread structural properties and tissue distributions, all known tyrosinases depend on a type 3 copper center active site, consisting of two copper atoms each coordinated by 3 histidine residues.[64] The addition of a second hydroxyl to certain tyrosine residues (which in effect, generates a catechol moiety)

grants protein chains a couple remarkable functionalities that are not observed in proteins whose composition is restricted to the 20 standard amino acids contained in the genetic code (**Figure 7**).

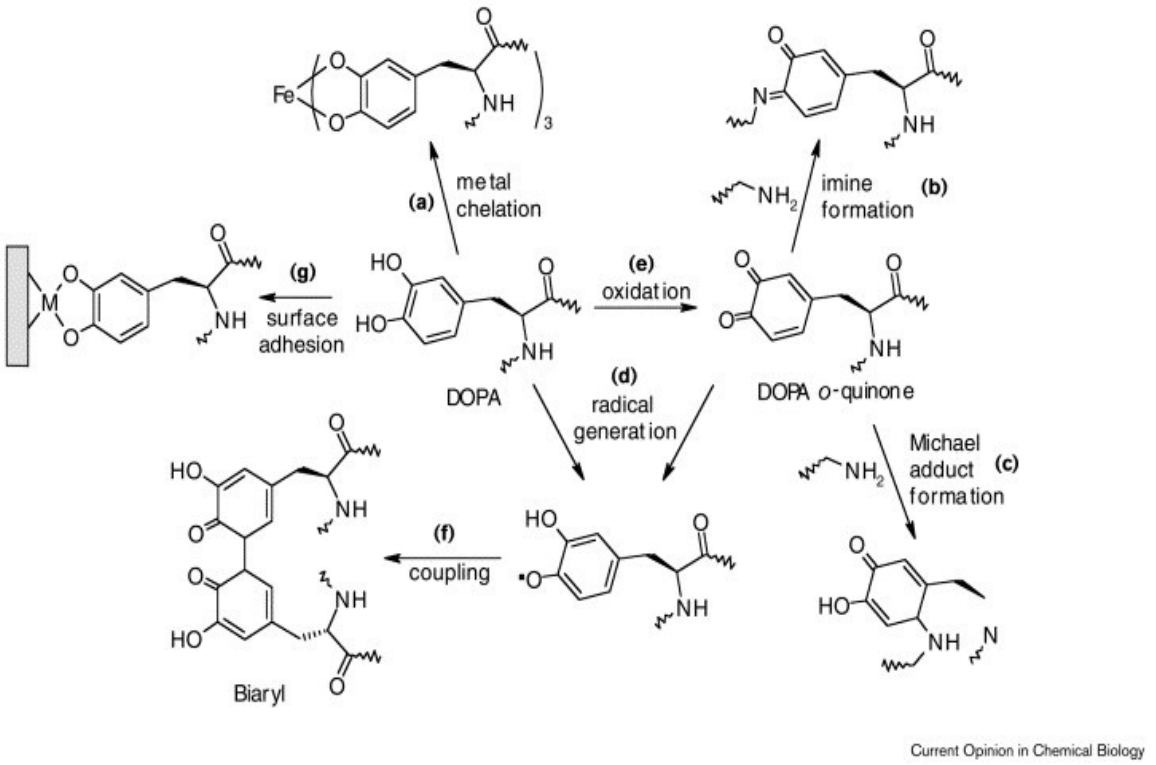

Figure 7: Possible reaction pathways for peptidyl DOPA. Depending on the intended purpose of the DOPA residues, these transformations may have positive or negative effects on the function of byssal threads. Reproduced from Deming et al. (1999) with permission.[65]

The first notable property of DOPA is its propensity to interact with metal ions; at sufficiently high pH, the hydroxyls of DOPA begin to deprotonate, and the resulting catecholate becomes a potent chelator of many di and trivalent transition metal cations.[66] The second property is DOPA's predisposition for a 2 electron oxidation process to form DOPA-quinone above neutral pH.[67] This species is highly reactive towards nucleophilic amino acids via Michael addition, Schiff base, and even free radical reaction mechanisms.[68] Interestingly, these two forms (DOPA and DOPA-quinone) have rather different properties: DOPA is not very susceptible to nucleophilic attack, whereas DOPA-quinone demonstrates lackluster metal binding capacity. *Mytilids* utilize both

compounds with an amazing degree of control, manipulating their concentrations, their oxidative state, and their specific role all throughout their byssal threads.[69]

Our deeper examination into the structure-function relationships of the byssus will begin with the byssal thread core, which is primarily comprised of the collagenous preCol proteins (**Figure 8C**).

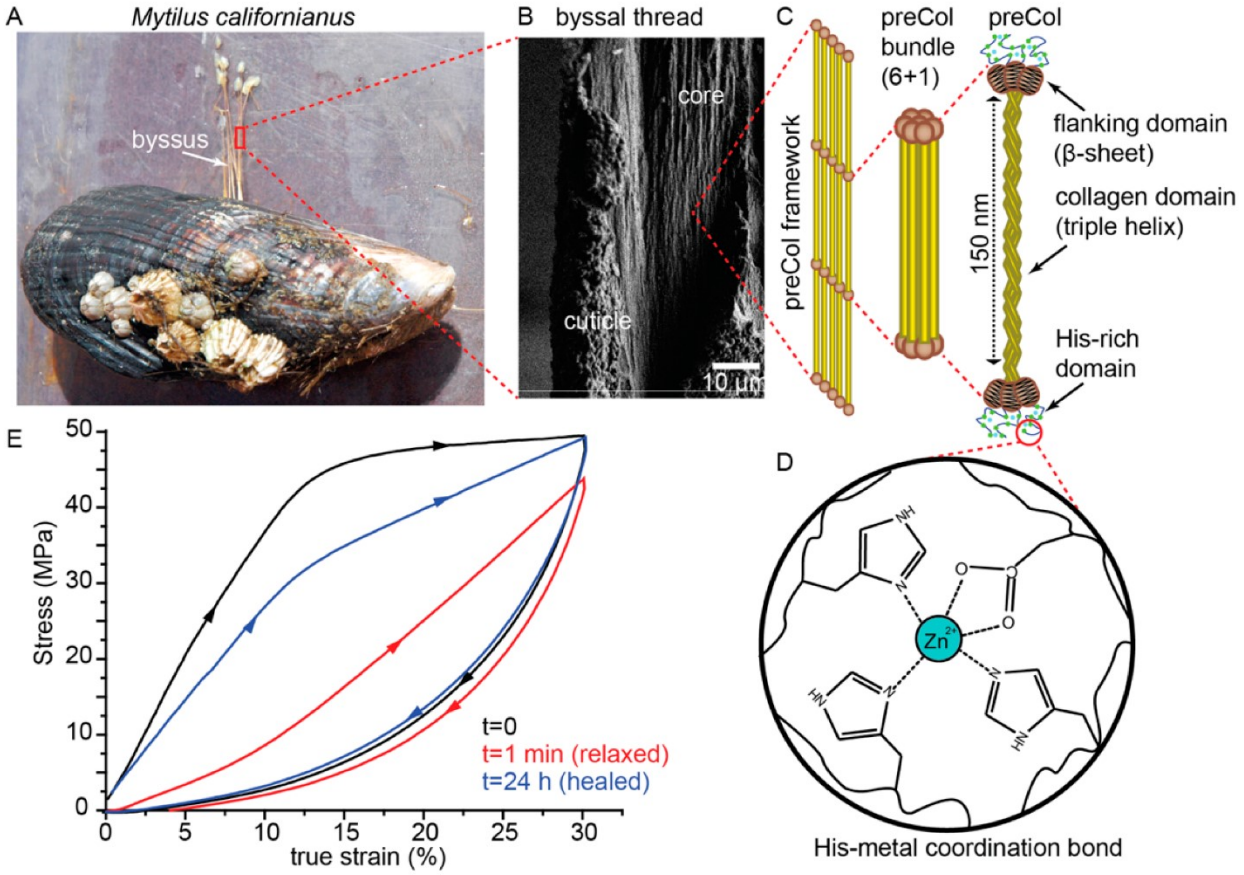

Figure 8: A) An image of a Mytilus californianus mussel and its byssus. B) SEM image of the core surrounded by the protective cuticle. C) The hierarchical organization of preCOL proteins in the thread core. D) Histidine-rich domains (HRDs) at the end of preCOL fibers coordinate Zn^{II} and provide self-healing behavior. E) Cyclic tensile testing of threads, highlighting the large hysteresis and intrinsic self-healing of overloaded threads. Reproduced from Zechel et al. (2019) under a CC BY 4.0 license.[70]

Three preCol protein chain assemble to form a collagen triple helix, which then combines with 6 other triple helices to make a 6 +1 hexagonal preCol bundle, which is the functional building block of the thread core.[71] The central collagen motif of these fibrils gives the distal thread an initial stiffness matching that of vertebrate tendon; however, unlike vertebrate tendon which fails

around 12% strain, byssal threads undergo a yield point around 15% strain followed by an extended yield plateau.[72] Remarkably, byssal threads can be extended up to 100% strain before failure,[73] and despite apparent molecular-level damage evidenced by decreased load-bearing capacities after extension beyond the yield point, near-complete recoveries of the initial mechanical properties can be achieved after a 24 hour rest period in a so-called self-healing response (**Figure 8E**).[74] Whilst the behavior pre-yield is likely modulated by the collagen domains of preCol, the high capacity for elongation is attributed to the flanking and histidine-rich domains of the preCol fibrils .[50] In the distal region of the thread, the flanking domains have sequence similarity to silk proteins with propensity for forming β-sheet structures.[50] As threads are strained and the yield point is reached, these β-sheet flanking domains begin to unfold, providing hidden length that allows for further extension.[75] Lastly, the self-healing capacity exhibited by the threads is attributed to roughly 100 amino acid long histidine-rich domains (HRDs) that cap the flanking domains at both ends of the preCols.[53] Histidines of different preCol fibrils combine with divalent cations such as Zn^{II} to form metal coordination complexes, providing sacrificial bonds that can break and reform, rather than the protein backbone which would be permanently damaged upon breakage (**Figure 8D**).[76] In addition to the preCols, there are smaller amounts of matrix proteins such as TMP-1 and PTMP-1. TMP-1 is found in the distal parts of the thread, and contains numerous asparagine-glycine repeats, which are known to spontaneously deamidate and generate numerous negative charges; while the exact role of TMP-1 remains unknown, the deamidation of Asn residues is thought to play an important role in the maturation of secreted threads.[54] Conversely, PTMP-1 does not possess Asn-Gly repeats, rather containing 2 separate von Willebrand factor-like domains; these domains have high affinities for collagen, and have been shown to affect collagen fibril assembly and morphology *in vitro*.[77]

As mentioned, the plaque contains a complex slew of proteins working together to generate the observed wet adhesion (**Figure 9D**). At the interface, Mfp-3 and Mfp-5 (**Figure 9E**) utilize their large amounts of DOPA to generate adhesion via mechanisms such as H-bonding, metal coordination, cation-pi interactions and hydrophobic interactions.[56] Critical to generating this adhesion is displacing water and cations that may interfere with the process; DOPA removes water while protonated lysine residues sweep across surfaces, displacing common ions such as

Na^+ to make adhesion more favourable.[56] The large amount of DOPA present in the interfacial proteins requires an impressive array of adaptations preventing premature oxidation and subsequent decline of adhesive capacity. First, secretion of these proteins in conducted at acidic pH, allowing DOPAs to form coordination bonds to metal oxides present on surfaces before the basic pH of seawater oxidizes DOPA to its less adhesive quinone form.[78]

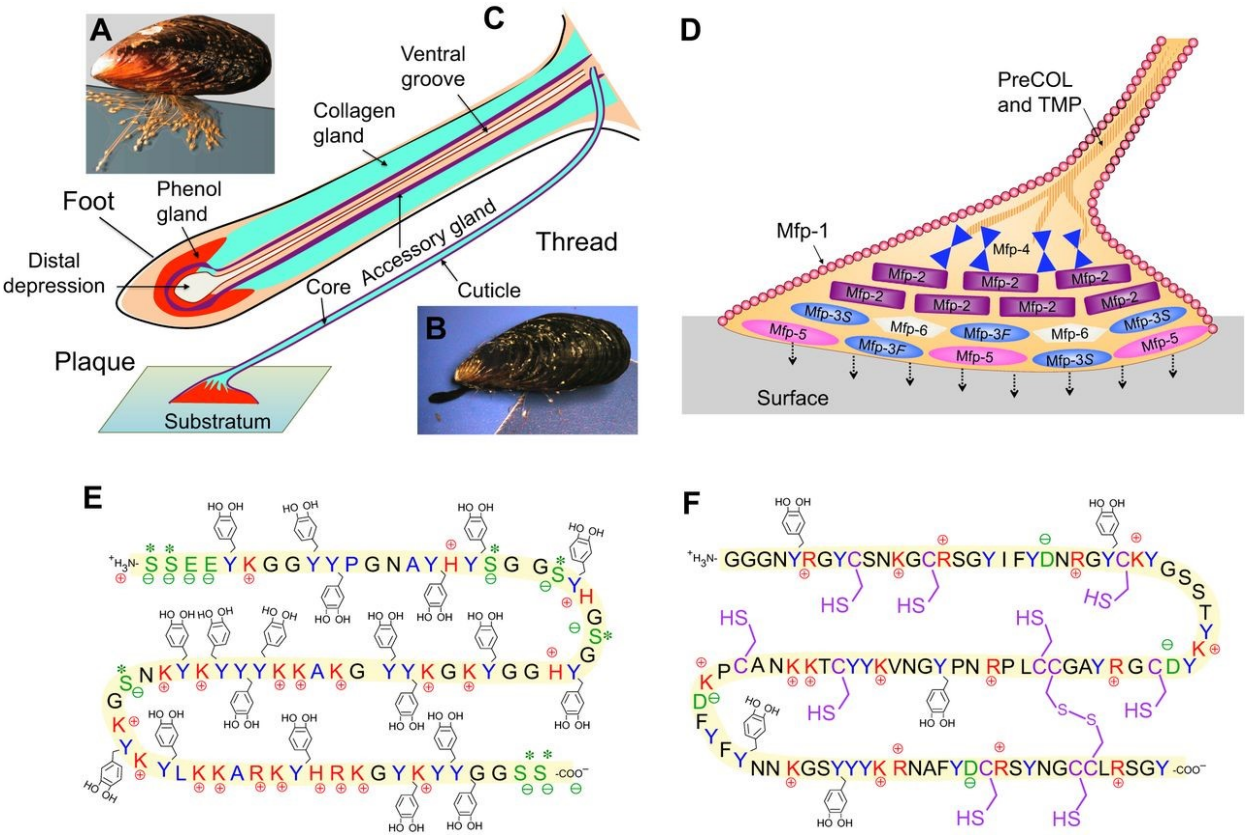

Figure 9: An overview of adhesion in the plaque. A) A mussel and its byssus. B) The mussel foot is pictured on the left of the image. C) A diagram of the mussel foot, its glands, and a thread it will typically produce. D) Distribution of foot proteins in the adhesive plaque. E) Sequence of Mfp-5, highlighting the prominence of DOPA and lysine residues. S* indicates a post-translational modification to phosphoserine. F) Sequence of Mfp-6, highlighting the prominence of cysteine residues. Reproduced from Waite et al. (2017) with permission.[56]

Secondly, these DOPA-rich proteins are granted further resistance to oxidation via co-secretion with anti-oxidants. Mfp-6 (**Figure 9F**) is one of these critical antioxidants, providing as much as 9 cysteine residues in thiol form that are thought to provide reducing equivalents for dopaquinone that forms in the plaque.[45, 79] Despite rapid equilibration of the plaque structure to the pH of

seawater,[78] a reducing environment generated by Mfp-6 is maintained within the plaque for months.[80] Perhaps the most compelling evidence to date of how redox capacity is maintained is via seawater-induced liquid-liquid phase separations (LLPS) of Mfp-3 and Mfp-6; through LLPS, these proteins seem to form reservoirs in the plaque that show exceptional insulation against the oxidative stress besieging their surroundings.[81] This plaque chemistry occurs within a scaffold comprised of Mfp-2 that resembles a porous, open-cell foam; this protein forms a highly-crosslinked dual length scale mesh (with a dense mesh of 100 nm, and larger interpenetrating pores of 1 μm) that houses aforementioned protein constituents.[82] This scaffold and its contents require an equally impressive interface connecting it to the wholly different structure of the thread core. As one ascends the thread from the substrate-plaque interface, the foamy scaffold begins to interdigitate with the core's descending preCOL fibers. Mediating this interaction between scaffold and core is Mfp-4, a bifunctional protein linker. The N-terminal half of the protein contains around 36 tandem His-rich undecapeptides, whereas the C-terminal half is comprised of 16 aspartate-rich undecapeptides; the N terminus shows strong affinity for Cu^{2+}, allowing interaction with the HRDs of PreCOL fibers, whereas the C terminus preferentially binds Ca^{2+}, allowing for a different coupling interaction to occur with the plaque scaffold.[58] Much like the core, this entire plaque assembly is sheathed by the cuticle – the primary focus of this thesis.

While the thin cuticle coating appears unassuming at a glance, its chemistry allows for some interesting structural features. Although DOPA has been discussed in some detail, another small preamble is warranted before proceeding, as the DOPA in the cuticle is not used for adhesion, but for cohesion. This cohesion is achieved by load-bearing catecholato-transition metal complexes, which have only been identified in a few other biological systems and in these cases typically for non-material functions. For example, a well characterized set of catecholato-transition metal complexes have been observed in bacterial siderophores. Siderophores are chelating molecules released by bacteria into their local environment to scavenge Fe^{III}.[83] This can be achieved by several chemistries, but the most powerful siderophores (such as enterobactin, **Figure 10**) employ 3 catechol groups to form complexes with ferric transferrin ($K_a \approx 10^{51}$ at neutral pH, which is 25 orders of magnitude greater than EDTA), after which transfer of the Fe^{III} particle to the siderophore occurs.[84] This affinity is so large, that enterobactin can effectively solubilize

Fe^{III} from insoluble rust particles, as well as strip tightly bound iron from heme proteins.[83, 85] While it is not related to its biological function, enterobactin-vanadium complexes have also been produced *in vitro* showing incredibly strong binding stabilities.[86]

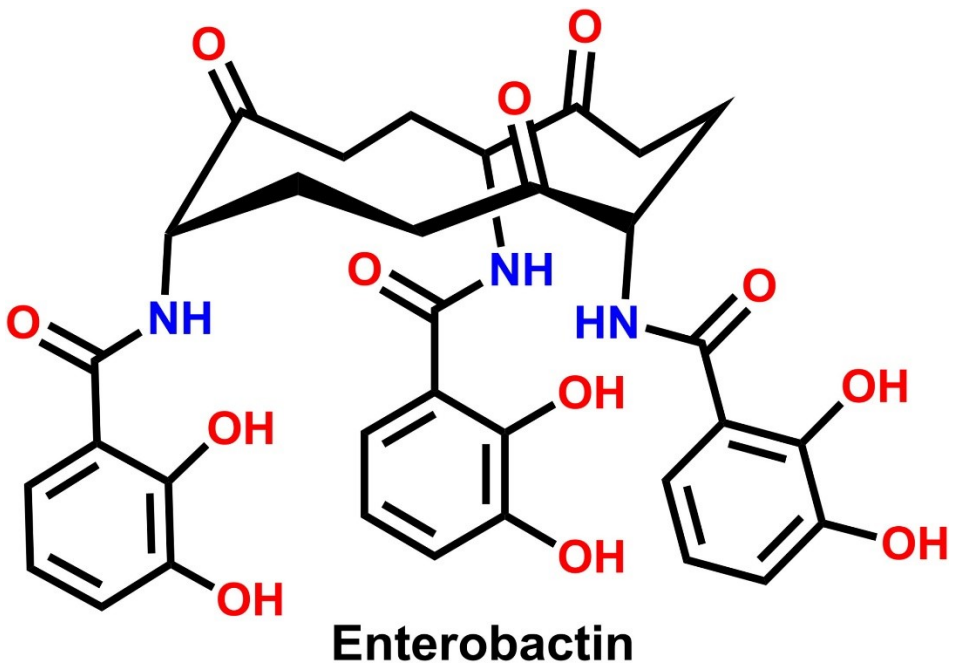

Figure 10: The structure of Enterobactin. It is a catecholate siderophore commonly found in Gram-negative bacteria such as E. coli. Reproduced from Rohini et al. (2019) with permission.[87]

Additionally, vanadium hyperaccumulation and complexation has been observed in ascidians (also known as sea squirts or tunicates) with some species concentrate vanadium up to 10^6 times higher than what is found in seawater.[88] Vanadium-binding was originally ascribed to tunichrome, a highly reducing family of peptide-based blood pigments rich in catechol or gallic acid; however, later studies clearly demonstrated that tunichrome and vanadium are compartmentally segregated and never interact *in vivo*.[89-90] Rather, the vanadium in ascidians may be complexed with a protein known as vanabin (that does not contain catechols) and stored in highly acidic vacuoles at a low oxidation state.[88]

In this light, it is clear that the mussel's use of cohesive catechol-metal coordination complexes for a mechanical function is distinctive in biology and has not been identified in other biological materials based on our literature exploration. Moreover, the manner in which they are

used is quite complex and highly interesting from a biomimetics perspective. The cuticle of mussels living in intertidal zones (like the Mytilids) frequently contains a biphasic structure, consisting of hard granules interspersed within a softer matrix (**Figure 11**).[42] Based on recent findings, the DOPA-rich mussel protein Mfp-1 is believed to be primarily localized within the granules, while the matrix contains putative Cys-rich proteins.[61]

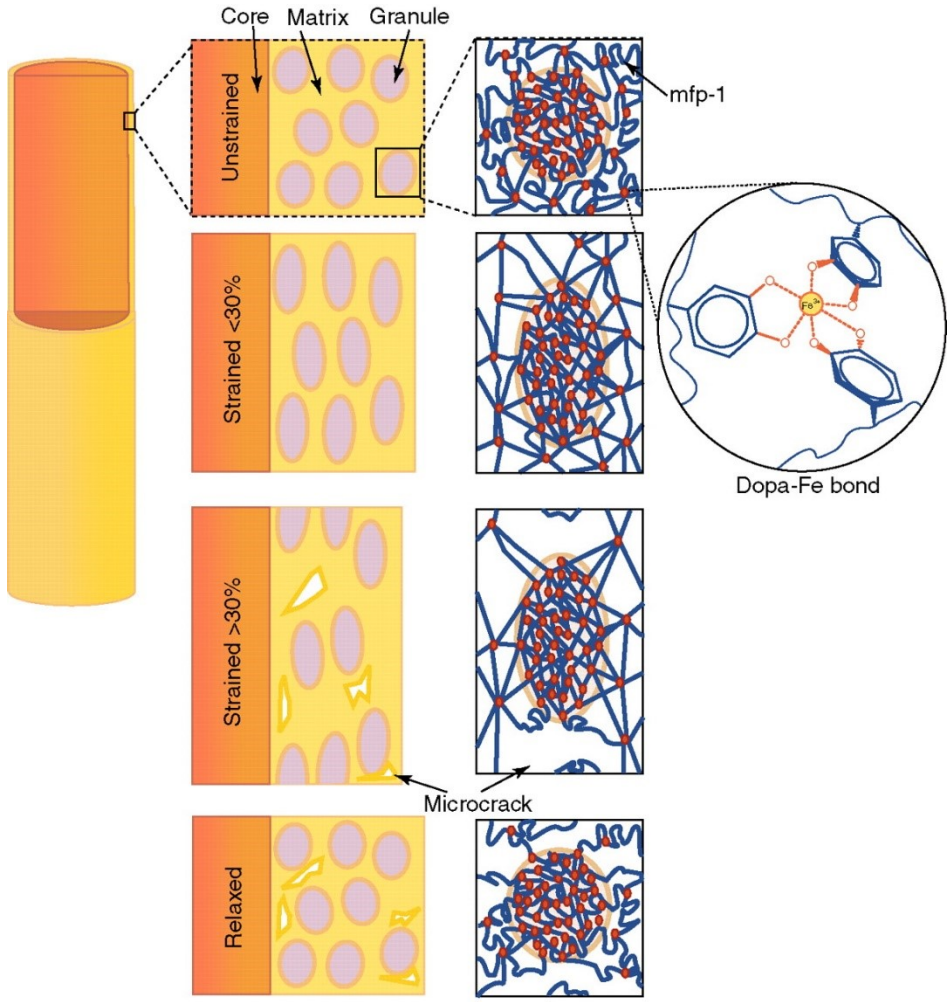

Figure 11: A model depicting cuticle composition and resultant behavior in response to applied strains. The higher crosslink density in the granules vs the matrix result in microcrack formation in the latter when strained beyond 30%. Reproduced from Harrington et al. (2010) with permission.[91]

As stated before, Mfp-1 contains a large amount of DOPA (up to 15 mol%) and unlike the plaque, this DOPA is not used for adhesion to surfaces, but rather for cohesion between Mfp-1 proteins.[91] In literature, this cohesion is thought to be achieved via the formation of complexes between

DOPAs and metal ions such as Fe^{III} and V^{III}, which have been observed first *in vitro,* and then *in situ* using Resonance Raman Spectroscopy.[91-93] Different DOPA-metal cross-linking densities in the granule versus the matrix were previously proposed as the mechanism providing the cuticle with a remarkable capacity for handling large strains. Initially, the cuticle accommodates strains up to 30% by unravelling protein chains, at which point further strain shifts the load to the DOPA-metal complexes; as the cross-link density is lower in the matrix, these cross-links break first, creating microtears and delaying failure of the material.[91] However, recent evidence suggests that Mfp-1 is co-localized with V in the granules, while the matrix is enriched in both Cys- and possibly DOPA-rich proteins and Fe, suggesting a different mechanism may be at work (**Figure 12**).[61] As DOPA-vanadium complexes have bond lifetimes an order of magnitude larger than DOPA-iron, it stands to reason that these complexes would confer a higher mechanical stability under viscoelastic loading than the DOPA-iron.[94] Additionally, it has been proposed that the cysteine residues in the matrix might interact with the Fe ions creating a second class of load-bearing metal complexes with different mechanical stability;[61] however, this remains to be verified experimentally.

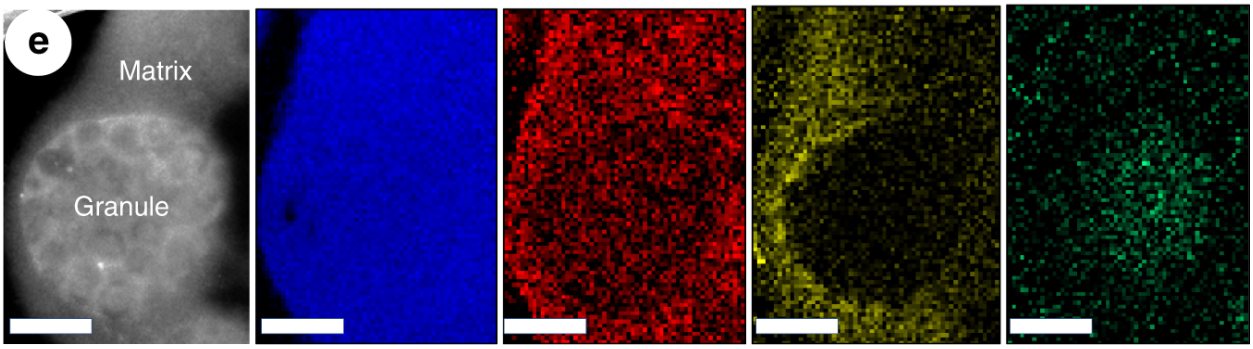

Figure 12: Scanning transmission electron microscopy high-angle annular dark-field imaging (STEM HAADF) compositional analysis of native cuticle sections. Energy-dispersive X-ray spectroscopy (EDS) is utilized to analyze the elemental composition of the cuticle sections. A notable anticorrelation between Vanadium and Iron signals is observed in the granule portion of the cuticle, which is thought to contain greater amounts of Mfp-1. Adapted from Jehle et al. (2020) under a CC BY 4.0 license.[61]

Importantly, these very recent findings pose a great challenge to the longstanding belief that the DOPA-metal bonds in the cuticle are predominantly iron-based.[61] Thus, as both DOPA-iron and DOPA-vanadium complexes likely bear relevance in the study of the Mytilid cuticle, it is

well worth discussing characteristic properties of these complexes that may be useful in later studies. An immediately obvious property of these complexes is their production of vivid colours and characteristic UV-Vis spectra as a function of increasing pH; using both pyrocatechol and a DOPA-containing peptide, these colour changes were correlated to increasing catechol coordination of the metal ion, providing an important fingerprinting method for identifying these complexes in different systems.[66, 95] For catecholato-iron complexes in Mytilids, the metal ion of interest is the highly stable Fe^{III}, as the other common ion Fe^{II} has significantly lower affinity for catechol ligands.[96-97] Conversely, catecholato-vanadium complexes are less understood in Mytilids, owing to the stability of catechol complexation with V^{III}, V^{IV}, and V^{V}, and the accessible transitions between oxidation states that can occur in biological systems.[90] References to vanadium in mussel literature have not probed the oxidation state of the ion, inviting future experimentation on the topic.[93] When Fe^{III} complexes achieve their tris-catecholato forms, their geometries are most commonly described as octahedral, however, crystallographic studies have demonstrated that the inability of catechol to achieve 90° O-Fe-O bond angles forces distortions towards trigonal prismatic geometries.[98] Similar crystallographic evidence also exists for V^{III} complexes and V^{IV} complexes, with the latter having a severe twist angle of 39° (where 60° is perfectly octahedral and 0° is perfectly trigonal prismatic).[99-100]

1.7. Assembly of Mussel Byssal Threads

Mussel byssal threads are fabricated by a specialized organ known as the mussel foot in a rapid assembly process in which liquid protein precursors synthesized and stockpiled in the foot are secreted into a narrow groove running along its ventral side (**Figure 13**).[44] The foot contains 3 principal types of glands that mass produce the protein building blocks for the different thread structures: the core gland produces protein precursors for the byssal core, while the plaque and cuticle glands do the same for their respective structures.[44] It is important to note that older literature has used a variety of names for these glands: the core gland has been called the white or collagen gland,[101-102] while the plaque gland has been the purple or phenol gland,[103] and the cuticle gland has been the referred to as the enzyme or accessory gland.[104-105] For brevity, the generally understood method of assembly for all tissues will be discussed. Within the glands,

protein precursors are produced, modified, and subsequently packaged into secretory vesicles. Typically, these vesicles are secreted into the groove of the foot with a high degree of control to produce a thread.

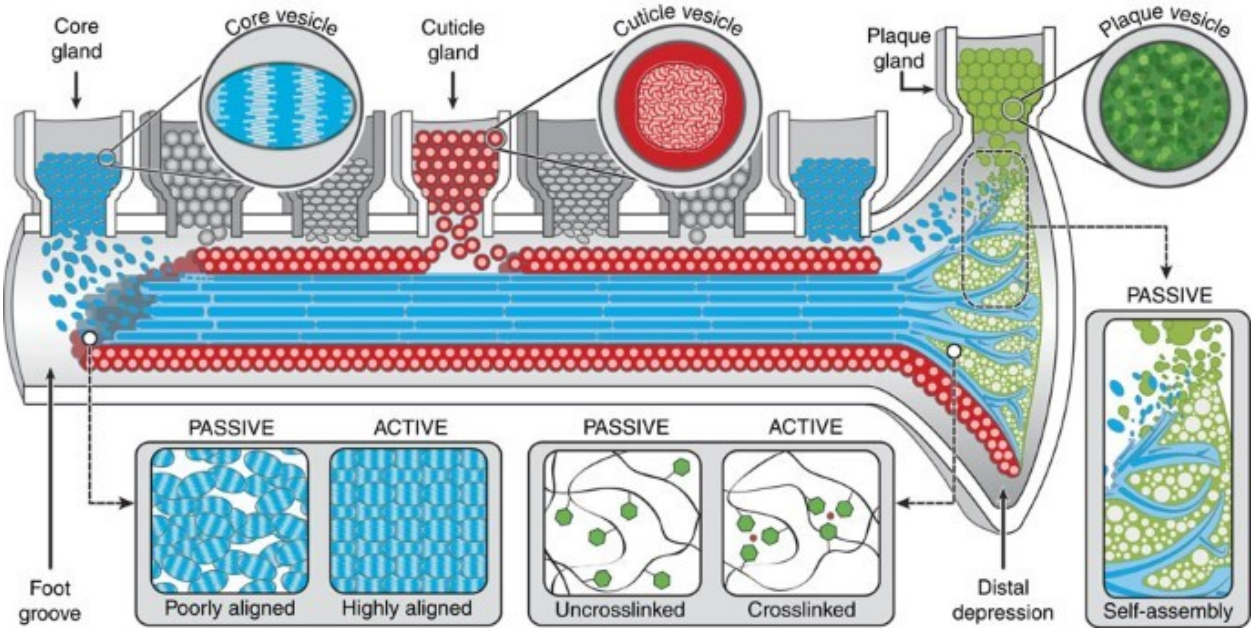

Figure 13: An illustrative model of the byssal assembly process. Vesicles containing different structural materials are secreted from glandular tissue in the foot, into the foot groove. In this groove, the mussel employs active control (such as release of metal ions and flexion of the foot), and passive controls (such as properties intrinsic to the proteins being secreted). Reproduced from Priemel et al. (2017) under a CC BY 4.0 license.[44]

One of the best ways to study the formation process is by capturing "snapshots" of the formation process within the mussel foot; currently, this is a novel idea with few literature examples. One such example is a previously mentioned study that has captured remarkable electron microscopy images of the cuticle vesicle secretion process (**Figure 14**).[61] An indirect way to study the formation process is artificial induction of threads via injection of 0.56 M KCl into the pedal nerve of the foot, causing uncontrolled contraction of the foot muscles and subsequent vesicle release into the ventral groove of the foot.[103] Threads formed in this manner are an important contrast to ones produced naturally, contributing valuable insight into how the mussel might need to control the formation process. Overall, the individual protein precursors do possess some capacity for self-assembly, however, it is also clear that some form of secretory control is required of the mussel to produce optimal threads.[44] Exactly how the process must be

controlled remains unclear. Among the core, cuticle, and plaque formation processes, the plaque has arguably received the most study, likely owing to the relatively higher accessibility of the plaque during formation, as opposed to the core and cuticle which are produced deeper within the foot.

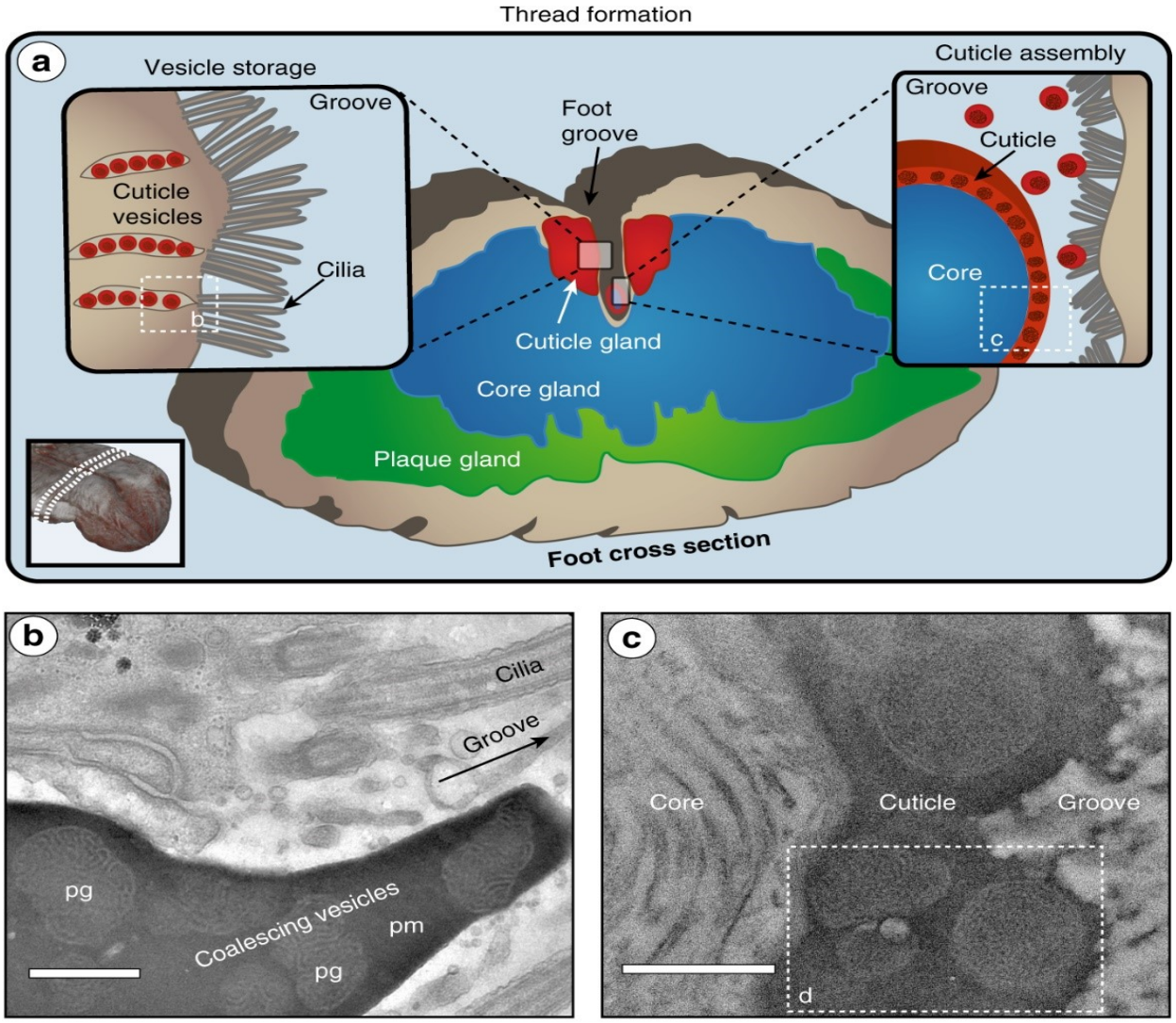

Figure 14: Transmission electron microscopy – energy-dispersive X-ray spectroscopy (TEM-EDS) studies of byssal cuticle secretion. a) Schematic depicting a transverse cross-section of the mussel foot, with the approximate location of the glandular tissue surrounding the foot groove. b) a TEM image showing coalescing cuticle vesicles about to be secreted into the foot groove. pg indicates protogranule and pm indicates protomatrix, which are the fluid precursors to the solid, cured cuticle. c) a TEM image of the cuticle secreted and deposited around the core of an induced thread. d refers to another panel that was not included in this figure. All scale bars are 500 nm. Adapted from Jehle et al. (2020) under a CC BY 4.0 license.[61]

As evidenced in **Figure 15B-J**, many steps must occur sequentially to form the functional plaque. Despite few studies detailing the formation processes of the core and cuticle, the nature of their protein constituents do hint towards the presence of similar steps such as pH and redox control, coacervation and curing.[61] Beyond thread formation, redox processes likely remain active in all parts of mature threads, shielding residues such as DOPA from the oxidizing environment present in seawater (**Figure 15F**).[81]

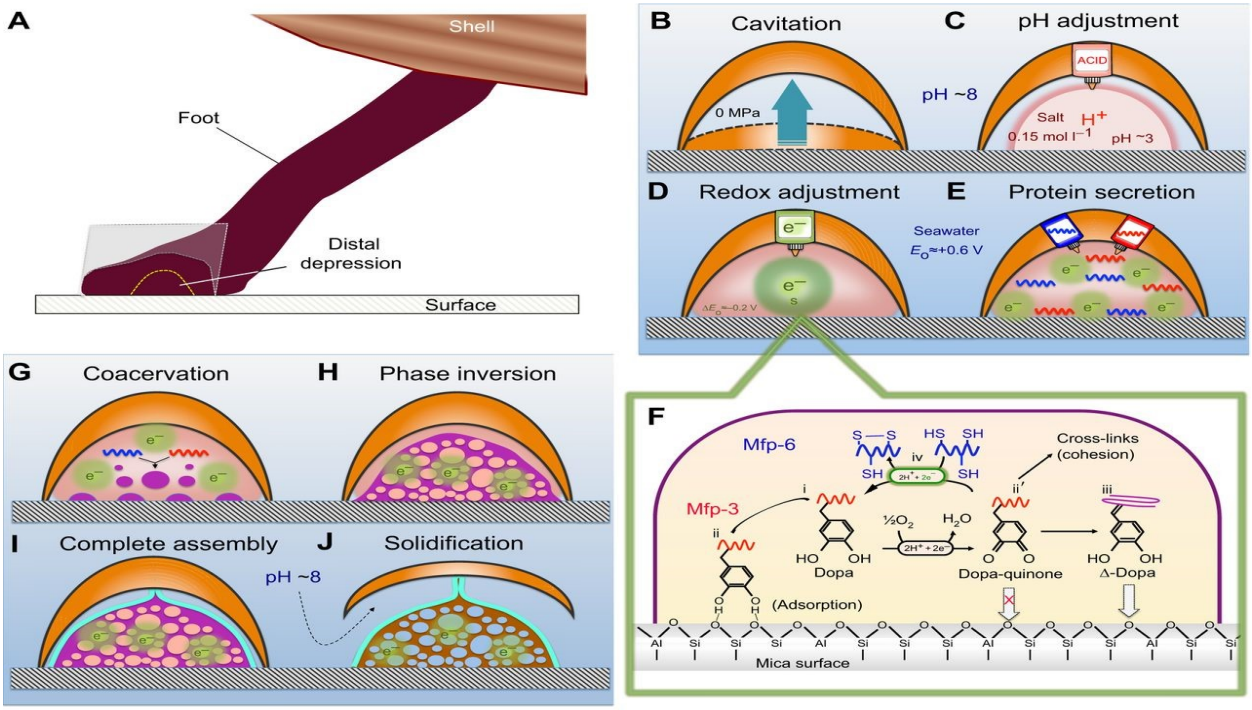

Figure 15: A) depicts a mussel in the process of thread production. B) through J) show the sequential steps involved in plaque maturation. F) demonstrates the redox reactions that are thought to occur between plaque proteins Mfp-3 and Mfp-6. Reproduced with permission from Waite et al. (2017).[56]

Coacervation, alternatively known as Liquid-Liquid Phase Separation (LLPS), deserves its own small section, as the evidence for its role in byssus assembly builds up. Briefly, coacervates are condensed fluid phases of molecules (in this case proteins) which have phase separated from the solvent phase due to favorable intermolecular interactions, including hydrophobic forces and electrostatic interactions.[106] Because of this, coacervates are highly dense, dynamic phases that exhibit fluid-like characteristics including surface wetting, coalescence with other coacervate droplets and exchange of contents with their surroundings.[107] Overall, coacervates are grouped

into complex and simple coacervate categories. A complex coacervate typically involves two or more oppositely charged polyelectrolytes being driven together by associative forces (like proteins and nucleic acids coming together in the nucleolus) under appropriate conditions of pH and ionic strength. Simple coacervates, on the other hand, are typically generated by segregative forces (such as reduction of repulsive interactions in β-amyloid fibers)[108] or hydrophobic forces between non-polar molecules in an aqueous environments (such as in the case of tropoelastin).[109] Currently, it is believed that mussels predominantly employ simple coacervates to achieve their underwater adhesion, but much of their properties and exact functions remain unknown.[110-111] The implications of coacervation on Mfp-1 and our experiments will be discussed in a later section.

1.8. Goals of This Thesis

As evidenced, our understanding of byssal structure, function and assembly has steadily been improving over time. Despite this, many unknowns still exist, especially in the context of the mussel cuticle. Given the bio-inspiration already provided by the mussel byssus (in domains such as wet adhesion and self-healing polymers), it stands to reason that furthering our understanding of its structure and function will continue to drive innovations in the field of sustainable materials with remarkable properties. In this thesis, I have two primary objectives: First, I aim to provide evidence of additional protein constituents forming the cuticle (e.g. the putative Cys-rich proteins in the matrix) by studying sections of byssal cuticle using Tip-Enhanced Raman spectroscopy (TERS). Secondly, I intend to perform an in-depth *in vitro* functional analysis of Mfp-1-metal binding, specifically comparing iron vs. vanadium binding with the goal of understanding how and why vanadium appears to be the preferred binding metal in cuticle formation and function. In this second objective, I will utilize rMfp-1-DOPA (a recombinantly expressed truncated form of the Mfp-1 cuticle protein modified with DOPA) analyzing iron and vanadium binding using UV-Vis spectroscopy, resonance Raman spectroscopy, electron paramagnetic resonance (EPR) and surface force apparatus (SFA) techniques. Before continuing to the next part of this thesis, a brief overview of the techniques utilized is provided for readers who may not be familiar with them.

1.9. Resonance Raman and TERS

Resonance Raman occurs when a molecule of interest possesses a chromophore with electronic transition that can be stimulated by an incident laser; when this occurs, the vibrations of certain bonds in proximity to the chromophore can couple with the electronic transition, resulting in a Raman signal enhancement of up to 10^5 for those vibrations (**Figure 16**).[112]

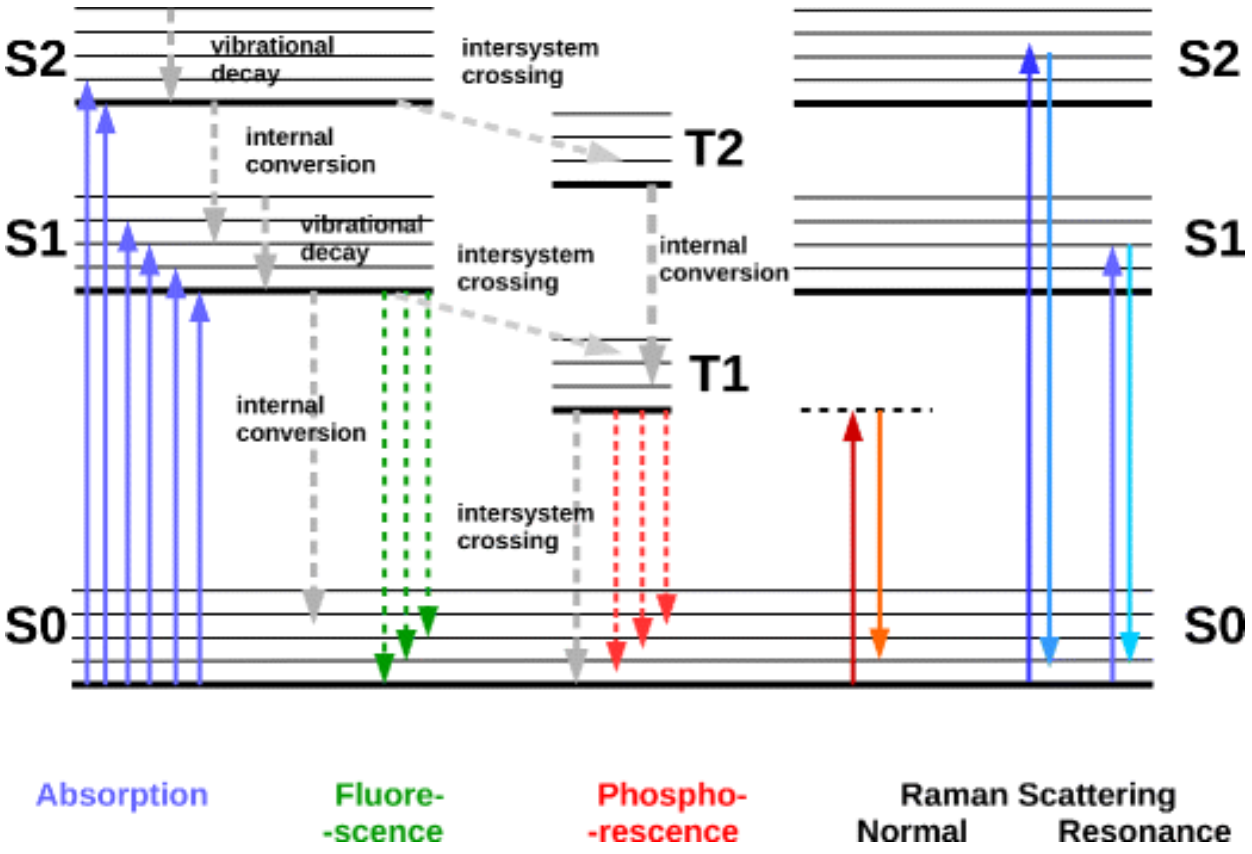

Figure 16: A Jablonski diagram highlighting various absorption and emission phenomena. The far right example depicts Resonance Raman, where a change in electronic energy (i.e. S0 to S1) is coupled with a change in vibrational energy, resulting in a highly amplified Raman signal. Reproduced from Antosiewicz et al. (2016) under a CC BY 4.0 license.[113]

TERS, on the other hand, generates an enhanced Raman signal by bringing an atomically sharp gold or silver tip in contact to the sample, and subsequently focusing a laser on the tip, in a manner analogous to more traditional surface enhanced Raman spectroscopy (SERS).[114-115] Due to a strong surface plasmon resonance localized at the tip-sample interface, the local electric field is significantly increased, providing the possibility of enhancing the Raman signal by a factor of up to 10^{14}.[115] Notably, these tips also function as Atomic Force Microscopy (AFM) probes, and

the coupling of AFM and TERS is synergistic, allowing for simultaneous surface mapping and vibrational inspection of a sample. What makes TERS uniquely powerful method is the localization of its resonance amplification; while the resolution of Resonance Raman using optical microscopes is limited to approximately half the wavelength of laser light used (roughly 250 nm for a green laser), TERS generates an enhanced Raman signal from molecules in a 10-20 nm radius of the tip.[116] This allows for detection of signals and trends in samples that would be lost in ensemble Raman measurements, such as distinguishing between bovine pancreatic enzymes RNAse A and B on a coated surface, which differ only by their glycosylation patterns.[117] On the other hand, the spectra produced with TERS, which arise from interactions of the tips with only a very small number of molecules or even just specific vibrational modes of molecules, can deviate significantly from traditional ensemble Raman spectra, which present an average of the vibrational modes of a very large number of molecules. Here, TERS data acquired from cuticle sections will be analyzed in detail in hopes of elucidating key fingerprints from other protein constituents in the cuticle. Resonance Raman will be used on rMfp-1-DOPA samples to better understand the nature of the metal-DOPA complexes this protein forms.

1.10. Surface Force Apparatus (SFA)

SFA is an instrument used to measure forces of adhesion or cohesion between various molecules and substrates (**Figure 17**). For cohesion experiments, 2 mica plates of the SFA are coated with a polymer of interest and brought into contact. After a certain contact time, they are pulled apart to determine the cohesion energy; throughout the experiment, the distance between the 2 plates and the forces between them are constantly monitored by optical interferometry and piezoelectric motors, respectively.[118] SFA will be used to measure the strength of cohesion of rMfp-1 films in the presence of differing metal ions.

1.11. UV-Vis Spectroscopy

UV-Vis spectroscopy is a technique that probes the light absorbance of a molecule in the 200-900 nm range, which is often useful for identifying analytes, and determining their concentrations.[119] The wavelength of light absorbed is proportional to the energy required to

excite an electron from the ground electronic state to an excited electronic state (**Figure 16**), while the intensity of the absorption (expressed as an extinction coefficient ε) is related to how "allowed" the transition is based on orbital and symmetry considerations (i.e. transitions between different orbital types p -> d would be favoured, while d -> d would not).[120]

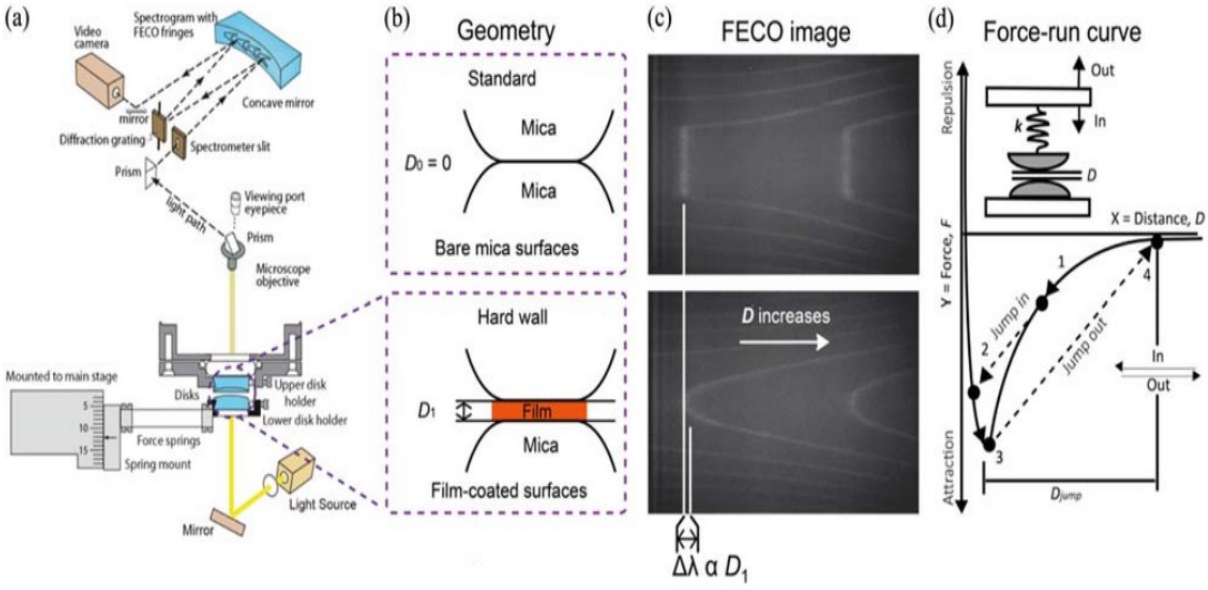

Figure 17: a) shows a schematic view of the SFA instrument. b) demonstrates the experimental difference between bare mica plates and plates covered with a film. c) shows how this difference is resolved using Fringes of Equal Chromatic Order (FECO). d) demonstrates the experimental setup, and how it translates to a force-distance profile. Reproduced from Oh et al. (2014) with permission.[121]

While many molecules can be studied using UV-Vis, molecules containing transition metal ion complexes are particularly well suited for this method, as they almost always produce d -> d transitions, and occasionally highly intense (ε in the 1000s) Ligand to Metal Charge Transfer (LMCT) bands that are very useful for determining the extent of coordination of metal ions.[95] This thesis will utilize UV-Vis to identify complexes formed between rMfp-1-DOPA and metal ions, as well as to study the complex formation pH dependence between rMfp-1-DOPA and other catechols.

1.12. Electron Paramagnetic Resonance (EPR) Spectroscopy

EPR is a technique that excites unpaired electron spins to elucidate structural information about a molecule. At its simplest, a molecule is placed in a magnetic field, the energy level

degeneracy of an unpaired spin +1/2 or -1/2 electron is lifted, and microwaves can be used to stimulate transitions between the two energy levels (**Figure 18**).[122]

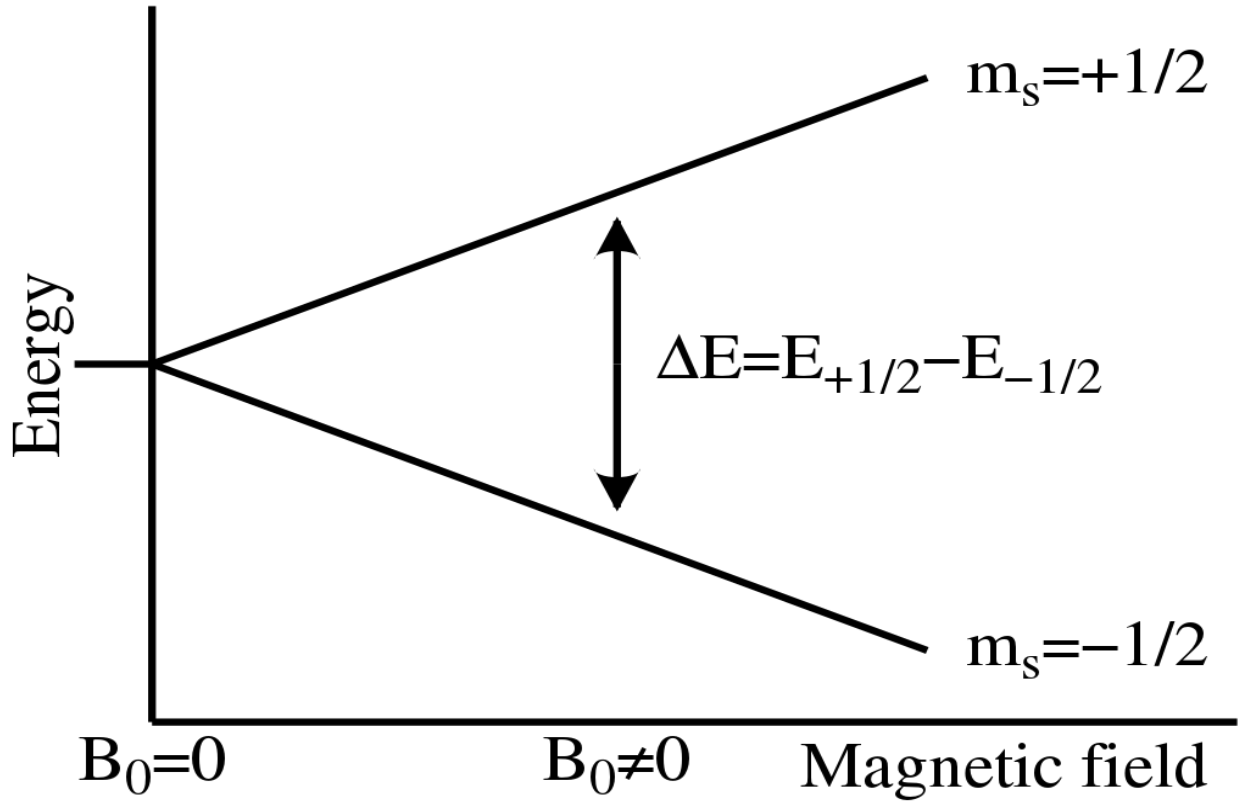

Figure 18: An energy level diagram of an electron in a magnetic field. As the strength of the magnetic field increases, so does the energy difference between spin ½ and spin – ½ states. When this difference matches the energy provided by the microwave source, a transition occurs and is marked by an absorbance in energy that forms the basis of EPR analysis. "EPR Splitting" by Krishnavedala is reproduced under Public Domain.

The exact absorbance by the electron is highly dependent on parameters from its chemical environment, such as non-zero nuclear spins in its vicinity and ligands which can perturb the electronic symmetry of the molecule, making this an excellent tool for probing the molecular environment of an analyte of interest. The critical requirement for EPR is the need for unpaired electrons, thereby limiting its use to organic radicals, excited states, and certain transition metal ions. Considering vanadium, VO^{IV} is an excellent $3d^1$ EPR probe, being the only paramagnetic metal ion complex that gives consistently sharp spectra in room-temperature and frozen solutions.[123] EPR will be utilized to better understand the formation and speciation of metal complexes formed by rMfp-1-DOPA in the presence of V^{III} and VO^{IV}.

2. Experimental

2.1. Recombinant Expression of truncated Mfp-1

2.1.1. Plasmid Design and Transfection into Bacterial Strains

The plasmid to produce rMfp-1 (AKPSYPPTYK$_{12}$) was obtained as a generous gift from Dr. Dong Soo Hwang (Pohang University of Science and Technology). The DNA coding for the 120 amino acid rMFP-1 sequence was contained in a pET-28a(+) plasmid (EMD Millipore), inserted between the NcoI and HindIII restriction digest sites (**Figure 19**). Competent *E. coli* stocks of DH5α and BL21(DE3) were kindly provided by Dr. Christopher Thibodeaux (McGill University).

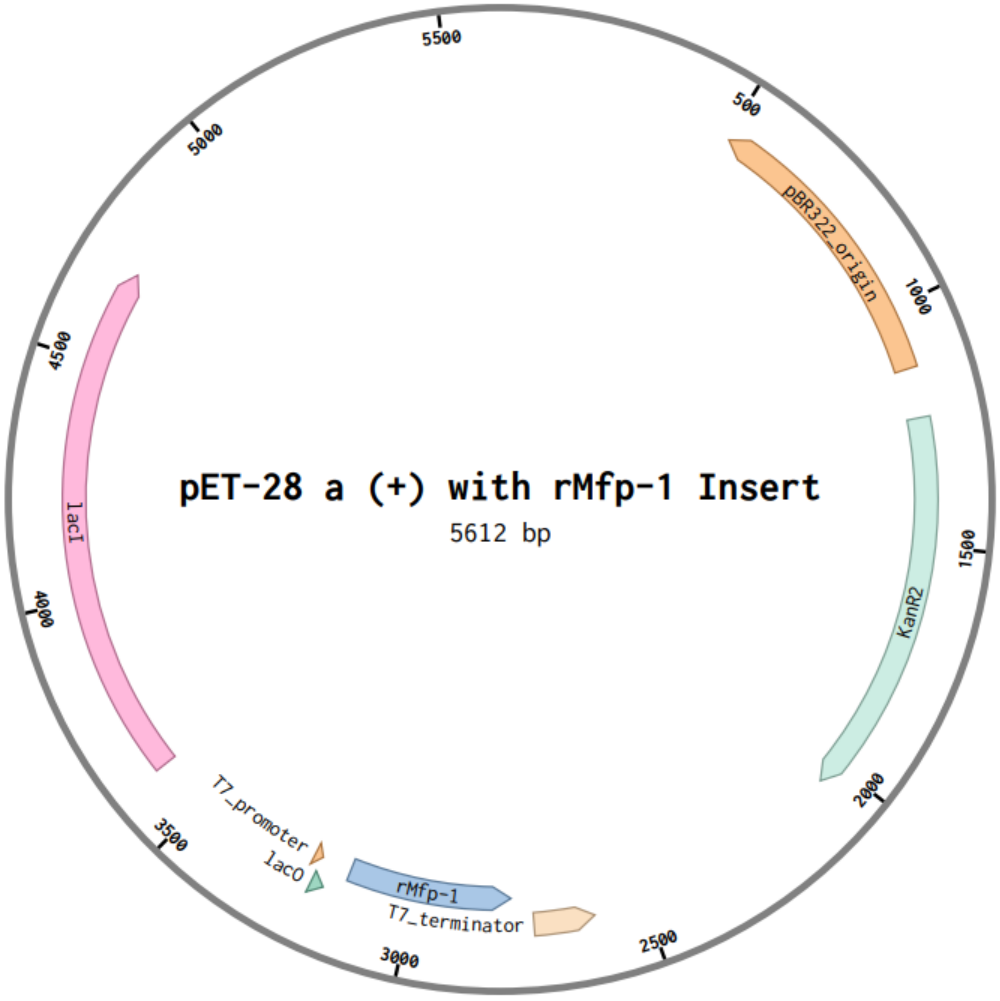

Figure 19: A plasmid map of pET-28a(+) containing the rMfp-1 sequence. Critical sequences are labelled in the plasmid map, such as those for the lac operon, the kanamycin resistance gene, promoters and terminators of transcription, and the origin of replication.

Transformation of DH5α cells was performed using the following protocol. 100 µL of chemically competent *E. coli* DH5α were thawed on ice for 30 minutes. 10 µL of a solution containing 10 µg/mL plasmid was added, and the solution was kept on ice for an additional 30 minutes. After, the cells were placed in water bath at 42 °C for 30 seconds to allow uptake of the plasmid, and immediately placed back on ice for 2 minutes. Afterwards, 1 mL of Luria broth (LB) liquid media was added to the cells and they were incubated for 1 hour in a shaker at 37 °C and 160 rpm. 100 µL of the final mixture was then applied to LB-Agar plates doped with 50 µg/mL kanamycin sulfate. The liquid was spread evenly, the plates were allowed to dry and were placed upside down in an incubator at 37 °C for 18 hours. Successful colonies were selected to undergo colony PCR, as well as Sanger sequencing to confirm the presence of the correct insert.

2.1.2. Colony PCR, Plasmid Purification and Sequencing

Around 10 transformed bacterial colonies were selected for further analysis. Colony PCR was first performed using *Taq* DNA Polymerase (NEB), as it is a rapid and inexpensive method of confirming whether a colony contains the DNA of interest or not. Roughly half of each colony was scraped from the agar plate and was suspended in 50 µL Milli-Q water. The mixtures were placed in 95 °C water for 5 minutes to lyse the bacterial cells and were then centrifuged at 10,000 *g* for 1 minute to pellet any large cellular components that might interfere with the PCR reaction.

From here, the PCR reactions were assembled on ice in 8 x 200 µL PCR strip tubes. The reactions comprised of the following components: 0.13 µL of *Taq* DNA Polymerase (NEB), 2.5 µL of 10X Standard *Taq* Reaction Buffer (NEB), 0.5 µL of 10 mM dNTPs (kindly provided by Dr. Maureen McKeague), 0.5 uL each of 10 µM standard T7 and T7 Terminal primers (ThermoFisher), 5 µL of supernatant of the boiled and centrifuged cells, and 15 µL of HyClone nuclease-free water (ThermoFisher), for a final volume of 25.13 µL. The tubes were transferred to a thermocycler preheated to 95 °C, where the following amplification protocol was utilized: Initial denaturation for 30 seconds at 95 °C, 30 elongation cycles (each comprising of 30 seconds at 95 °C for denaturation, 30 seconds at 41 °C for primer annealing, and 60 seconds at 68 °C for elongation), followed by final extension for 5 minutes at 68 °C. The reactions were stored at 4 °C until further manipulation.

The results of the PCR reactions were studied using agarose gel electrophoresis. Gels were prepared at a concentration of 1.5 % w/v agarose/1X TAE buffer, with TAE buffer comprising of 40 mM Tris, 20 mM acetic acid and 1 mM EDTA adjusted to pH 8.3. The gel mixture was heated until the agarose was dissolved and allowed to cool until the reaction vessel could be comfortably held. The mixture was poured into a casting tray, a comb was inserted to form wells, and was allowed to solidify for 30 minutes. The comb was removed from the gel, and the gel was placed in the electrophoresis chamber, fully submerged by 1X TAE. The PCR samples were thoroughly mixed with 6X Purple Loading Buffer (NEB) and were loaded into the wells. One well per gel was loaded with Quick-Load Purple 1kb Plus DNA Ladder (NEB) as a reference marker. As the gel ran, the purple loading dye separates into pink and blue bands; the gel was run at 80 V until the pink band migrated about 75% of the way through the gel. At this point, the gel was removed from the electrophoresis chamber, and submerged in a plastic container containing 1X SYBR Safe DNA Stain (ThermoFisher) in 1X TAE. The gel was shaken in the dark at 30 RPM and room temperature for 30 minutes and was then imaged using ChemiDoc XRS+ System (BioRad) with the excitation wavelength optimized for SYBR Safe. Lanes containing fragments of lengths around 550 bp indicated a plasmid with an rMfp-1 sequence inside, whereas fragments of mass around 320 bp indicated a plain pET-28a(+) plasmid.

The samples containing 550 bp fragments were prepared for sequencing to confirm the appropriate DNA sequence was present without mutations. To do this, the remaining halves of the colonies were scraped off their plates and suspended in 1 mL of LB liquid media containing 50 µg/mL kanamycin sulfate. The tubes were incubated for 18 hours in a shaker at 37 °C and 160 rpm. 0.5 mL was utilized in a plasmid miniprep to extract purified plasmid DNA using the NEB Monarch Plasmid Miniprep kit. The purified plasmids were submitted for Sanger sequencing using T7 and T7 Terminal primers at the McGill University and Genome Quebec Innovation Center. After sequence verification, plasmid DNA was utilized to transform *E. coli* BL21(DE3) for recombinant expression of the protein. Glycerol stocks were made for both DH5α and BL21 for long term storage. Briefly, glycerol and overnight LB cultures of the respective plasmid-containing bacteria were mixed in a 1:3 ratio, allowed to sit for 30 minutes to permit cellular uptake of glycerol and subsequently flash frozen in liquid nitrogen for long-term storage at -80 °C.

2.1.3. Recombinant Expression

10 L of Terrific Broth (TB) was prepared following the Cold Spring Harbor Protocol. To summarize, Terrific Broth is composed of 12 g/L of tryptone, 24 g/L yeast extract, 4 mL/L glycerol, and 17 mM KH2PO4 and 72 mM K2HPO4. All reagents except the phosphate salts were mixed together and dissolved in approximately 9 L of dH_2O. The phosphate salts were dissolved separately as 10X stocks in dH_2O. The solutions were autoclaved for 15 minutes at 15 PSI and 151 °C and were allowed to cool to room temperature. At this point, the phosphate solution was added to the other components, yielding the final TB solution with the appropriate concentrations. Kanamycin sulfate was added to yield a final concentration of 50 μg/mL.

Meanwhile, an innoculation stock of *E. coli* BL21 was grown to inoculate the 10 L of TB. A sterile pipette tip was firmly scraped along the surface of the Bl21-rMfp-1 -80 °C glycerol stock to collect bacteria. The tip was then swirled in 100 mL of TB containing 50 μg/mL kanamycin sulfate to suspend the bacteria in solution. This solution was then incubated at 37 °C and 160 rpm to grow the *E. coli* until the optical density at 600 nm (OD_{600}) reached a value of 0.6, yielding the ready-to-use innoculation stock.

The OD_{600} = 0.6 stock was used to inoculate the 10 L of TB, yielding a concentration of 10 mL stock / L of TB. The 10 L of inoculated TB was then incubated at 37 °C and 160 rpm until the OD_{600} reached a value of 0.6. With the *E. coli* sufficiently multiplied, the expression of rMfp-1 contained on the bacterial plasmids was induced via the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. The *E. coli* were allowed to express for 5 hours at the same conditions as before. 100 uL aliquots of the reaction mixture were collected every 2 hours after induction to monitor protein expression via SDS-PAGE (**Figure 20**). All SDS-PAGE gels were cast using a 4% acrylamide/bis-acrylamide stacking gel and 12% acrylamide/bis-acrylamide separation gel. For sample preparation, proteins solutions were mixed with water and 4X Loading Buffer (prepared in house, consisting of 50 mM Tris pH 6.8, 2% SDS, 10% glycerol, 1% BME, 12.5 mM EDTA and 0.02% Bromophenol Blue) in a 1:2:1 ratio. The mixes were boiled for 5 minutes to fully denature the proteins, and 20 μL was pipetted into each well, and the gel was run for 45 minutes at 200 V. Gels were then stained in Coomassie Brilliant

Blue R-250 solution (Sigma) for 1 hour, and placed overnight in destaining solution comprising of 50% methanol (%v/v) and 10% acetic acid (%v/v) in water. After expression, the bacterial culture was centrifuged at 5000 *g* and 4 °C for 10 minutes, yielding a wet *E. coli* pellet. The wet weight of the pellet was recorded for use in subsequent processing, and the pellet was stored in a plastic bag at -80 °C until needed.

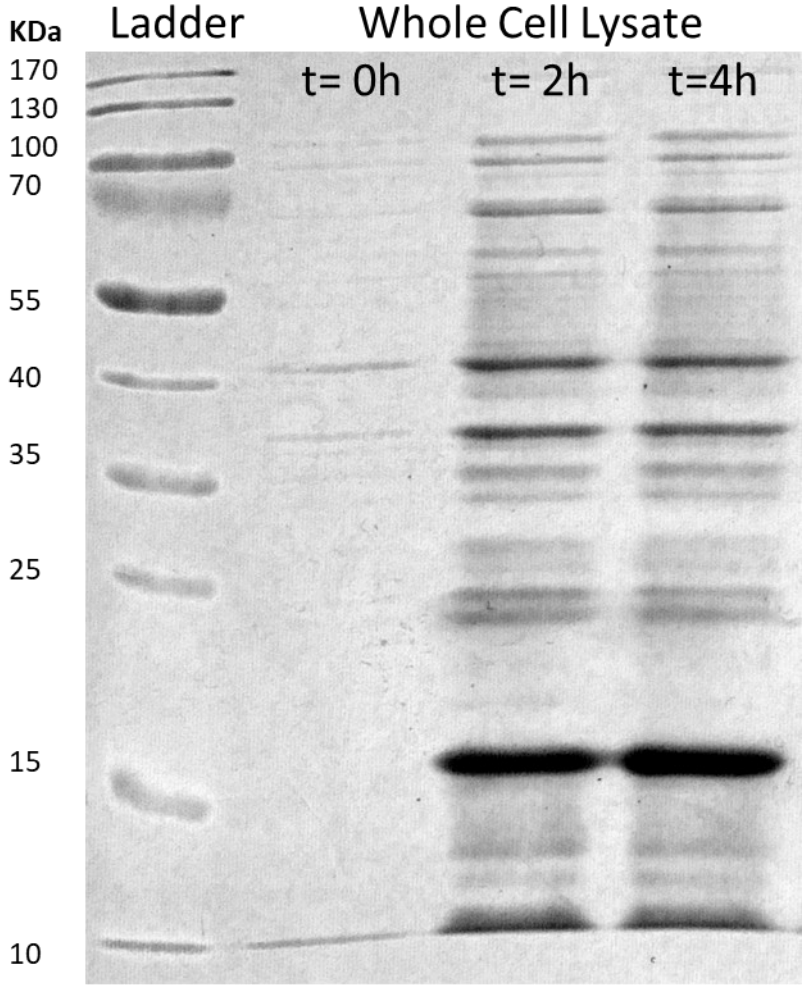

Figure 20: SDS-PAGE gel of rMfp-1 expression process. t = 0h marks the addition of IPTG to induce expression. rMfp-1 can be seen appearing slightly above the 15 kDA ladder protein.

2.1.4. Inclusion Body Purification and Solubilization

Previous literature has found that expression at 37 °C, coupled with the intrinsic properties of the protein sequence results in rMfp-1 being primarily expressed as dense, insoluble aggregates called inclusion bodies.[124] This density and insolubility was exploited to

purify rMfp-1 without the use of any affinity tags. To begin, the bacterial pellet was removed from the freezer, and was forcefully hit with a hammer several times to break it up into smaller pieces. These pieces were quickly transferred to a beaker containing a stir bar before defrosting for subsequent steps. 4 ml of TNE per gram *E. coli* (wet weight) was added to the beaker, which was then stirred to defrost and resuspend the *E. coli* at 37 °C. Meanwhile, a solution of egg lysozyme (BioBasic) was prepared in TNE at a concentration of 10 mg/mL. After the *E. coli* was completely defrosted and resuspended, 80 uL of lysozyme solution was added per gram *E. coli* (wet weight) in solution, and the mixture was incubated for 30 minutes. As the breakdown of the bacterial cell walls progressed, the solution became very viscous due to large amounts of DNA being released. At this point, the entire beaker was placed in the -20 °C freezer for several hours to allow for freezing of the solution. This was done to help further break down the cells and shear some of the DNA. The process was repeated 2 more times starting at the addition of lysozyme solution, as the complete lysis of the cells is critical for the quality of the purification. After this was completed, the solution still contained significant amounts of unlysed cells; to deal with this, the solution was kept on ice and lysis was continued using a probe sonicator (Sonic Dismembrator Model 500, Fisher Scientific) operating at 75% amplitude, at a cycle of 6 seconds on: 4 seconds off for a total of 5 minutes on time. The resulting suspension was centrifuged at 12000 rpm for 20 minutes to pellet all of the insoluble components in solution. The pellet was then resuspended in a refrigerated solution of TNE containing 2% Triton, in order to help solubilize the cell wall and membrane components that pellet with the inclusion bodies. This solution was first homogenized using a Teflon pestle and mortar and was then sonicated at the same settings as before. Centrifugation was repeated, yielding a slightly smaller pellet and a yellow supernatant containing solubilized cellular components. The steps from resuspension of the pellet to centrifugation constitute one “wash” cycle; these washes with TNE-2% Triton were repeated several times, until the supernatant after centrifugation did not undergo a visible colour change. At this point, the solution used to wash the pellet was changed to TNE containing 2M urea, to solubilize any protein contaminants adhering to the surface of the inclusion bodies. The wash cycle was repeated once using TNE-2M urea, then once with plain TNE to remove urea, and lastly, once with Milli-Q water to remove excess salts, yielding a pellet consisting of fairly pure inclusion

bodies that were ready for solubilization. Inclusion body purity was assessed by sampling both supernatant and pellet after washing steps (**Figure 21**). The pellet was resuspended in Milli-Q and homogenized using the Teflon mortar and pestle and sonication and was then transferred to a beaker containing a stir bar. To solubilize the inclusion bodies, glacial acetic acid was added over the span of 30 minutes to yield a final concentration of 20% acetic acid under constant stirring.

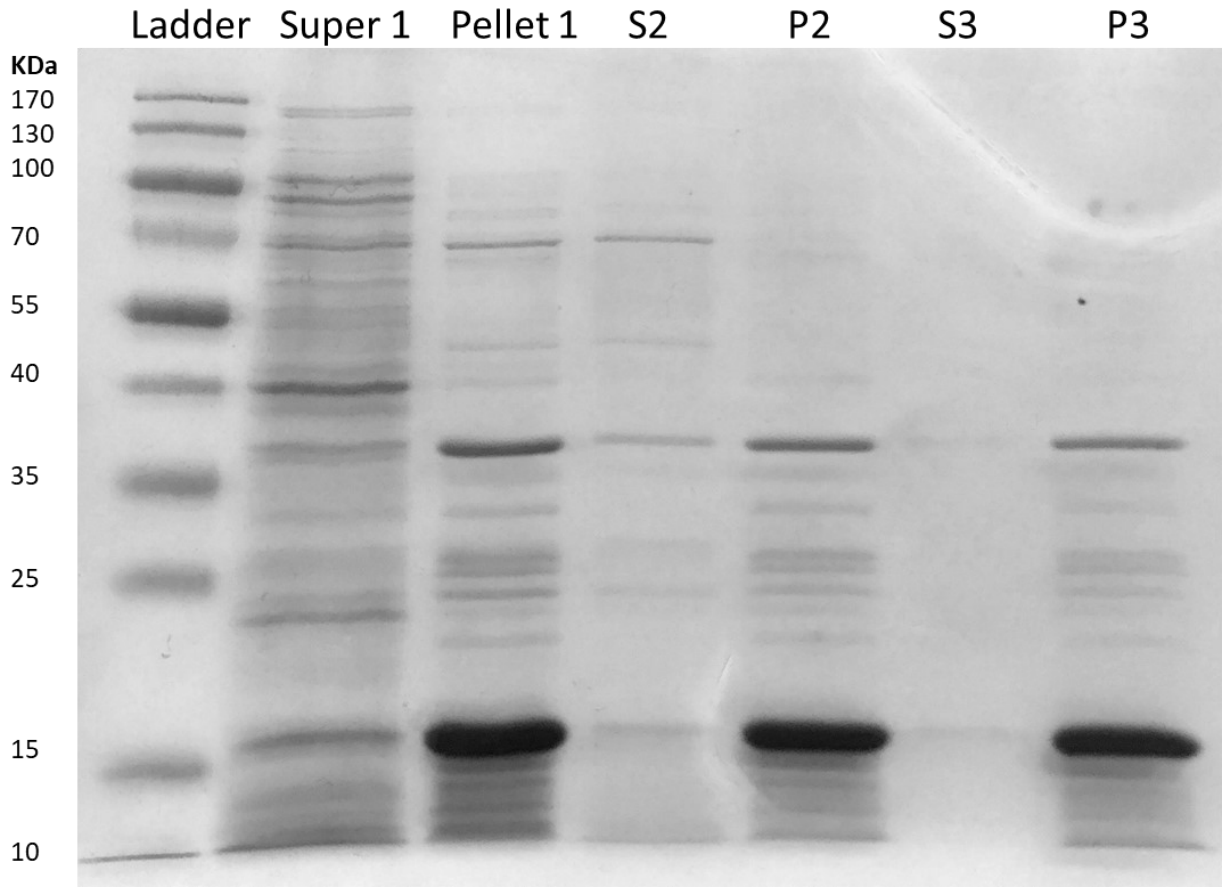

Figure 21: SDS-PAGE gel of the rMfp-1 inclusion bodies purification process. Being in inclusion body form, rMfp-1 is predominantly found in the pellets after centrifugation, and once again found slightly above the 15 kDa ladder protein. The numbers above columns of the gel indicate the washing step after which the aliquots were collected.

The final volume of the solubilization solution was approximately 100 mL, and the solubilization was run for a minimum of 24 hours. This solution was then centrifuged at 12,000 rpm for 20 minutes and supernatant containing rMfp-1 was collected. Typically, a reduction in solution volume to approximately 30-40 mL was necessary to facilitate further manipulations;

this was achieved by centrifuging the solubilization supernatant in batches using an Amicon Ultra filter (15 mL, 10 kDa MWCO) (Sigma). The retentate of this filtrate (containing rMfp-1) was subjected to a final purification using prep-scale HPLC. This was done utilizing a Supelco Discovery BIO Wide Pore C8 column (10 µm particle size, 25 cm x 21.2 mm) with the gradient elution of water and acetonitrile in **Figure 22**.

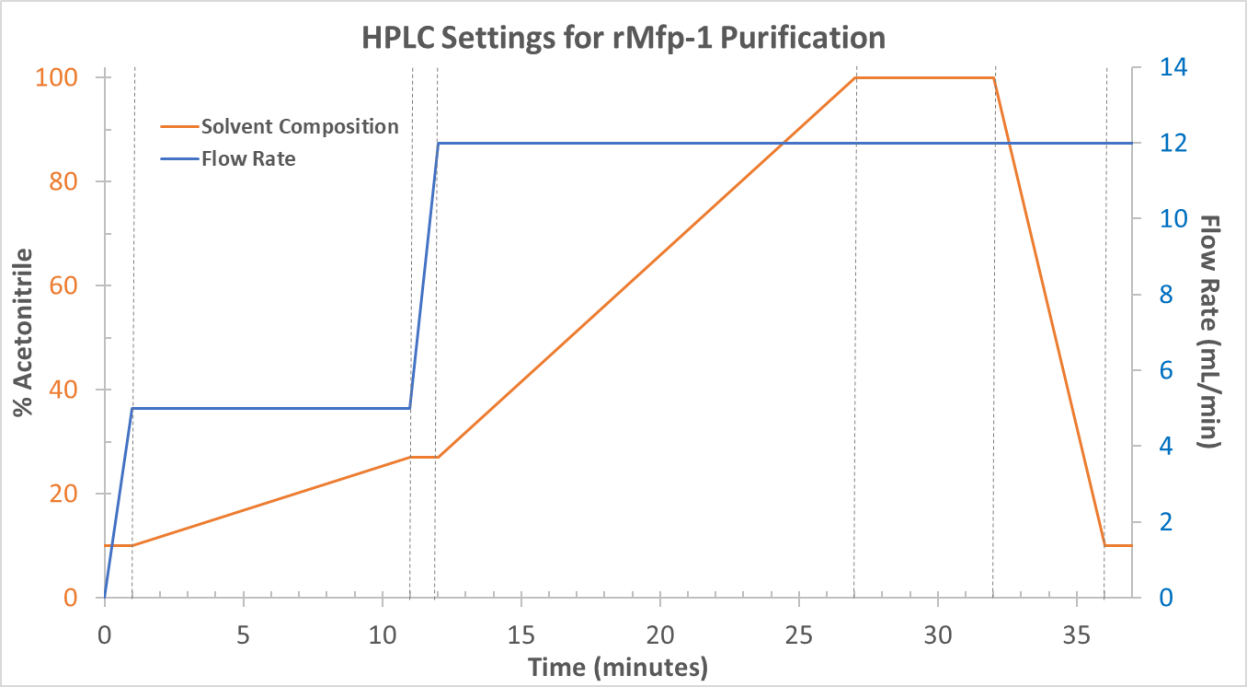

Figure 22: HPLC elution gradient utilized in purification of rMfp-1.

10 mL injections were performed, and the absorbance of the eluent was monitored at 280 nm; as rMfp-1 is the principal component of the solution at this point, a very large spike in absorbance is observed once per run, indicating the fraction(s) containing rMfp-1 (**Figure 23**). These fractions are pooled from all the runs, flash frozen in liquid N_2, and lyophilized for 24 hours to yield a granulated powder that is pure rMfp-1. This powder is stored in an airtight container at -20 °C until required for use. Lastly, these identical steps were also utilized for a rMfp-1 variant containing a terminal cysteine residue.

2.1.5. rMfp-1 Functionalization

Since rMfp-1 does not contain any DOPA residues, and native Mfp-1 does, post-translational modification must be performed to generate an rMfp-1 that can mimic the metal binding capacities of the native form. This is done *in vitro* using mushroom tyrosinase (Sigma). To begin, a buffer consisting of 0.1 M phosphate, 25 mM citric acid and 20 mM borate, at pH 6.8 is prepared. rMfp-1 is removed from the freezer, measured out, and then suspended in the buffer and dissolved at a concentration of 10 mg/mL. Simultaneously, mushroom tyrosinase is dissolved in a separate, but identical volume of buffer as rMfp-1 at a concentration of 250 units / µmol tyrosine that will be present in the final reaction mixture**.**

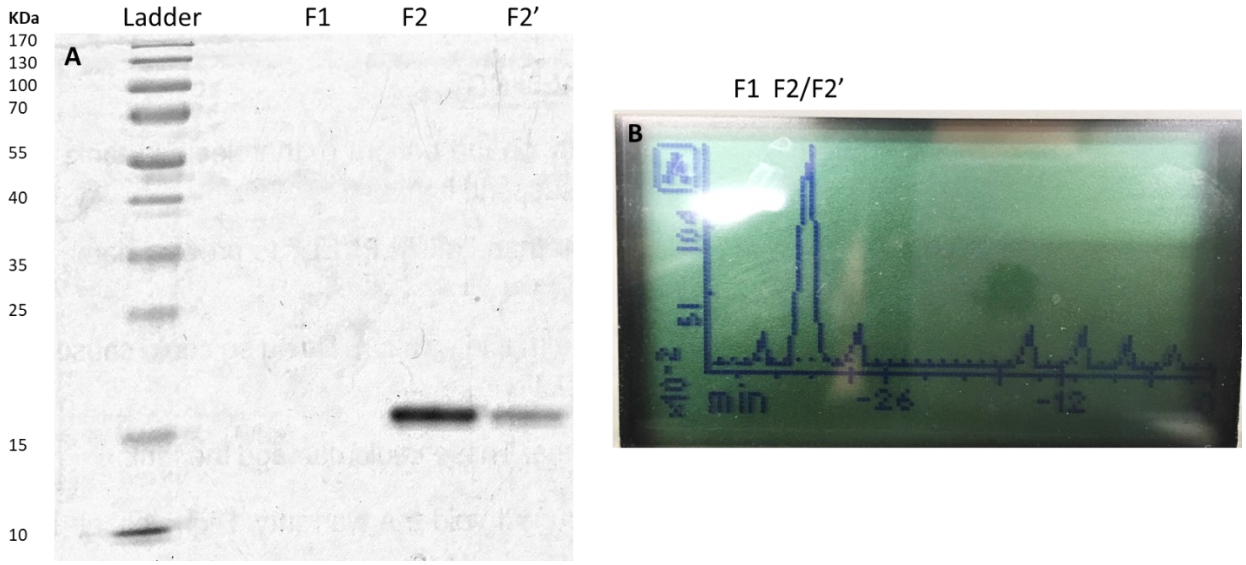

Figure 23: A) SDS-PAGE gel of eluting HPLC column fractions. After this chromatography step, rMfp-1 has > 95% purity. B) Chromatogram of column eluent showing the fractions sampled in the SDS-PAGE gel. As the HPLC system is very old and has a built-in control module, data cannot not be extracted, resigning operators to photos of the LCD.

These two solutions are then combined in equal parts, yielding a solution with rMfp-1 at a concentration of 5 mg/mL and mushroom tyrosinase at 125 units / µmol tyrosine. The reaction is covered in foil and gently shaken at room temperature for 24 hours. After this time has elapsed, 1/5th reaction volume of glacial acetic acid is added to stop the reaction. As the solution now contains both rMfp-1 alongside mushroom tyrosinase, the mixture must be separated using column chromatography. This is done using a gel filtration column (HiLoad 16/600 Superdex 200 pg, GE Life Sciences) running on an FPLC system (BioRad). The reaction mixture is centrifuged at 12,000 *g* for 10 minutes, and the supernatant is loaded into a syringe to be loaded onto the

column. After injection, the proteins are separated via isocratic elution utilizing a mobile phase of 1% acetic acid with 150 mM NaCl. The fractions and the associated absorbance chromatogram are collected, and the samples are run on an SDS-PAGE gel to confirm separation. The fractions containing rMfp-1 were pooled and dialyzed against 1% acetic acid to remove excess salts.

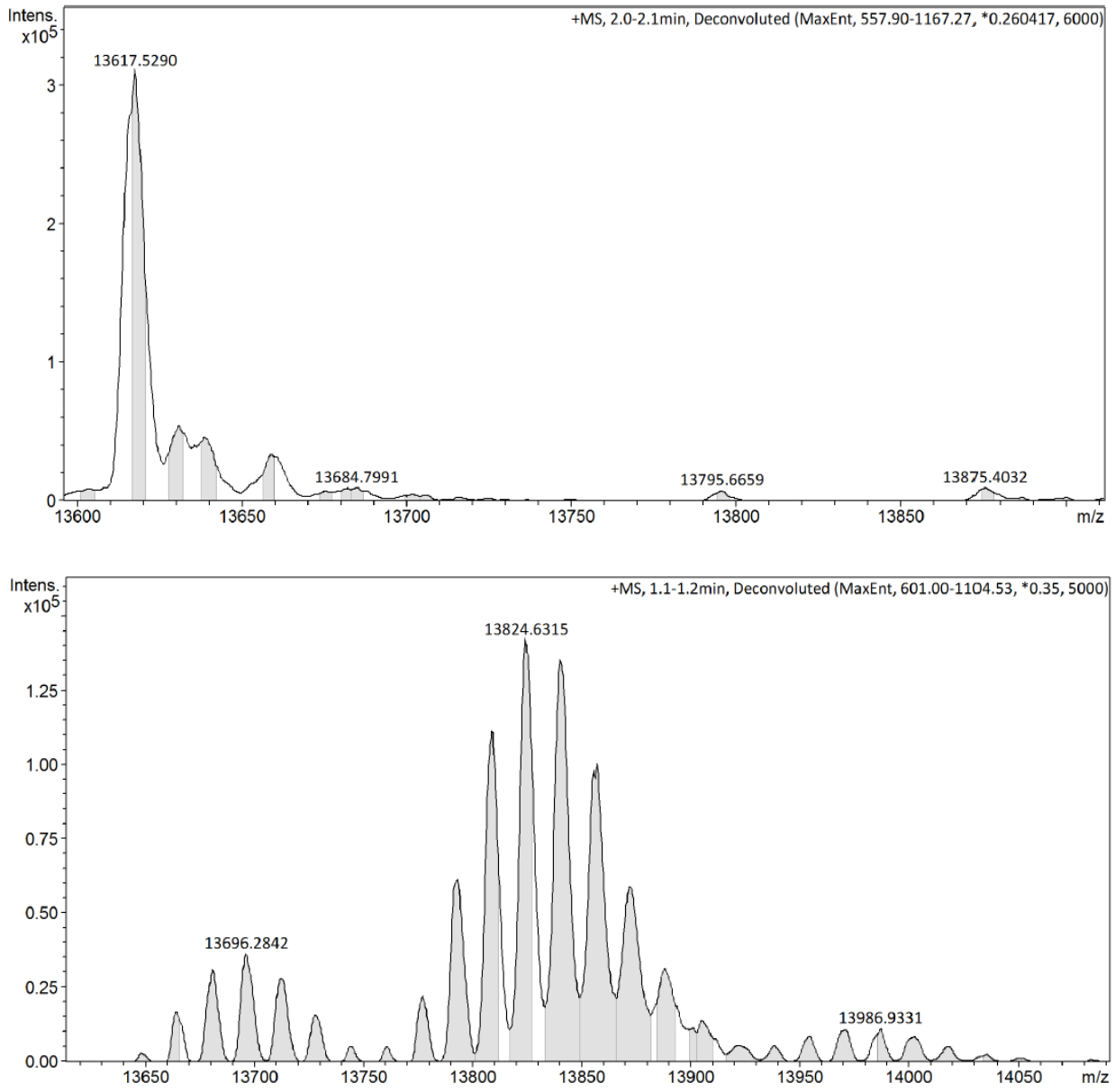

Figure 24: Plots of rMfp-1 before and after treatment with mushroom tyrosinase. The top graph shows a sample of plain rMfp-1 as obtained from recombinant expression in E. coli. The bottom graph shows functionalized rMfp-1-DOPA. The Peak positions are localized according to the formula 13616 + 16n Da, where n is the number of DOPA residues present in a protein.

They were then submitted for high-resolution mass spectrometry (maXis Impact Q-TOF, Bruker Daltonics) to confirm the extent of functionalization with DOPA residues (**Figure 24**). The mol% of DOPA was calculated based on the Q-TOF MS results (**7.2 DOPA Content Calculation**), and the pooled fractions were lyophilized; the resulting solid was stored at -20 °C until required for further experiments.

2.2. UV-Vis Spectroscopy and pH Dependence Studies

2.2.1. Sample Preparation

Stocks of catechol (Sigma), PEG-DOPA (10 kDa 4-arm PEG, with 1 DOPA on each arm, kindly synthesized by Mostafa Rammal) and rMfp-1-DOPA were prepared to a concentration of 3 mM catechol moieties. The pH of these solutions was adjusted to pH 3.0, to inhibit the oxidation of DOPA. Stocks of relevant metal solutions ($FeCl_3$, VCl_3, V^{IV} produced from oxidized VCl_3) were prepared to a concentration of 100 mM.

Working solutions were made by combining 1 µL of metal stock with 100 µL of catechol-containing stock, yielding a 3:1 molar ratio of catechol : metal ion. For Fe^{III} containing solutions, a threefold dilution was performed before measurement owing to the high extinction coefficients of Fe^{III}-catecholato complexes. Base titrations were carried out using additions of NaOH. In these studies, 1 equivalent of NaOH is defined relative to the concentration of the metal ion in solutions (i.e. 2 equivalents of NaOH added to a solution of 1 mM VCl_3 would imply 2 mM of NaOH has been added to the solution). For all experiments, 2 samples were prepared in parallel: one was used for the UV-Vis measurements while the other was used for pH measurements using a micro pH electrode.

2.2.2. Data Acquisition and Analysis

UV-Vis spectroscopy was performed using a dual-beam spectrophotometer (Cary 100 Bio, Agilent) utilizing a small volume fused quartz cuvette. The absorbance was recorded in the range of 375 to 800 nm. Data was collected using the “Scan” app contained in the Cary 100 Bio Suite. Spectra were obtained after every base addition, and all measurements were performed in triplicate.

pH was measured using a pH meter (FiveEasy, Mettler Toledo) equipped with a microelectrode (LE422, Mettler Toledo). Electrode calibration was performed using a 3-point curve pH 4.01, 7.00 and 10.01 standards. Readings were collected in triplicate using automatic mode. Data analysis was performed in Excel. Triplicate absorbance measurements were averaged and had their blank absorbances subtracted. Spectra were then plotted, and using the "equivalents added" parameter, were correlated to the solution pH that was obtained in parallel.

2.3. Resonance Raman Spectroscopy Studies

2.3.1. Sample Preparation

Working samples of rMfp-1-DOPA and V^{III}, VO^{IV}, and Fe^{III} were prepared in a similar manner to the UV-Vis samples described before. Instead of aliquoting samples into a cuvette, 2 µL were spotted on a quartz coverslip and allowed to air dry.

2.3.2. Data Acquisition and Analysis

The dried samples were imaged using a confocal Raman microscope (Alpha 300R, WITec). A 532-nm green laser was used at a power between 1 and 5 mW and focused with a 100× objective (Zeiss, numerical aperture [NA] = 0.9). Spectra were collected with a thermoelectrically cooled CCD detector behind a 600 g/mm grating. Spectra were obtained using an integration time of 5 seconds and 10 acquisitions. Data was collected using WITec ControlFIVE 5.1 software and processed using WITec Plus software to remove cosmic rays, provide background correction and smooth the spectra.

2.4. Electron Paramagnetic Resonance Spectroscopy Studies

2.4.1. Sample Preparation

Working samples of rMfp-1-DOPA and V^{3+} and VO^{2+} were prepared in a similar manner to the UV-Vis samples described above. 10 µL of sample was aliquoted into a 4 mm OD glass EPR tube and flash frozen using liquid N_2. After running the samples in their initial states (pH 3.0), they were defrosted and a small amount of concentrated sodium acetate buffer (pH 5.5) was added to achieve a final concentration of 5 mM. The samples were refrozen to be run again at the raised pH.

2.4.2. Data Acquisition

EPR was performed on the samples using an X-Band EPR Spectrometer (Elexsys E580, Bruker). Spectra were collected on frozen samples using a microwave frequency of 9.33 GHz, a central g value of 2.00 with a 2000 G window, a modulation amplitude of 20 and an attenuation value of 20 dB at a temperature of 210 K. The number of acquisitions varied between 4 to 16 scans, depending on the strength of the signal and the resulting signal-to-noise ratio. Data collection was done using the Bruker Xepr software, and baseline correction was also performed using this software prior to exporting data.

2.4.3. Data Analysis

Acquired EPR data were analyzed using the EasySpin EPR Spectral Simulation Package for MatLab.[125] Spectral assignment was based on collecting EPR parameters of catecholato-vanadium complexes from literature, using them as starting points for simulating a spectrum and then refining the parameters via a fitting algorithm. All simulations were done using the `pepper` function, which is built for simulating continuous wave, field swept EPR of solid systems (such as the frozen solutions in our case). Across all samples, the `pepper` simulation is passed to the `esfit` function, which takes the simulated `pepper` parameters and optimizes them to achieve the best fit. The optimization is done using a Nelder/Mead downhill simplex which minimizes the RMSD between the experimental and simulated spectra by varying the simulation parameters.

One important difference between samples is how many components (i.e. different vanadium species) comprise their total spectrum. When simulating spectra, different species are simulated independently and then linearly combined. As a result, an additional weighting parameter is generated besides parameters that derive from quantum mechanical computations on the spin system.

2.5. Surface Force Apparatus Studies

The normal force-distance profiles and adhesion forces of rMfp-1-DOPA were determined using a SFA. The typical experimental setup and working principle of SFA have been reported previously.[126-128] Briefly, thin back-silvered mica sheets (thickness 1-5 μm) were first glued onto

cylindrical silica disks (radius R = 2 cm). The functionalized rMfp-1-DOPA coatings were prepared by a dip coating method. Two glued mica surfaces were dipped into a prepared rMfp-1-DOPA solution (20 μg/mL in 150 mM acetic acid buffer, pH 3.0) for 30 min. After that, the mica surfaces were thoroughly rinsed by the buffer solution (150 mM acetic acid buffer, pH 3.0) to remove unbound or loosely bound proteins followed by gently blow-drying using nitrogen gas. The protein-coated mica surfaces were then mounted into the SFA chamber in a crossed-cylinder geometry, the interaction of which is equivalent to a sphere of radius R approaching a flat surface when their separation D is much smaller than R based on the ''Derjaguin approximation".[129] Buffer solution conditions for the experiments included: (1) 100 mM sodium acetate + 250 mM potassium nitrate + 1 mM bis-tris buffer, pH 5.5 and (2) 100 mM sodium nitrate + 250 mM potassium nitrate + 1 mM trizma buffer, pH 9. To prevent DOPA from oxidizing at pH 9.0, the whole SFA chamber was evacuated and purged with argon three times followed by being strictly sealed under argon protection. The testing buffer solution containing desired metal ion (*i.e.,* Fe^{3+}, VO^{2+} and V^{3+}) was degassed with argon for 15 mins, and then was carefully injected between two surfaces from the syringe entry port on the front plate. The system was kept equilibrated for 30 min before force measurements. The normal interaction forces F between the two curved surfaces were measured as a function of surface separation distance D. D could be monitored in real time and *in situ* using an optical technique called multiple beam interferometry (MBI) by employing the fringes of equal chromatic order (FECO).[130] The reference distance (D = 0) was defined as the contact of two bare mica surfaces in air. The coating thickness D_T could be determined via the shift of FECO wavelength before and after coating. F was determined according to the Hooke's law.[131] During a typical force measurement, two surfaces were first brought into contact (known as "approach"), were kept in contact for a certain time, and were finally pulled apart (known as "separation"). Cohesion forces are measured when the two attractive surfaces and suddenly jump apart while being separated (known as "jump out"). The measured adhesion F_{ad} is related to the adhesion energy per unit area W_{ad} for two flat surfaces of the same materials based on the Johnson−Kendall−Roberts (JKR) model $F_{ad}/R = 1.5\pi W_{ad}$, which is generally applied for soft deformable surfaces with relatively large curvature and strong adhesion.[127, 132-133]

2.6. TERS Studies

TERS measurements, including sample preparation, were performed by collaborators at the Max Planck Institute of Colloids and Interfaces and the Friedrich-Schiller University in Germany. All data analyses described were performed by myself.

2.6.1. Sample Preparation

Longitudinal sections of byssal cuticle were prepared using a cryo-microtome (HM560, Microm). To begin, a large drop of water was frozen on the sample holder, and a flat area measuring 1 cm^2 was cut using the cryo-microtome. Next, the sample holder was unmounted and brought to a horizontal orientation while maintaining cryo-conditions. A few pieces of thread were laid on the flat area and the area was briefly thawed and refrozen by holding a finger in close proximity, thereby allowing the threads to stick to the flat area. Water was added drop by drop (allowing for complete freezing between subsequent additions) to fully embed the threads. The samples were then cut into 1 µm sections using a knife, at a sample temperature of -20°C and knife angle of 6°. Sections were then thawed in water and picked up using a loop of hair affixed to a steel dissecting needle. From here, they were transferred to a glass slide prepared with Piranha solution using the hair loop. Good adhesion to the surface was achieved after drying the sample.

2.6.2. Data Acquisition

TERS spectra and AFM images of cuticle sections were acquired on a Nanowizard I (JPK Instruments AG) mounted on inverted microscope (Olympus IX70). The following objective conditions were used: oil immersion, 60X, NA = 1,45 (Olympus), mounted on a Piezo (PIFOC, Physik Instrumente) synchronizing the z-movement of the TERS tip and the laser focus, a confocal Raman spectrometer (LabRam HR, Horiba Jobin Yvon) with a liquid nitrogen-cooled charge-coupled device camera (ISA Spectrum One, Horiba Jobin Yvon), and a 530.9 nm Krypton ion laser (Innova 300c, Coherent) at a power of 880 µW as the excitation source. Each TERS spectrum was accumulated for 10 s. Spectra were collected from the cuticle section surface with a lateral distance step that varied from 0.5 to 10 nm. The lateral steps were controlled by an additional

100 X 100 µm sample scanning stage (P-734, Physik Instrumente). To confirm that the TERS tip was not contaminated during the acquisition of spectra of the cuticle surface, a reference spectrum from the glass slide was recorded at the end of each measurement.

2.6.3. Data Analysis

Data was obtained in a series of text documents containing wavenumbers alongside their respective signal intensities. Each spectrum from each grid was contained in its own text document, with hundreds of text documents comprising the data for a singular grid region. The vast amounts of data and its formatting required a lot of pre-processing prior to analysis, for which several MatLab scripts were designed. All of the raw spectra were first imported and subjected to such pre-processing scripts. These scripts will be described below and will then be referenced in subsequent sections as they were utilized.

2.6.3.1. Pre-Processing: Baseline Corrections

Baseline corrections were performed utilizing an Asymmetric Least Squares algorithm (ASLS).[134] In brief, ASLS requires an input of spectral data, alongside 2 parameters, λ and p. λ is the smoothing parameter typically ranging from 1 x 10^2-10^9; the larger the magnitude of λ, the smoother the estimated baseline will be. p is the asymmetry parameter, spanning $0 < p < 1$; the value of p is the weighting assigned to positive residuals (typically a small value < 0.01), whereas $1 - p$ is the weighting assigned to negative residuals from the calculated baseline. This asymmetric weighing of residuals proceeds iteratively until the calculated baseline minimizes the negative residuals.[135] The resulting estimated baseline is finally subtracted from the initial data, yielding a corrected spectrum.

Primary ASLS was performed on all spectra using values of λ = 1000 and p = 0.01, for removal of large baseline deflections arising from fluorescence of the tissue samples. Secondary ASLS was also performed, utilizing values of λ = 10,000 and p = 0.1, to remove a broad silicon overtone peak.

2.6.3.2. Pre-Processing: Silicon Peak Removal

All TERS spectra contain a strong and fairly broad silicon background peak centered at 520 cm^{-1} which can be confounding to certain methods of data analysis. Removal of this feature for datasets requiring it was done by setting the intensity of the spectral region of 500 – 540 cm^{-1} to zero.[117]

2.6.3.3. Pre-Processing: Normalization

Two main methods of normalization were utilized. For datasets that did not have their silicon peak removed, the magnitude of this peak was used to normalize all intensity values in the spectrum. For datasets that no longer had a silicon peak, row normalization was used. Row normalization is the process of summing of all intensity values in a spectrum and dividing each intensity by this value, such that the final total sum of intensity values is 1.

2.6.3.4. Pre-Processing: Smoothing

Peak fitting on data sets was hindered by random noise present throughout the spectra. To overcome this, a gentle smoothing was utilized using a built-in MatLab Savitsky-Golay function, `sgolayfilt.` The smoothing polynomial order chosen was 2 (quadratic), with a framelength of 13 data points. These values were chosen based on the least amount of smoothing required, while permitting the accurate fitting of peaks.

2.6.3.5. Pre-Processing: Interpolation

Peak fitting was also hampered by the sparsity of spectral data; due to instrumental limitations, intensity values were only recorded roughly every 4 cm^{-1}. As a result, peaks requiring a fit would sometimes have as few 2 or 3 points available, which is insufficient for accurate fits.

To remedy this, spectra were interpolated using `griddedInterpolant`, a built-in MatLab function. From the choices of interpolation methods, 'makima' was chosen, owing to its ability to preserve the original peak shapes without introducing unnecessary undulations, and its intrinsic overshoot prevention. Spectral traces were interpolated from 512 datapoints to 2048.

2.6.3.6. Analysis using Principal Component Analysis

Data for Principal Component Analysis (PCA), was pre-processed using Primary and Secondary ASLS, followed by silicon peak removal and row normalization.

The MatLab function `pcacov` was utilized to conduct PCA, as opposed to the more standard `pca`, (TERS grids smaller than 22 x 22 resulted in `pca` assuming the larger matrix dimension was always the observations; this assumption is incorrect for smaller grids TERS, thus requiring the use of a different function). To utilize `pcacov`, the data was arranged in a *n* x *p* matrix, where each row *n* represented a spectrum in a grid (aka 1 observation), and each column *p* contained the intensity of Raman scattering for each wavenumber (aka 1 variable). Next, the mean of each variable was calculated, and was subtracted from every observation in the variable. Having had the means subtracted from its variables, the *n* x *p* matrix had its a *p* x *p* covariance matrix calculated using the `cov` function. Finally, it was this covariance matrix that was used as input for the `pcacov` function to conduct the PCA on the dataset.

`pcacov` returns 3 variables that are used to examine the data: *COEFF, latent,* and *explained*. *COEFF* is a *p* x *p* matrix containing the principal component coefficients (aka loadings). Plotting columns of *COEFF* against the original wavenumbers yields plots of the principal components comprising the data. *latent* is a vector that contains the variances of principal components and *explained* is a vector containing the percentage of total variance explained per principal component. These parameters can be plotted against the number of principal components to generate scree plots. Together, plots of principal component loadings alongside scree plots were examined to reveal relationships between variables (i.e. peaks at different wavenumbers that tend to appear together).

2.6.3.7. Analysis using Area Under the Curve

Data for Area Under the Curve (AUC), was pre-processed using Primary and Secondary ASLS, followed by silicon peak removal and row normalization. For this analysis, 6 separate regions of the TERS spectra were selected for integration of AUC: 400-480 cm^{-1} for putative cysteine-iron interactions, 650-700 cm^{-1} and 745-800 cm^{-1} for C-S stretches, 830-850 cm^{-1} for tyrosine ring stretches, 995-1010 cm^{-1} for phenylalanine and 1600-1615 cm^{-1} for general aromatic signals. The centroid and size of these regions were chosen based on previous publications and

simulations of TERS spectra.[136] AUC integration was then performed on the spectral regions of interest using the MatLab function `trapz` for every spectrum in a grid. The integrated AUCs per grid per region were then plotted using the MatLab function `boxplot.` Boxplots of regions of interest were then visually compared across grids for defining trends across different parts of cuticle morphology. AUC data was also utilized in heatmaps (whose generation is described in a subsequent section).

2.6.3.8. Analysis using Peak Fits

Data for Peak Fits was pre-processed using Primary and Secondary ASLS, followed by normalization to the silicon peak, Savitsky-Golay smoothing and interpolation. An in-house automatic peak fitting algorithm was designed as the built-in MatLab functions were not reliably selecting appropriate peaks. First, a fitting range was provided to the algorithm. Secondly, the algorithm attempted to fit the range with a summation of 1 to 4 Gaussian peak shapes. Third, the algorithm examined the R^2 value of the Gaussian summations; the first summation with an $R^2 > 0.99$ was selected for further use. If one was not found, the criterion was lowered to the first summation with $R^2 > 0.95$, and if found, it was chosen with a note being made in an output text file. If an appropriate fit was not found even with the lowered criterion, the grid position was discarded and treated as though there were either no peaks in the region, or there were too many peak shapes to make a confident assignment. Fourth, the chosen Gaussian summation had the parameters of all its Gaussians (*a, b, c*) extracted, which were then subjected to checks. The *a* values (corresponding to Gaussian height) were checked against a threshold value (in this case, the average signal intensity for the entire spectrum) to ensure the peaks was large enough. The *b* values (corresponding to the center of the Gaussian peak) were checked to make sure they fell within the original fitting range. Failing any of these checks resulted in that Gaussian fit being deleted. Fifth, if there was Gaussian fit remaining that meets all check criteria, a final fit of a singular Gaussian was done on the narrowed down *b* region. This was done because the initial Gaussian summation excels at finding accurate *b* values, whereas the *a* values of individual Gaussian tend to be slightly underestimated. The final Gaussian fit was then written to a master array of successfully fitted peaks, alongside its position in the TERS grid. The algorithm repeated

for each spectrum in the TERS grid, and the resultant data was passed to the Heatmap algorithm to visualize the localization of specific TERS signals in specific regions of cuticle morphology.

2.6.3.9. Analysis using Heatmaps

Heatmaps required generation of a numerical grid, subsequent assignment of data to the grid and generation of a visual heatmap. Numerical grids were generated using for loops, taking the x and y dimensions of the TERS grid as input and outputting an empty array of the same dimension. Next, data (AUC or peak intensity) from each spectrum was assigned to the same position in the numerical grid (i.e. spectrum 3 in a TERS grid was assigned to position 3,1 in the numerical grid).

To eventually visualize this data, a color gradient correlating to the magnitude of the data contained in a numerical grid position was required. This gradient was generated using the MatLab `colormap` function and the *jet* colour scheme. The gradient consisted of 255 individual points from one colour extreme to the other. After colour gradient generation, the highest and lowest data values in the grid were extracted and set to correspond to index 255 and 0 of the colour gradient. A column vector of 255 equally spaced points was then generated between the two data extremes using the `linspace` function. Iteration through the grid data was done and each data point was matched to the closest value (and corresponding index number) in the previously generated column vector; the index value (ranging from 0-255) now attached to each data point specified a colour in the colour gradient that is linearly correlated to magnitude of the data value. The visual grid was finally generated using the `scatter` function, utilizing the numerical grid and the assigned colour for each position. These heatmaps were then visually examined against the AFM images of the TERS grid to detect trends across different areas of cuticle morphology.

3. Results

3.1. Tip-enhanced Raman spectroscopy (TERS)

TERS measurements of byssus cuticle sections were performed previously by collaborators at the Max Planck Institute of Colloids and Interfaces and the Friedrich Schiller University Jena, Germany. However, the spectral data acquired were never analyzed. Here, I re-examined this data with the specific goal of identifying compositional differences between the granule and matrix regions of the cuticle. Given the immense quantity of data and the well-known tendency of TERS spectra to show only specific vibrational bands of parts of molecules, rather than the average spectra acquired from bulk Raman measurements, data analysis required development of specific statistical approaches for extraction of meaningful data.

Analysis began by attempting principal component analysis (PCA) on TERS grids. PCA becomes very useful when only 2 or 3 principal components (PCs) can describe a large amount of variance across spectra (i.e. > 50%). This has been demonstrated successfully in a publication seeking to discriminate between RNase A and B proteins on gold nanoplates, with the best principal components explaining 42% of variance across spectra.[117] The method of choice to visualize how much variance is being explained by individual PCs is through a scree plot (**Figure 25A**). The scree plot below is representative of all TERS grids studied, with no individual component describing more than 10% of the variance across spectra in a grid. As can be seen in the orange trace, approximately 7-8 PCs are necessary to account for 40% of variance, whereas only 1 PC can describe the same proportion of variance in certain literature. When a few PCs account for a large proportion of variance, plotting these PCs can yield some highly anti-correlated spectra that are diagnostic of the differing protein signals present in a sample. In our case (**Figure 25B**), many PCs have large amounts of overlap that inhibit that diagnostic power of this technique. While PCA has been successful in highly controlled scenarios containing few proteins or a singular protein with different morphology, this method seems to struggle to simplify the data from our tissue sections. Overall, the complexity of our samples (containing a multitude of proteins, many of which are unknown or incompletely characterized to this day) hindered the ability of PCA to simplify our data in meaningful ways, prompting us to try alternative methods.

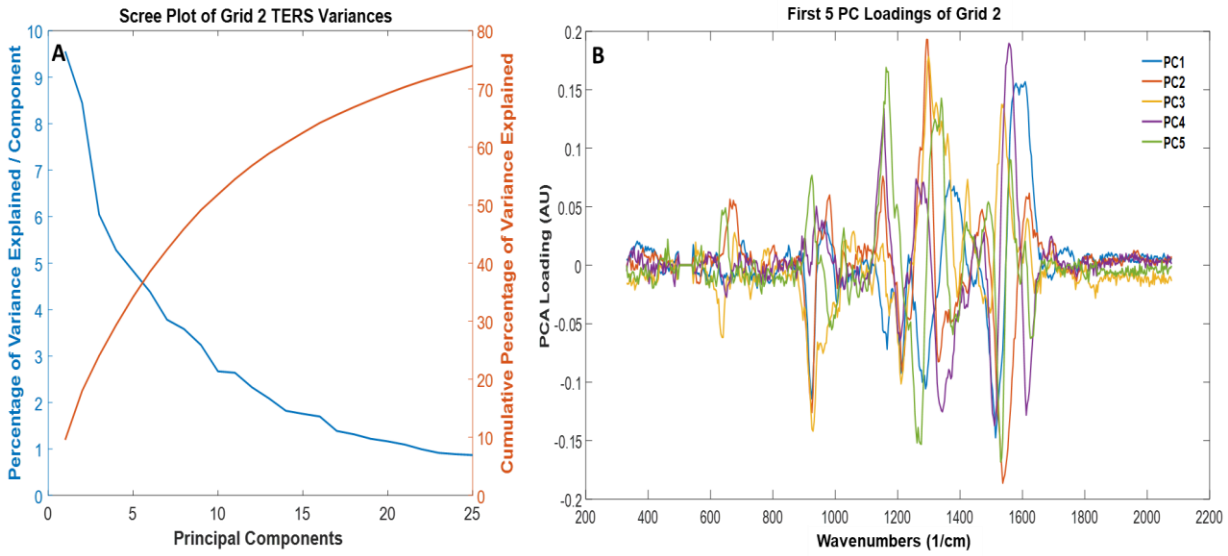

Figure 25: A sample of PCA analysis performed on TERS grids. A) Scree plot of a sample TERS grid demonstrating the variance explained by the first 25 calculated principal components. B) A plot of the first 5 principal components describing the Grid 2 dataset.

Another method that has found success in TERS analysis in Area Under the Curve (AUC) analysis. By selecting spectral regions of interest, and integrating the AUC across every spectrum in a grid, one can visualize differences between grids.[117] Chosen spectral regions included putative Fe-S interactions (400-480 cm^{-1}), C-S bonds (650-700 and 745-806 cm^{-1}), tyrosine signals (830-850 cm^{-1}), phenylalanine signals (995-1010 cm^{-1}) and aromatic amino acid signals (1600-1615 cm^{-1}) – as these peaks might be expected based on prior knowledge of the biochemical features of the proteins in the cuticle.[136] We envisioned for this to be particularly useful by characterizing "granule-only" and "matrix-only" TERS grids and studying the difference in AUC between the samples. Using the aforementioned spectral regions, one could hopefully gain insight about compositional differences between granule and matrix. One major roadblock in this approach was that when the TERS spectra were acquired several years ago (by a collaborator and previous student in the Harrington Lab), many were taken of "granule-only" regions (**Figure 26**) and none were taken of "matrix-only" regions. Without a strong contrasting comparison of "matrix-only" regions, AUC also struggled to gain traction and provide meaningful data. As with PCA, the highly complex nature of a tissue sections makes confident analysis challenging. Another complicating characteristic that becomes apparent in AUC analysis is the stochastic nature of TERS measurements, which naturally generates highly variable amounts of outliers for every

single spectral range examined. From here, we decided to move to a novel, highly targeted approach for compositional differences we were already aware of.

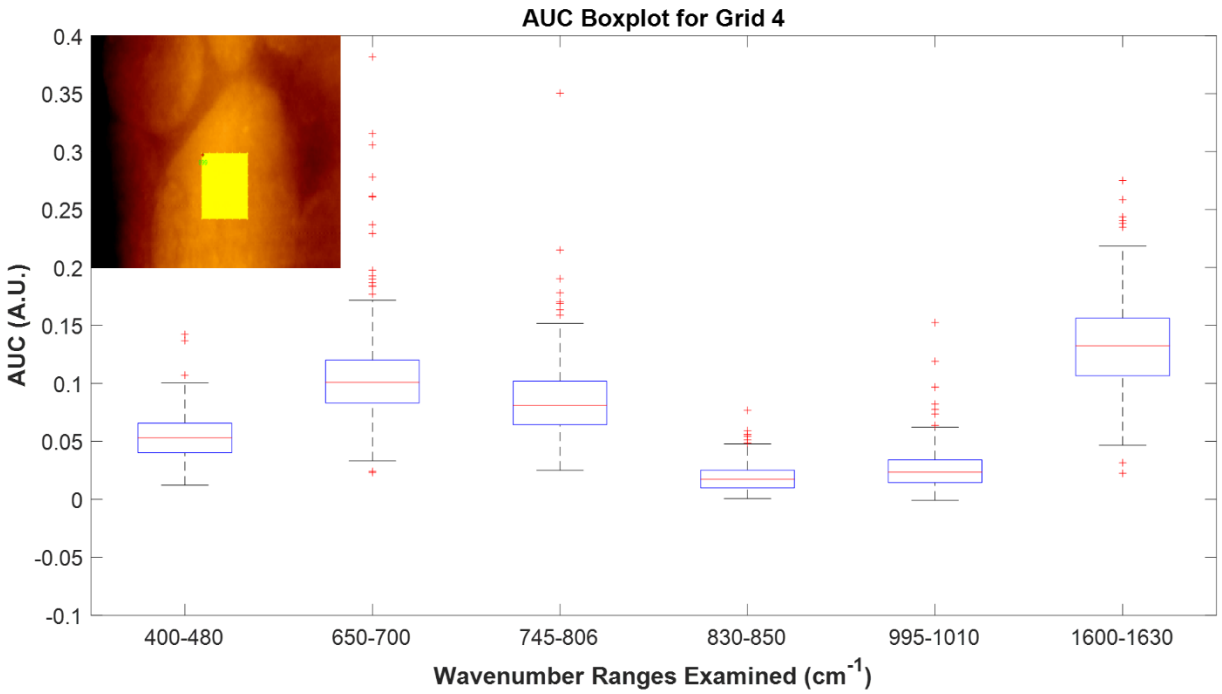

Figure 26: An Area Under the Curve Boxplot for selected spectral regions of interest. The blue boxes represent the interquartile range (IQR), the red line the median and the black whiskers represent data with a maximum distance of 1.5 x IQR outside the IQR. The red crosses represent outliers beyond the whisker range. The inset shows the AFM image with the overlaid TERS grid in yellow.

To target our TERS analysis more specifically, we turned to transcriptomic evidence which has found that the entire sequence of Mfp-1, the principal protein constituent of the cuticle, does not contain a single tryptophan residue.[49] However, several of the putative cysteine-rich cuticle proteins (mfp-17 and mfp-19) each contain 4 mol% tryptophan.[49] This is a valuable realization, as tryptophan is typically a highly Raman active residue whose C=C indole ring stretch signal can be found around 1550 cm^{-1}.[137] Such a signal has been detected in our group via Raman spectroscopy of cuticle vesicles purified from the mussel foot, supporting it being a viable target to use in further TERS analysis (**Figure 27**)**.** This area of protein Raman spectra is typically quite silent, making assignment simpler. If a tryptophan residue is observed, then it is positive confirmation of a protein other than Mfp-1, and based on our current understanding of the cuticle, these tryptophan signals should be localized preferentially in the cuticle matrix.

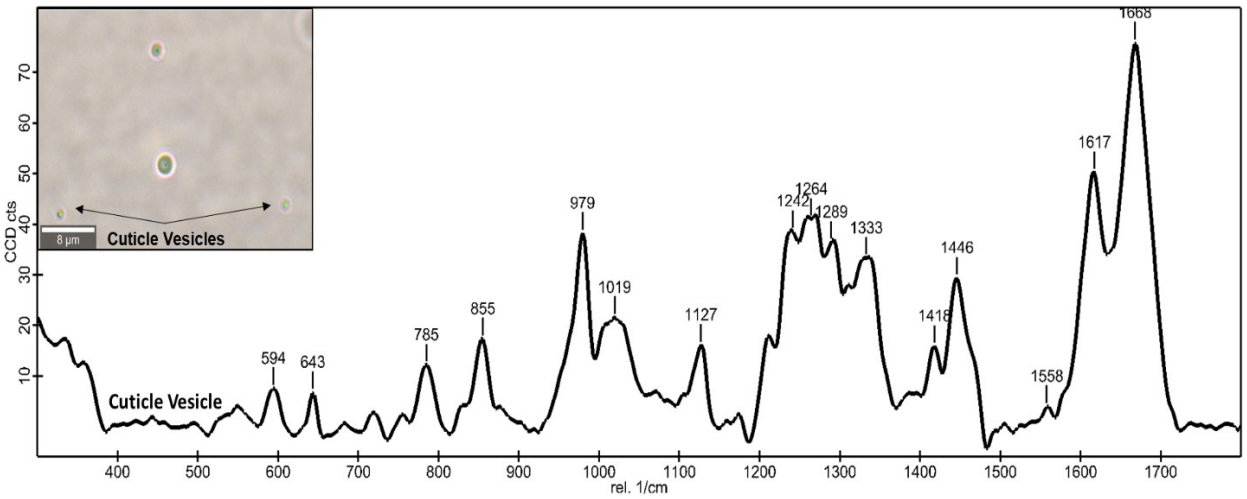

Figure 27: A Raman spectrum of a cuticle vesicle purified from the foot of Mytilus edulis. Of particular interest in the peak labelled at 1558 cm^{-1}, indicating a C-C stretch of the tryptophan indole ring.

By seeing which spectra in a grid possess a tryptophan signal, we can assign them to a heatmap, thereby preserving valuable spatial information about where signals come from in our sample. **Figure 28** illustrates this process and provides the final heatmap for Grid 4 (**Figure 28C**). What comes as a surprise for this grid and all others is that the tryptophan signal does not appear as expected. As Grid 4 is exclusively granule, one would expect little tryptophan signal, however, we see many unexpected hotspots throughout the grid. Even grids that encompass both granule and matrix do not show any noticeable correlation in the location of tryptophan signals. Manipulating parameters such as the fitting range, or the intensity criteria for tryptophan peak detection did not produce marked improvements. Overall, the lack of consistent and tangible spectral differences between different areas of cuticle morphology cause complications signals having been found ranging from 1484 to 1608 cm^{-1},[136] and the common appearance of 2 to 3 peaks in the fitting range, which under normal Raman conditions should only have a 1550 cm^{-1} tryptophan signal. At the surface level, TERS appears to be an attractive avenue with which to explore cuticle composition; however, there were some major challenges given the complex nature of the sample. It is certainly possible that as the technique continues to mature, the scope of samples suited to the method will expand. Under the time constraints of a Masters thesis and the current scope of the TERS field, the complexity of the cuticle tissue sections and the unknowns in cuticle

composition proved a difficult barrier to overcome for new and existing methods of TERS data analysis.

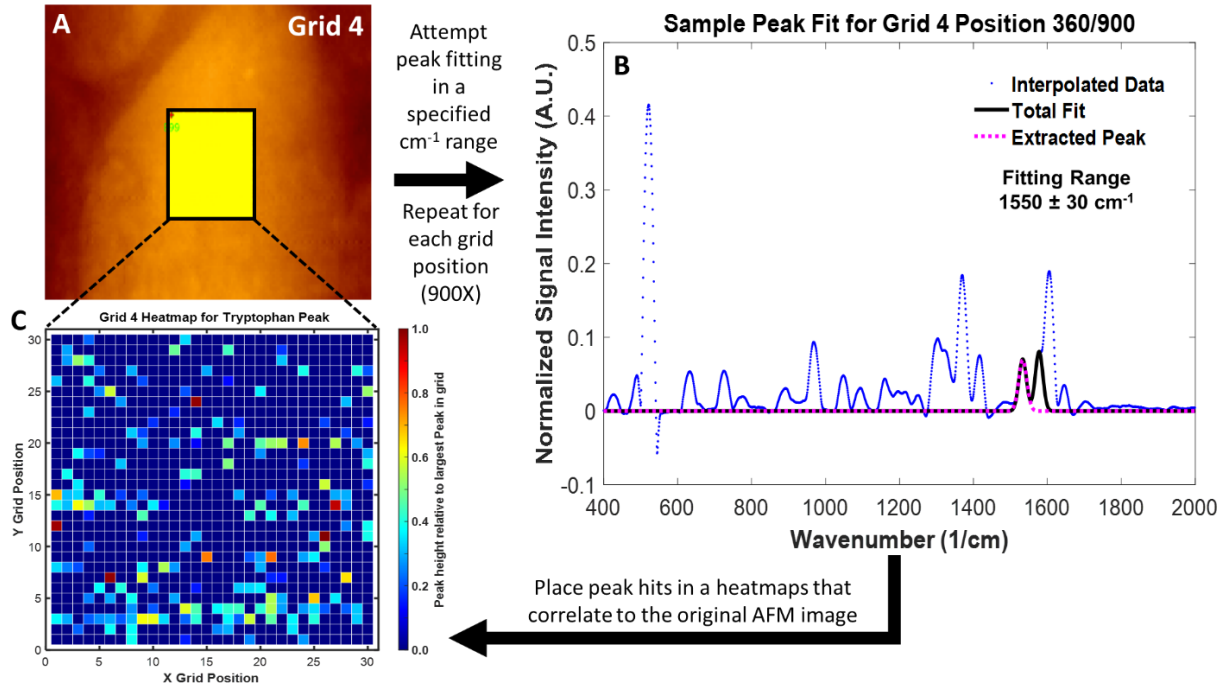

Figure 28: A workflow describing the process of peak fitting and assignment to a heatmap. A) An AFM image of a section of cuticle, with the yellow overlay showing the grid area in which TERS spectra are collected. B) The spectra from one grid position, showcasing the automatic fitting process in which spectra are analyzed and peaks are extracted from. C) The final heatmap from analyzing Grid 4. "Hits" are normalized to the most intense peak detected in the fitting range region. The darkest blue squares in the heatmap indicate an absence of a peak or peaks that did not match the provided criteria.

3.2. Recombinant expression and DOPA functionalization of truncated rMFP-1

Shifting from the preceding *in situ* spectroscopic studies of the native thread material, the remainder of the thesis will focus on *in vitro* studies investigating the nature of metal binding of recombinant Mfp-1 (rMfp-1) with Fe and V. Following the experimental protocol described above, recombinant *E. coli* grown in 10L of Terrific Broth typically yielded a 25 g pellet (wet weight) after centrifugation. Solubilizing the rMfp-1 inclusion bodies from the pellet using 20% acetic acid yielded between 400 – 500 mg of recombinant protein after lyophilization. Subsequent functionalization with mushroom tyrosinase following the described protocol typically yielded percentage conversion rates of tyrosine to DOPA between 50 – 60 %.

3.3. Surface Force Apparatus (SFA)

Mechanical testing using SFA was performed to explore the cohesive properties of rMfp-1-DOPA in the presence of different metal ions under various buffering conditions, in order to provide some initial insight on how iron and vanadium interact differently with this protein. In the following figures, F_{ad}/R is a unit used to describe the cohesion of rMfp-1-DOPA films (pull-off force (F) normalized to R, the radius of the circular mica plates on which rMfp-1-DOPA is coated on). An immediately striking observation is that V^{III}-mediated cohesion forces (**Figure 29B**) are roughly two-fold higher than Fe^{III}-mediated cohesion (**Figure 29A**) under the same buffering conditions (pH 5.5) given the same incubation and contact times. A possible explanation for this could be the intrinsic (albeit less than an order of magnitude) difference in binding affinity that catechols have for Fe^{III} versus V^{III}.[66]

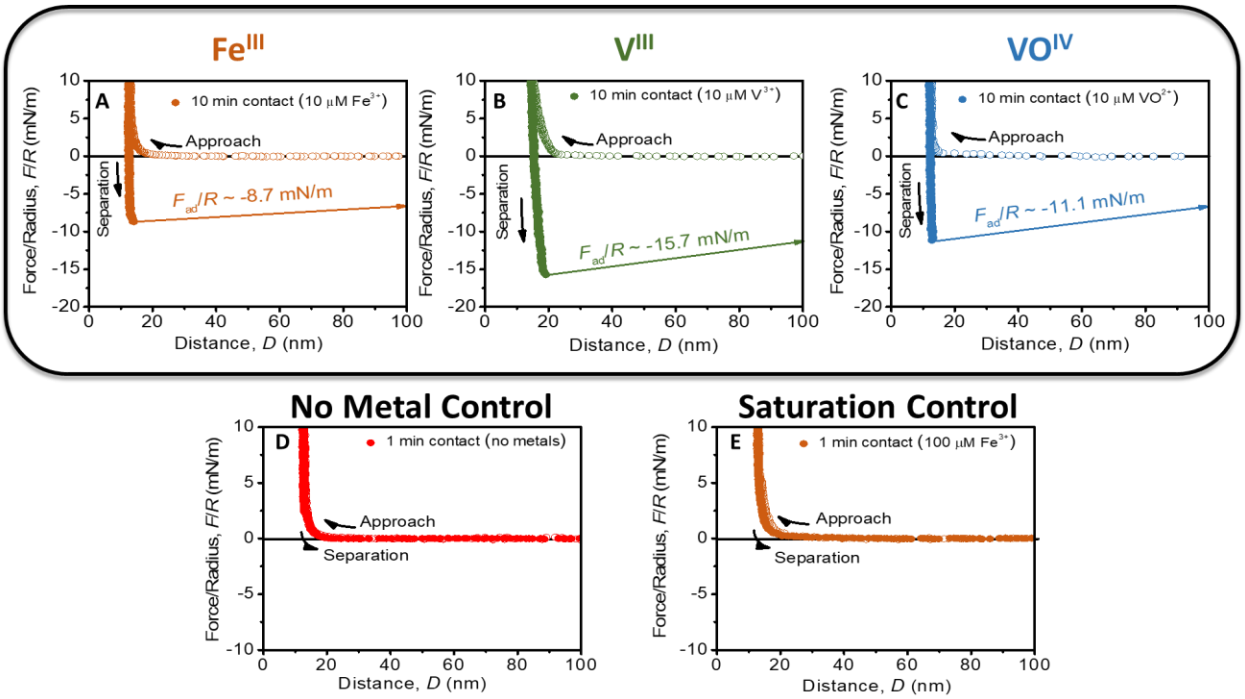

Figure 29: Metal-mediated bridging adhesion in symmetric rMfp-1-DOPA films on mica at pH 5.5. Mica plates are dip-coated in a 20 µg/mL rMfp-1-DOPA solution in 150 mM AcOH for 30 minutes, and subsequently incubated with metal solutions for 30 minutes . A) 10 minute contact following incubation with 10 µM Fe^{III}. B) 10 minute contact following incubation with 10 µM V^{III}. C) 10 minute contact following incubation with 10 µM VO^{IV}. D) 1 minute contact in the absence of metals. E) 1 minute contact for incubation with 100 µM Fe^{III} (enough to provide a metal ion for every DOPA residue in the sample).

Between the two extremes of cohesion strength lies VO^{IV} (**Figure 29C**), with an intermediate value for cohesion strength; in the following sections, UV-Vis and EPR measurements will be used to provide further explanation for this observation. In the absence of added metal ions (**Figure 29D**), and in the overabundance of metal ions relative to DOPA residues (**Figure 29E**), no cohesion between rMfp-1-DOPA films can be detected. It is interesting to note that this difference in adhesion force persists even at higher pH values. SFA data was also obtained at pH 9, which mimics the pH range at which Mfp-1 is functions at in nature. At this high pH, V^{III} (**Figure 30B)** maintains its mechanical superiority over Fe^{III} (**Figure 30A**) and VO^{IV} (**Figure 30C**) in terms of the cohesive forces produced. Importantly, these experiments were performed in an oxygen-free environment to prevent the oxidation of DOPA-catechol in rMfp-1-DOPA to DOPA quinone, which could lead to oxidative cross-linking of the protein films. Thus, we can conclude that the increased force arises from the formation of DOPA-metal coordination cross-links between the layers. These SFA results, including trials having different contact times, are succinctly summarized in the bar plot in **Figure 31,** in which the data is also calculated in terms of adhesion energy. It can be seen that V^{III}-mediated cohesion dominates across both pH levels and all contact times. The adhesion energies at pH 5.5 for rMfp-1-DOPA and V^{III} ($W_{ad} \approx 7.6$ mJ/m^2) are significantly larger than most other mussel protein interactions tested using symmetrical films in SFA.

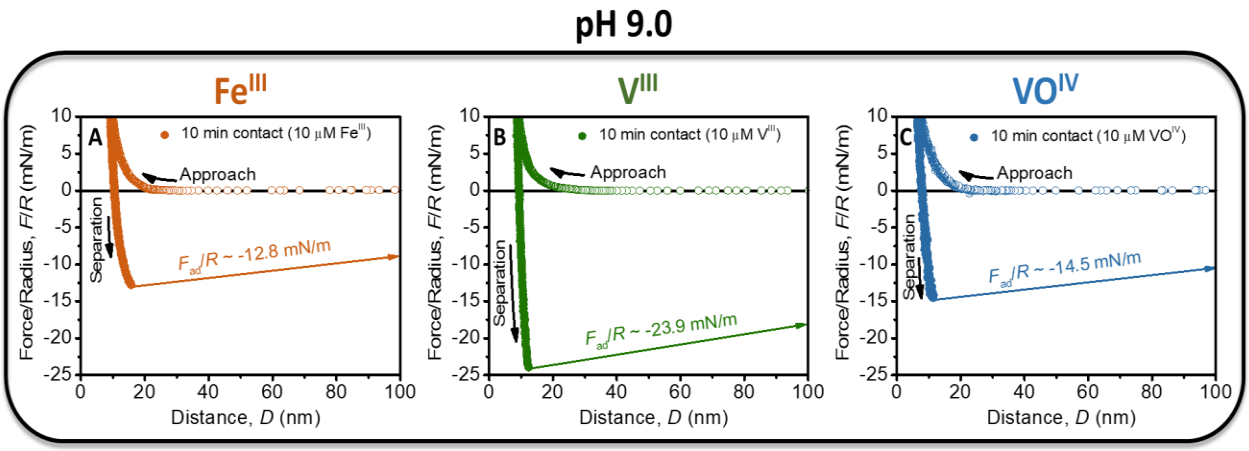

Figure 30: Metal-mediated bridging adhesion in symmetric rMfp-1-DOPA films on mica at pH 9.0. Mica plates are dip-coated in a 20 µg/mL rMfp-1-DOPA solution in 150 mM AcOH for 30 minutes, and subsequently incubated with metal solutions for 30 minutes . A) 10 minute contact following incubation with 10 µM Fe^{III}. B) 10 minute contact following incubation with 10 µM V^{III}. C) 10 minute contact following incubation with 10 µM VO^{IV}.

Examples include mfp-1 symmetric films with Fe^{III} (W_{ad} ≈ 4.2 mJ/m^2),[127] mfp-2 symmetric films with Fe^{III} (W_{ad} ≈ 2.2 mJ/m^2),[124] mfp-3f symmetric films with type 1 collagen (W_{ad} ≈ 5.4 mJ/m^2)[138] and mfp-5 symmetric films without bridging ions or molecules (W_{ad} ≈ 2.5 mJ/m^2).[139]

The difference in cohesion between Fe^{III} and V^{III} prompted us to examine how the metal ions behave when they must compete with one another for rMfp-1-DOPA binding sites. Initially, Fe^{III} was incubated with rMfp-1-DOPA and its cohesion was measured (**Figure 32A**), producing a value in line with the previous Fe^{III} trial and previously published data.[127] Then, the same mica plates were thoroughly rinsed with buffer and incubated with V^{III} (**Figure 32B**). Remarkably, the cohesion strength increased by 90%, reaching -16.3 mN/m, approaching values produced by fresh rMfp-1-DOPA films incubated with V^{III}. This suggests that V^{III} has the capacity to displace Fe^{III} from catechol binding sites in rMfp-1-DOPA under acidic pH conditions relevant for cuticle formation.

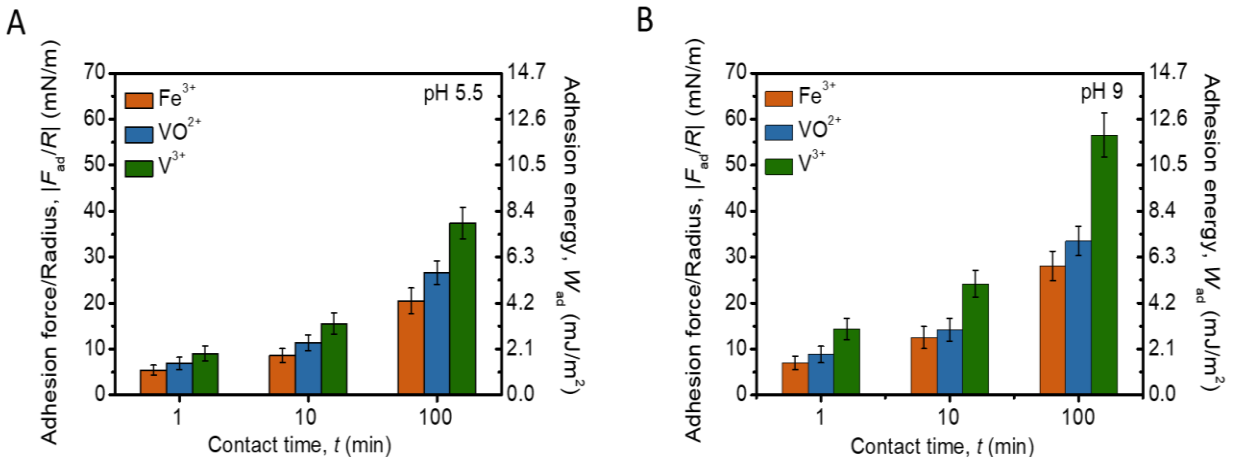

Figure 31: Bar graphs comparing Normalized Adhesion Forces (F_{ad}/R) and Adhesion Energy (mJ/m^2) of rMfp-1-DOPA across contact times and metals. A) Comparison of samples at pH 5.5. B) Comparison of samples at pH 9.0.

Rinsing once again and re-incubating with Fe^{III} yields only a minute decrease in the cohesion strength, demonstrating that Fe^{III} is much less effective at displacing V^{III} from existing complexes (**Figure 32C**). Once again, the predisposition of V^{IV} to form tris complexes under these specific conditions (pH 5.5) as opposed to bis complexation for Fe^{III} could play a key role in the overall stability of a complex and its bond lifetime.[94] As SFA does not provide insight into the

nature of the complexes, further experimentation was performed using other methods to extract further information on the nature of the DOPA-metal complexes.

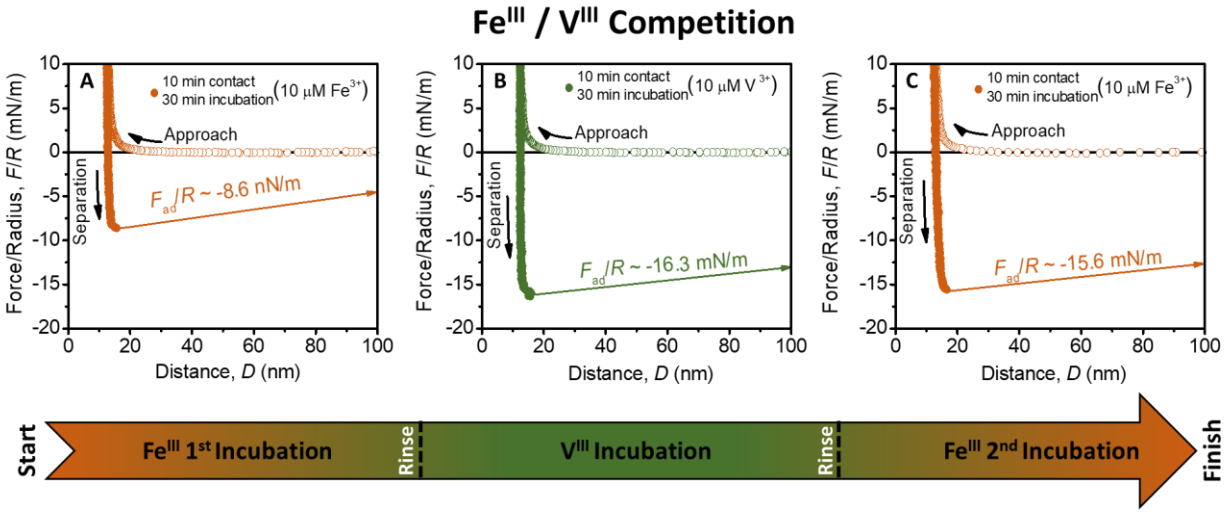

Figure 32: Competition of V^{III} with Fe^{III} for bridging adhesion on symmetric rMfp-1-DOPA films on mica. Experiments are done sequentially on the same set of mica plates, with thorough buffer rinsing between incubation with different metals. A) 10 minute contact following incubation with 10 μM Fe^{III}. B) 10 minute contact following rinse and incubation with 10 μM V^{III}. C) 10 minute contact following rinse and incubation with 10 μM Fe^{III}.

3.4. UV-Vis Spectroscopy

While UV-Vis spectra of catechols and small DOPA-containing peptides have been studied thoroughly, no comprehensive study has been performed on a mussel protein mimic, such as rMfp-1-DOPA.[66, 95] This thesis fills this gap in the literature with a departure from the simpler systems studied previously providing important insights into the metal binding modes of mfp-1. pH-dependent spectra were obtained for rMfp-1-DOPA, with catechol and PEG-DOPA (a previously reported mussel-inspired hydrogel material) serving as controls. Across all samples, a 3:1 ratio of catechol to metal ion was maintained and measurements were performed at a range of different pH values, as this is known to have a large effect on the ability of DOPA to bind metal ions.

Beginning with Fe^{III}, catechol displays clearly identifiable spectra for the mono, bis and tris Fe^{III}-catecholato complexes (**Figure 33A**), serving as the benchmark for the DOPA-containing samples being tested. The spectra acquired from 4-arm PEG-DOPA mixtures with metal ions

(**Figure 33D**) share strong similarities to the pure catechol spectra; however, a noticeable red-shift can be seen in the peak of the tris complex, possibly suggesting some increased difficulty in forming this complex. Interestingly, the rMfp-1-DOPA-Fe^{III} spectra (**Figure 33G**) exhibit significant deviations from the two controls. The first deviation is the lack of the mono complex spectral features at lower pH. Both in the controls and the literature[66, 95], the mono complex exhibits absorbances around 440 and 715 nm; however, this is not detected in the rMfp-1-DOPA sample. Instead, in this lower pH range, the bis complex is already observed. The tris complex presents the second deviation from the controls in the form of an even stronger red-shift than the PEG-DOPA sample. This phenomenon is not previously reported for catechol and DOPA-containing peptide complexes, nor in the other catechols tested in this thesis, suggesting the shift arises from a mechanism that is only accessible to rMfp-1-DOPA.[66, 95] This shift may originate from the presence of both bis and tris species that contribute to the UV-Vis spectrum. A conceivable explanation for this observation is stronger steric constraints as the DOPA is restricted within the protein backbone instead of the terminus of a PEG chain or as freely soluble catechol.

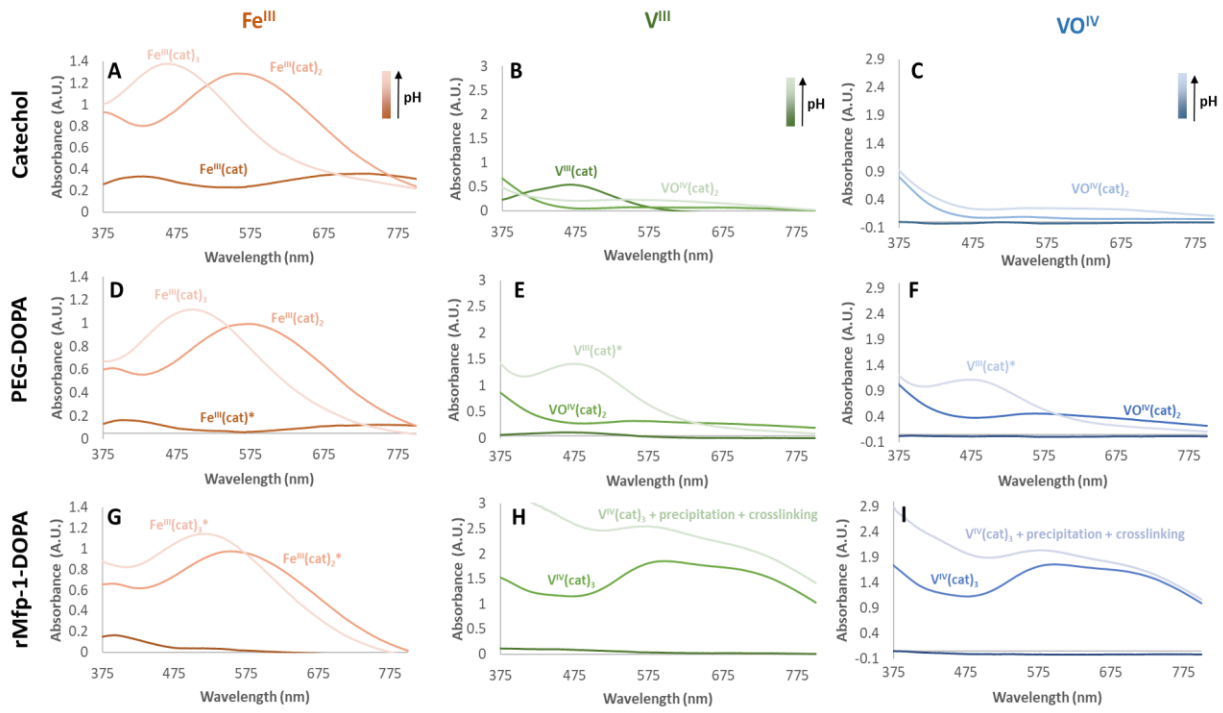

Figure 33: Key UV-Vis spectral changes as a function of pH in solutions starting from the described catechol and metal ion. Titration with NaOH from acidic starting conditions prompts the appearance of increasingly coordinated metal centers. All solutions contain a 3:1 ratio of

catechol to metal ion. Asterisks indicate spectra that either cannot be assigned with complete confidence or contain strong spectral contributions from other complexes. Importantly, the 3 chosen spectra in each plot do not represent the same pHs across plots, they merely show the appearance of new complexes as pH increases.

This could result in a scenario in which a third DOPA is sterically restricted from completing the tris complex even if the conditions would favor this, since it might simply be too far away, and the protein backbones cannot twist enough to bring it into close proximity. This effect is likely compounded by the mushroom tyrosinase reaction that is performed to create DOPA; a bimodal distribution of modified tyrosines was frequently observed in the MS data (**Figure 24**), hinting that mushroom tyrosinase converts certain residues preferentially over others.

Aside from which complexes are formed in the different sample, the pH at which they form also bears relevance. For the pH discussion, PEG-DOPA is omitted as its behavior is almost identical to catechol. Beginning with Fe^{III} (**Figure 34**), catechol shows a transition from the mono to the bis complex around pH 5.3, and the bis to tris transition occurs around pH 10.4. These values align well with previously reported values for Fe^{III}-catecholato complexes.[140]

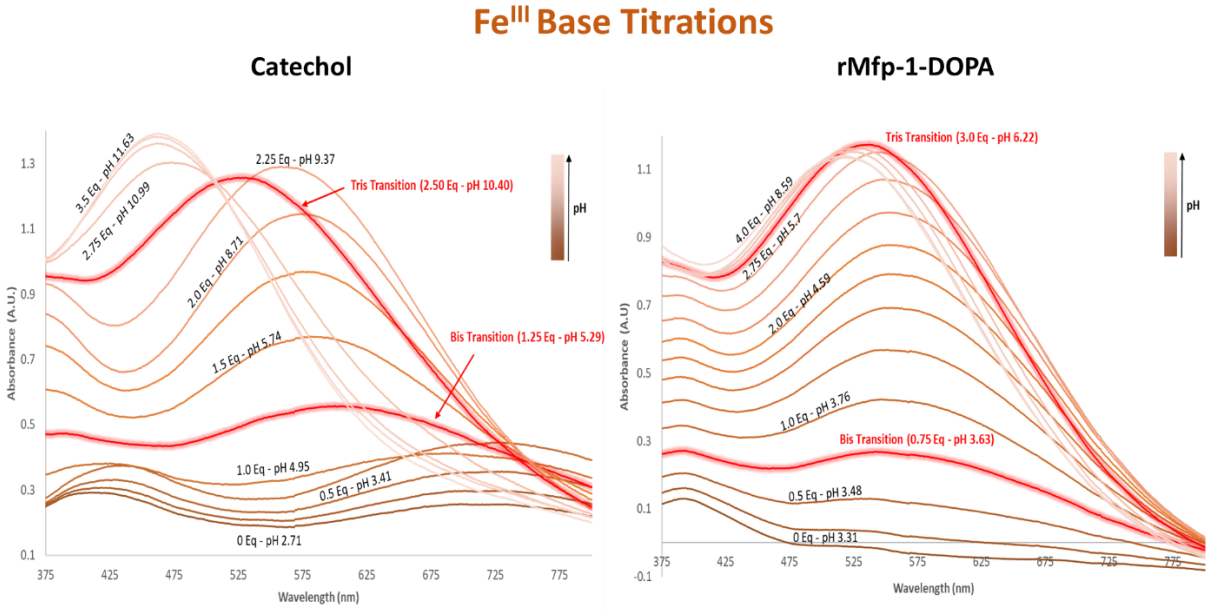

Figure 34: pH dependence of complex formation between catechol and rMfp-1-DOPA in the presence of Fe^{III}. Samples contain a 3:1 catechol : Fe^{III} ratio. Highlighted red lines indicate the point at which a transition to a new complex can confidently be assigned.

However, spectra measured from rMfp-1-DOPA mixed with Fe^{III} (3 DOPA: 1 Fe) indicate that the pH values at which the different forms of the complex appear are dramatically shifted to lower pH values - pH 3.6 for the mono to bis transition, and pH 6.2 for the bis to tris transition. rMfp-1-DOPA is only plotted to pH 8.59, as further base titration (up to pH 11) does not produce any spectral changes.

UV-Vis measurements of solutions made using V^{III} revealed discrepancies across spectra that were more prevalent than those observed with Fe^{III}. Catechol does form a mono complex with V^{III} at low pH, but it displays more enigmatic behavior as the pH rises (**Figure 33B**). In the presence of oxygen, V^{III} is reported to be an unstable species when it is not kept in highly acidic and reducing solutions,[90] oxidizing to VO^{IV} (the vanadyl ion) over the span of several days. As our samples approach neutral pH, the previously described distinct bis-DOPA-V^{III} spectrum with a peak at 400 nm does not appear.[96] Rather, absorbance starts to develop between 500-700 nm, which is characteristic of the presence of bis VO^{IV} or tris V^{IV} complexes. These peaks strongly indicate the occurrence of an oxidation event occurring to V^{III}, even as it is bound by catechol in a mono form. Such oxidation events have been observed in the past for V^{III}, and were prevented by performing experiments in an oxygen-free environment.[95] Thus, the remaining spectra after the disappearance of the mono-DOPA-V^{III} can be assigned as bis VO^{IV} owing to the low absorbance of the species in solution (the tris form has an extinction coefficient that should provide absorbances around 2 A.U. at these concentrations), and the literature observation that catechol struggles to displace the oxo ligand on VO^{IV}, requiring ratios of 50 : 1 catechol : VO^{IV} to reliably form $V^{IV}(cat)_3$.[95] Moving to PEG-DOPA (**Figure 33E**), some commonalities exist, such as the formation of the bis VO^{IV} complex, and the inability to progress to a tris V^{IV} complex at a 3:1 stoichiometry. Strangely, the V^{III} mono complex that was seen for catechol has a very low intensity when using PEG-DOPA, and as the pH enters the basic regime, a strong signal consistent with a mono V^{III} complex returns. While the return to a mono V^{III} complex seems strange at a glance, it may be supported by previous evidence that PEG can function as a reducing agent.[141] Similar to experiments with Fe^{III}, rMfp-1-DOPA (**Figure 33H**) deviates significantly from the controls when mixed with V^{III}. As with Fe^{III}, a mono complex is not observed at all, but remarkably a rapid formation of a tris V^{IV} complex with a characteristic high extinction coefficient is observed

even though this was not possible with catechol or PEG-DOPA at these same stoichiometric ratios of DOPA : metal. While it is possible that some bis VO^{IV} complex is also present, it is difficult to identify in the UV-Vis spectra due to the much smaller extinction coefficient and its superposition with the tris complex absorbance. As pH increases into the basic regime, the absorbance in the UV and near-UV regions sharply rises whilst the tris V^{IV} complex peaks remain mostly unperturbed. This could be the result of covalent cross-linking between oxidized DOPA residues (DOPA quinones forming C-C bonds) as well as precipitation, which can be observed upon visual inspection of the solution; both of these processes may form an important part of the cuticle curing process. The specific pH values at which the different molecules mixed with V^{III} form various complexes can be seen in **Figure 35**. The catechol samples exhibit a transition from mono V^{III} to bis VO^{IV} around pH 5.3, and no further transitions to recognizable spectra can be observed as pH continues to be raised. Conversely, rMfp-1-DOPA does not exhibit any V^{III} complexes, rather seeming to form bis VO^{IV} briefly at pH 3.3.

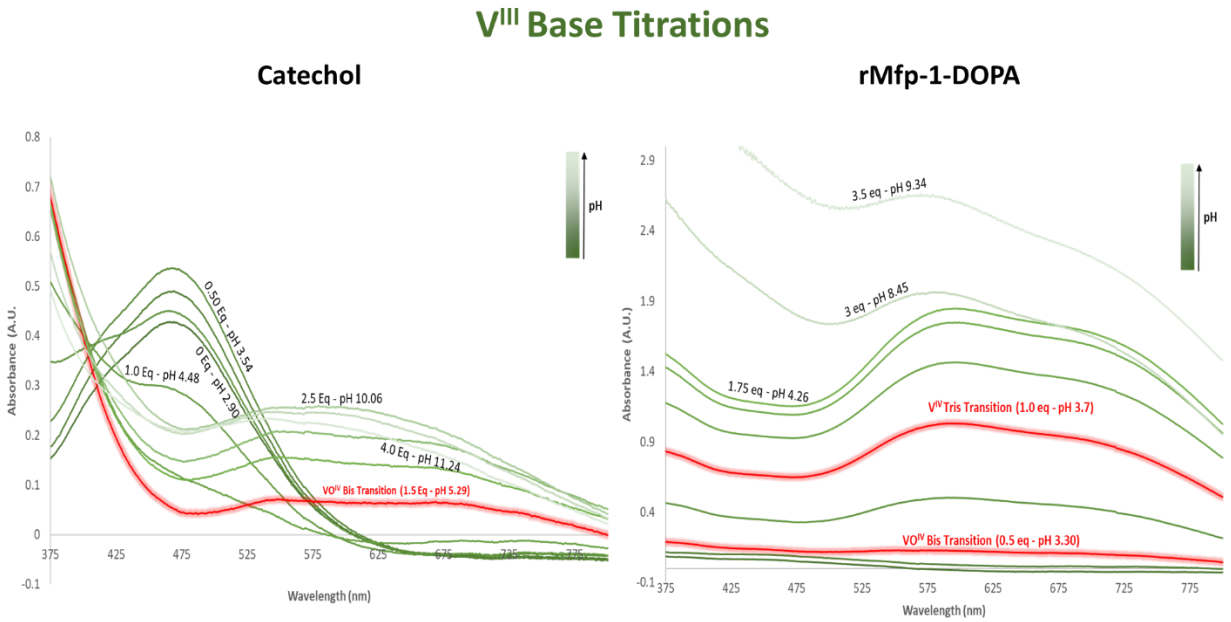

Figure 35: pH dependence of complex formation between Catechol and rMfp-1-DOPA in the presence of V^{III}. Samples contain a 3:1 catechol : V^{III} ratio. Highlighted red lines indicate the point at which a transition to a new complex can confidently be assigned.

This happens over a narrow pH range, with the highly intense tris V^{IV} spectra beginning to appear already around pH 3.7, with the strongest absorbances being attained around pH 4.3. Beyond

this, spectra remain fairly static until pH 8, at which point a sharp rise in absorbance across the spectrum was observed. This is attributed to covalent cross-linking and precipitation, a well-documented phenomenon in the curing of various mussel proteins.[56]

The last set of solutions were made with VO^{IV}, and for the most part, match very well with the spectra acquired with V^{III} at neutral and high pH. This is in line with the expectation that V^{III} does undergo oxidation to VO^{IV} in oxygenated solutions approaching neutral pH.[95] Catechol (**Figure 33C**) once again settles into a low intensity bis VO^{IV} spectrum by pH 3.5, but does not appear to form the tris V^{IV} complex. At low pH, the low extinction coefficients of VO^{IV} complexes make the identification of the mono complex difficult. PEG-DOPA mixed with VO^{IV} (**Figure 33F**) displays almost identical behavior to the V^{III} sample, once again exhibiting an apparent reduction back to V^{III} at basic pH. Lastly UV-Vis results from rMfp-1-DOPA (**Figure 33I**) mixed with VO^{IV} are also very similar to data acquired from the V^{III} sample. Since the starting ion here is VO^{IV}, rMfp-1-DOPA must displace the existing stable oxo ligand on VO^{IV}, which is something that catechol and PEG-DOPA are apparently unable to achieve at the stoichiometries measured in this experiment. As can be seen, rMfp-1-DOPA is able to form tris complexes at even slightly acidic pH, although it is notable that the maximum intensity of the spectra acquired from VO^{IV} is slightly lower than V^{III}. This suggests the presence of a more significant contribution of bis complexation, possibly consistent with the increased difficulty of oxo ligand removal in VO^{IV} that is not faced when starting with V^{III}. Addressing pH behavior more specifically, VO^{IV} (**Figure 36**) samples mirror the V^{III} samples quite well. The catechol control samples once again did not to form complexes beyond a bis VO^{IV} complex, even as pH rises to basic levels (i.e. tris complexation was not observed). rMfp-1-DOPA mixed with VO^{IV} also formed the tris V^{IV} species at low pH (3.8), with any mono or bis VO^{IV} complexes being difficult to discern. This pH-dependent spectroscopic data strongly suggest that peptidic nature of rMfp-1-DOPA significantly changes the chelation properties of the DOPA residues in the protein chain compared with simpler catechol species. Specifically, 1) the pH at which multi-ligand complexes is formed is significantly lowered, especially with vanadium. 2) DOPA residues are able to displace a VO^{IV} oxo atom at stoichiometries that are not possible with free catechol. This idea that the protein sequence and

structure might influence metal chelation, as well as possible mechanisms by which this might occur will be explored thoroughly in the discussion section.

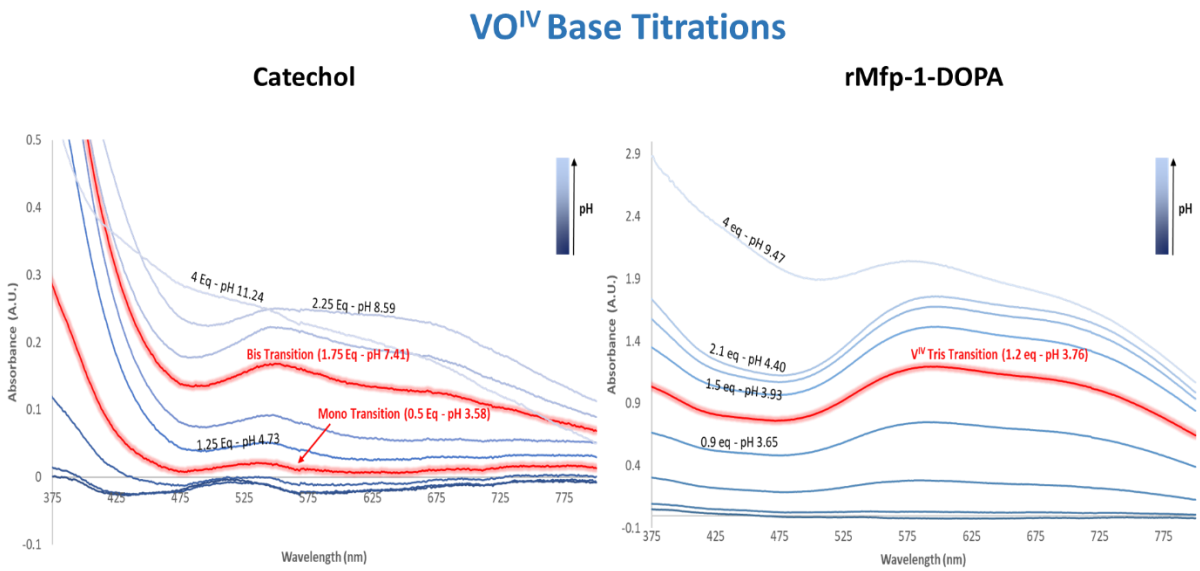

Figure 36: pH dependence of complex formation between Catechol and rMfp-1-DOPA in the presence of VO^{IV}. Samples contain a 3:1 catechol : VO^{IV} ratio. Highlighted red lines indicate the point at which a transition to a new complex can confidently be assigned.

3.5. Raman Spectroscopy

DOPA-metal complexes have been previously shown to exhibit a strong resonant Raman effect with appearance of a characteristic sets of peaks corresponding to vibrations within the complex, including metal-oxygen interactions spanning the 500-700 cm^{-1} range and catechol ring vibrations in the 1200-1500 cm^{-1} range.[142] In **Figure 37**, one can see these resonance peaks in all spectra measured including those from complexes of rMfp-1-DOPA with Fe and V, as well as from the native byssus cuticle. The exact positions of these peaks are influenced by the bond lengths and geometry of the complex. Thus, the uniformity of the spectra hint at a high consistency between the DOPA-metal complex structure formed using vanadium, even when starting with vanadium in different oxidation states considering that the positions of these peaks are highly sensitive to even small changes in coordination bond length.[143] The resonance Raman peaks originating from complexes made with Fe are notably shifted compared with the V complexation peaks, consistent with previous findings.[93] Correlating these observations to UV-Vis results

(**Figure 33**) supports the conclusion that both V^{III} and VO^{IV} end up as a highly similar complex (primarily tris-DOPA-V^{IV} based on the UV-Vis data).

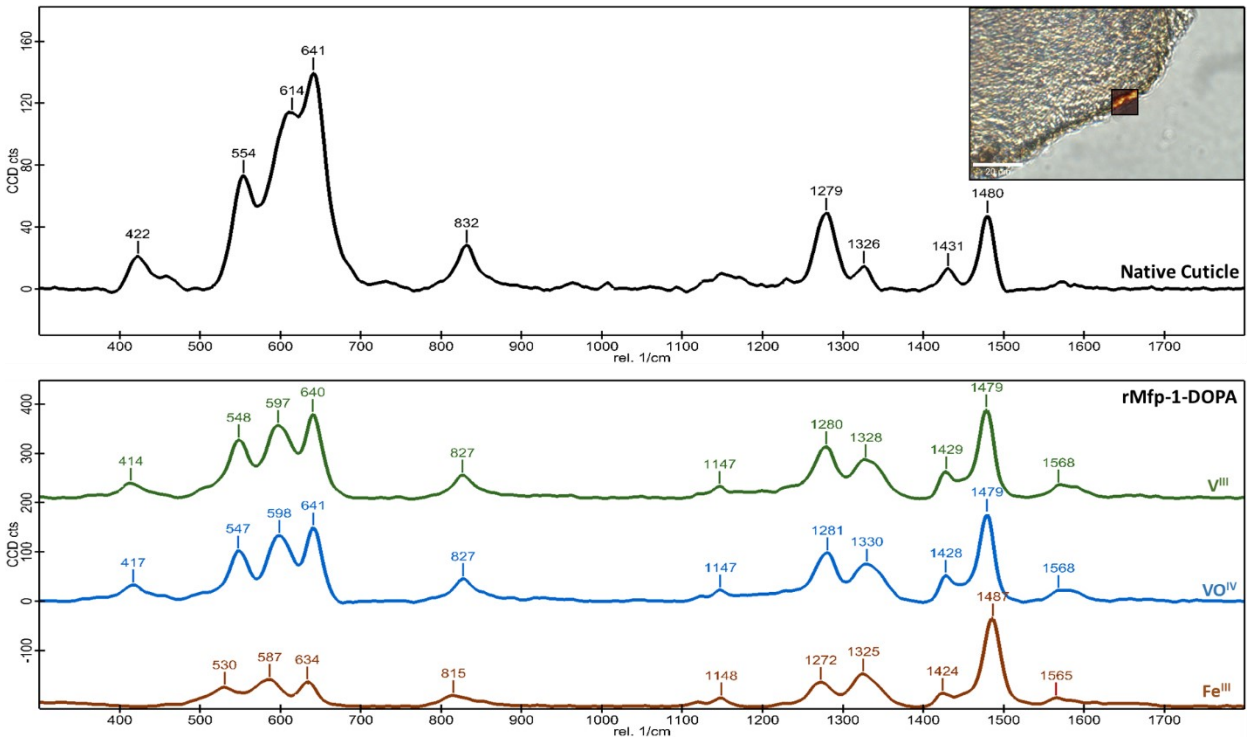

Figure 37: Resonance Raman spectroscopy of a native cuticle section, as well as dried rMfp-1-DOPA spots with various metal ions at pH 5.5. Prominent Resonance Raman peak positions are labelled to facilitate comparison between native cuticle and rMfp-1-DOPA samples.

Notably, the native cuticle shows a spectrum that more strongly resembles rMfp-1-DOPA with vanadium as opposed to iron; this observation lines up well with the recent TEM-EDX studies that illustrated the anti-correlation of iron and vanadium in the granular structure of the cuticle and the co-localization of V with the mfp-1 enriched granules.[61]

3.6. Electronic paramagnetic resonance (EPR) spectroscopy

EPR was used in order to further characterize the rMfp-1-metal complexes. Starting with r-Mfp-1-DOPA and VO^{IV} at ph 3.0 (3 DOPA : 1 V) (**Figure 38**), a well resolved, strongly anisotropic spectrum emerges. This spectrum, and its parameters (**Figure 40**) correlate well with literature reports of the electronically axially symmetric VO^{IV} pentaaqua complex that exists in acidic solutions (**Figure 39**). Such spectra are typically characterized by 2 overlapping 8-line EPR signals, with dramatically different Hyperfine Coupling Constants.[123] This supports the idea that rMfp-1-

DOPA does not form complexes with VO^{IV} at pH 3.0, as was evidenced by the UV-Vis data above. Additionally, whilst previous EPR measurements of whole threads have indicated the presence of carbon radicals, this data could be fitted without considering contributions from carbon radicals, which given the low pH are not be expected to be prominent.[93]

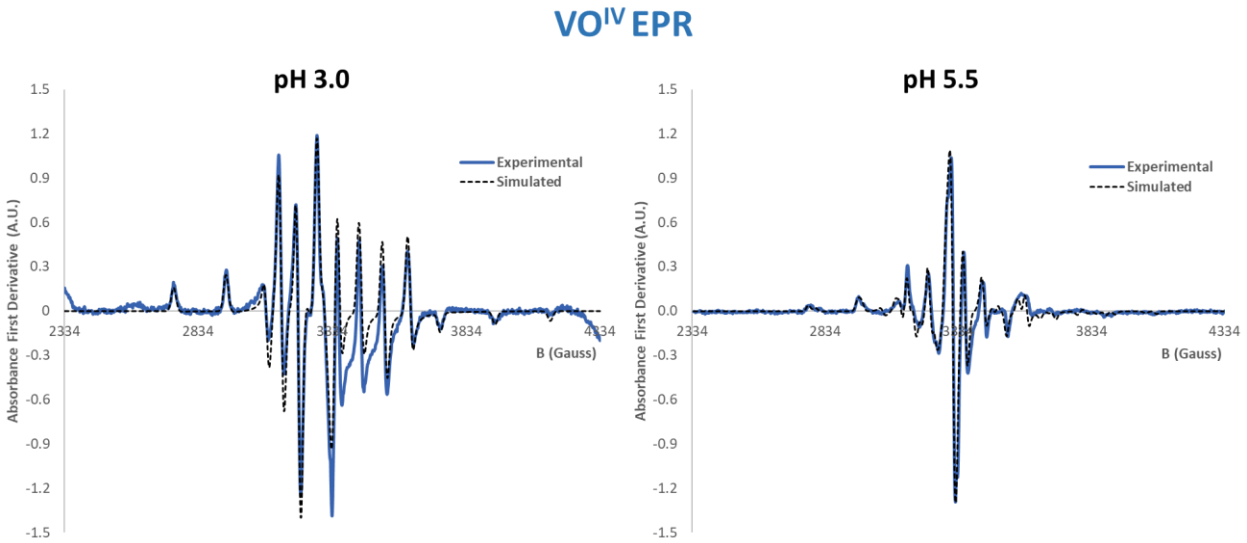

***Figure 38: Experimental and simulated EPR spectra of rMfp-1-DOPA samples containing V^{III}.* Samples contained a catechol : VO^{IV} ratio of 3:1.**

As the pH is raised to 5.5 (equivalent to the pH of the SFA measurements described above), the spectrum (**Figure 38**) undergoes dramatic changes. A carbon radical contributes strongly in the central part of the spectrum, and much of the anisotropy of the pH 3.0 sample is lost. While the spectrum supports the conclusion that a large part of the VO^{IV} is now in the form of $V^{IV}(DOPA)_3$, the remnants of anisotropic peaks on the fringes of the spectrum suggest the presence of some $VO^{IV}(DOPA)$ and $VO^{IV}(DOPA)_2$ (**Figure 40**). This gives important insight into the system that UV-Vis cannot, as the mono and bis complexes have small extinction coefficients that are dwarfed by the tris species.

Literature	g_0	g_x	g_y	g_z	A_0	A_x	A_y	A_z	Ref.
$[VO(H_2O)_5]^{2+}$	1.966	1.981	1.981	1.933	109	70	70	188	*144*
VO(cat)	1.968	1.976	1.976	1.943	96	63	63	170	*100*
$[VO(cat)_2]^{2-}$	1.975	1.981	1.981	1.947	82	50	50	154	*100*
$[V(cat)_3]^{2-}$	1.955	1.937	1.937	1.991	76	107	107	14	*100*

Figure 39: Literature parameters for V^{IV} EPR active nuclei relevant to the study of rMfp-1-DOPA. All A parameters are provided in units of 1 x 10^{-4} cm^{-1}. g_0 and A_0 parameters are obtained from liquid spectra when available, or otherwise calculated via the relation $x_0 = x_x + x_y + x_z$.

VO^{IV} pH 3.0 Sim	g_0	g_x	g_y	g_z	A_0	A_x	A_y	A_z	Weight
$[VO(H_2O)_5]^{2+}$	1.965	1.979	1.981	1.934	106	69	69	181	1
VO^{IV} pH 5.5 Sim									
VO(cat)	1.968	1.980	1.980	1.945	97	61	61	170	1.35
$[VO(cat)_2]^{2-}$	1.971	1.981	1.981	1.950	90	54	54	161	0.93
$[V(cat)_3]^{2-}$	1.953	1.933	1.941	1.986	74	123	98	1	0.85
Carbon Radical	2.011	-	-	-	-	-	-	-	0.18

Figure 40: Simulated parameters for EPR spectra of rMfp-1-DOPA containing VO^{IV}. All A parameters are provided in units of 1 x 10^{-4} cm^{-1}. g_0 and A_0 parameters are calculated via the relation $x_0 = x_x + x_y + x_z$. The weight column indicates the relative weighting of each species simulated to produce the best agreement with experimental data.

Moving to the samples containing r-Mfp-1-DOPA and V^{III}, several different behaviors were observed. The pH 3.0 sample once again exhibits a VO^{IV} pentaaqua complex, albeit with a lower intensity (**Figure 41 and Figure 42**). This is likely due to the slow oxidation process that produces VO^{IV} from V^{III} in the presence of oxygen, even in freshly prepared solutions.

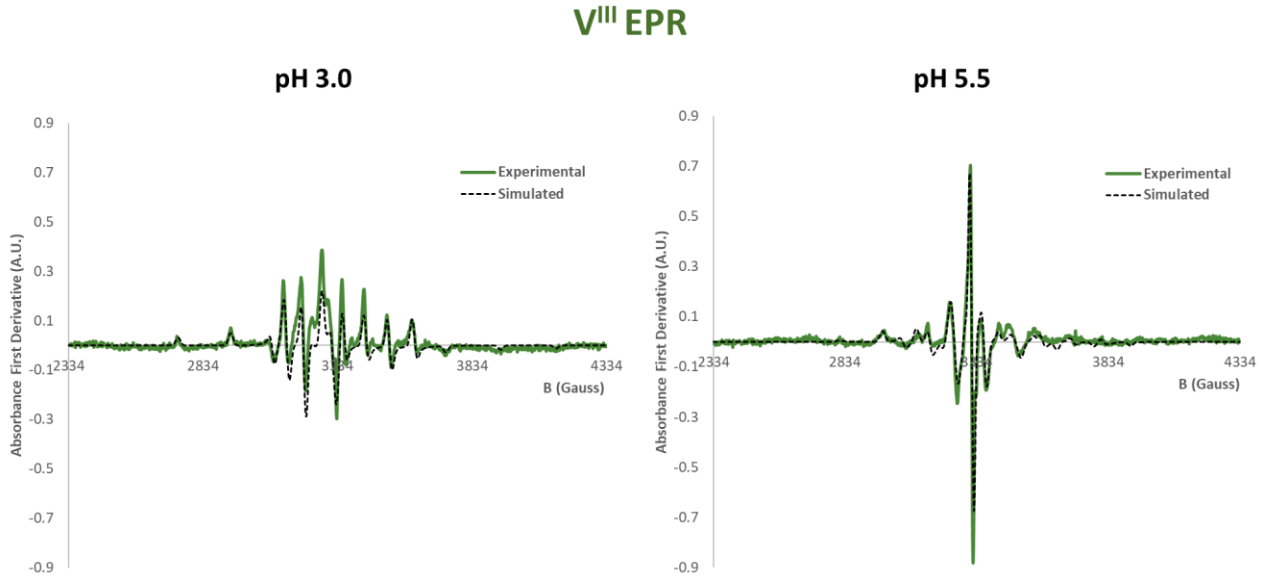

***Figure 41: Experimental and simulated EPR spectra of rMfp-1-DOPA samples containing V^{III}.* Sample contained a catechol: V^{III} ratio of 3:1.**

VO^{IV} is confidently assigned here due to V^{III}'s inability to be detected on X-band spectrometers; V^{III} nuclei have 2 unpaired electrons and a large zero-field splitting term in the spin Hamiltonian, generating an electron spin transition energy that is greater than what can be provided by an X-band microwave source. The absence of V^{III} complexes can be positively confirmed by examination of V^{III} UV-Vis spectra, which show a complete absence of defined peaks in the 375 – 500 nm region (the region that contains high extinction coefficient peaks for mono, bis and tris V^{III} complexes). As before, carbon radicals show negligible contribution to this low pH spectrum. At pH 5.5, a strong contribution from carbon radicals presents itself, and in this case, a much more stoichiometric conversion from V^{III} or VO^{IV} precursors to $V^{IV}(DOPA)_3$ occurs (**Figure 41 and Figure 42**). The almost complete disappearance of the anisotropic peaks present in other spectra is a good indicator of this process. While the simulations did assign some contribution from mono and bis VO^{IV} complexes to the final spectrum, their parameters are significantly removed from literature values; this most likely occurs due to lower concentrations of these species, and the resulting low signal intensity that the fitting algorithm must use for these species.

V^{III} pH 3.0 Sim	**g_0**	**g_x**	**g_y**	**g_z**	**A_0**	**A_x**	**A_y**	**A_z**	**Weight**
$[VO(H_2O)_5]^{2+}$	1.965	1.980	1.980	1.934	106	68	68	181	1
V^{III} pH 5.5 Sim									
VO(cat)	1.971	1.981	1.981	1.950	87	51	51	158	0.39
$[VO(cat)_2]^{2-}$	1.977	1.984	1.984	1.964	89	59	59	148	0.33
$[V(cat)_3]^{2-}$	1.953	1.933	1.939	1.986	75	121	98	5	1.72
Carbon Radical	2.011	-	-	-	-	-	-	-	0.20

Figure 42: Simulated parameters for EPR spectra of rMfp-1-DOPA containing V^{III}. All A parameters are provided in units of 1 x 10^{-4} cm^{-1}. g_0 and A_0 parameters are calculated via the relation $x_0 = x_x + x_y + x_z$. The weight column indicates the relative weighting of each species simulated to produce the best agreement with experimental data.

4. Discussion

The main goal of this thesis work was to obtain new insights into the structure, function, and formation of the mussel byssus cuticle. I aimed to address two major open questions in our understanding of this biological materials: 1) How are the different proteins that comprise the cuticle organized between matrix and granules? 2) What role do the different metal ions associated with the cuticle (e.g. Fe and V) play in the function and formation of the cuticle and how are they complexed by Mfp-1. To investigate question one, I analyzed TERS data that had been acquired from the matrix and granules of the cuticle. This method showed some potential in terms of increased spatial resolution of our compositional understanding of the cuticles; however, there were many challenges in terms of data acquisition and statistical analysis that must first be overcome before this method can yield novel insights into this material system. To address the second question, I examined the ability of recombinantly expressed Mfp-1 protein (rMfp-1-DOPA) to bind Fe and V, and the mechanical implications of this metal coordination cross-linking. This second approach yielded many interesting findings; thus, I primarily focus in the discussion on the implications arising from my work on metal binding of rMfp-1-DOPA, in particular the comparative studies of Fe vs. V binding.

To briefly summarize, experimentation on rMfp-1-DOPA produced several compelling findings. First, it was found using SFA that both V^{III} and VO^{IV} ions generate significantly stronger cohesive forces between rMfp-1-DOPA proteins than Fe^{III} ions, with V^{III} showing nearly twice the cohesive pullout force compare with Fe^{III} under a range of different pH conditions. Second, UV-Vis and Raman spectroscopy showed that the vanadium-based cohesion is most likely generated not by the V^{III} ions introduced, but by tris-DOPA-V^{IV} complexes that are generated through redox processes in solution. Third, these studies also highlighted the ability of DOPA residues in rMfp-1-DOPA to displace the oxo group for VO^{IV} at stoichiometric ratios of 3 DOPA : 1 VO^{IV}, something that is not possible with pure catechol except in large excess of catechol. Fourth, UV-Vis investigation also demonstrated that rMfp-1-DOPA can readily form bis and tris complexes with Fe or V at pHs as low as 3, which is not possible for pure catechol at the same 3:1 catechol : metal ratio. Both these findings suggest that the protein sequence surrounding the DOPA residues in mfp-1 contributes significantly to determining and finetuning the manner in which the catechol

moieties interact with different metal ions. Lastly, EPR measurements showed that despite both V^{III} and VO^{IV} converging to a tris coordinated DOPA-V^{IV} species in the presence of rMfp-1-DOPA, V^{III} undergoes a much more facile conversion, and produces a higher proportion tris- catecholato-V^{IV} complexes as a result, which may reflect challenges in displacing the oxo group of the VO^{IV} ion. While these results have answered some questions, they have also produced new questions that must be answered in future projects. For now, we must resign ourselves to hypotheses based on our existing knowledge.

4.1. Vanadium-induced Precipitation of rMfp-1-DOPA

An interesting phenomenon that is observed with rMfp-1-DOPA and both V^{III} and VO^{IV} is a sudden rise in near-UV absorption as pH rises above 8.0, which is attributed to precipitation and covalent cross-linking. rMfp-1-DOPA shows remarkable levels of solubility at all physiologically and environmentally relevant pH value for *Mytilus edulis*, except in the presence of vanadium ions that prompt precipitation at basic pH. While this made UV-vis spectroscopic data collection challenging under basic pH collection, upon further reflection it is conceivable that this inherent tendency to come out of solution may be functionally relevant during the cuticle formation. The formation process of the cuticle remains rather enigmatic, but one mandatory requirement is the phase transition from liquid protein precursor, to solid functional material. Given our observations, it stands to reason that vanadium might be a key player in this process that is triggered by the basic pH of seawater. Presently, a difficult question to answer is how this happens. One potential mechanism might be that tris complexes may favor solidification, whereas bis complexes do not. This was observed in experiments using PEG-DOPA-Fe hydrogels, in which only a viscous paste was formed at pH 8 due to the absence of tris complexes.[140] Another plausible way that vanadium drives this process is via generation of radicals that induce covalent cross-linking between amino acid residues.[145] In contrast to iron, vanadium ions readily generate carbon-centered radicals in rMfp-1-DOPA as pH approaches neutral territory (**Figure 38 and Figure 41**). These radicals may be the catalyst for the strong absorbance increases in the UV tail of UV-Vis spectra of vanadium ions with rMfp-1-DOPA (**Figure 35**). Possible products of this radical formation process include di-DOPA, lysyl-DOPA, histidyl-

DOPA and cysteinyl-DOPA.[69, 146] As the amount of cross-links increases between different rMfp-1-DOPA chains, the proteins begin to aggregate, and eventually fall out of solution, driving the cuticle curing process.

An important consideration is that these experiments were done in rather dilute protein conditions (250 μM). Conversely, the mussel can highly concentrate certain proteins (up to concentrations of 300 mg/mL or ≈ 22 mM for rMfp-1) in the previously introduced process of coacervation, and secrete them to produce threads.[81, 111] In this highly dense, coacervated form (observed both *in vitro,*[111] and through TEM studies of the cuticle formation process[61]), proteins gain access to properties that are inaccessible to typical protein solutions.[106] For Mfp-1, this could translate to easier "wetting" of the core structure during assembly, or facilitation of cross-linking and curing processes between densely packed protein strands within the confined microenvironment of the coacervate.

While our experimentation was done only with rMfp-1-DOPA, it also bears mentioning that the cuticle is currently believed to have other protein constituents, such as several putative cysteine-rich proteins that have been identified transcriptomically and through EDX analysis.[49, 61] Cysteines in byssal proteins are proposed to play important antioxidant roles, preventing the premature oxidation of DOPA that would harm a proteins' capacity for metal coordination bonds.[79] Along these lines, previous TEM-EDX studies indicate that the DOPA proteins are phase-separated from mfp-1, but still in very close contact, in a block copolymer-like gyroid configuration with layer thickness of ~20 nm.[61] The close association of the Cys-rich proteins and mfp-1 suggest there may be a delicate balance of vanadium-based metal coordination cross-linking (providing kinetically labile dynamic cross-linking) and oxidative covalent cross-linking via DOPA radical formation. While the antioxidant properties of the cysteine-rich proteins remains open to further exploration, previous studies in which metals were removed from the cuticle and then added back showed that up to 85% of the stiffness and hardness of the cuticle comes from the presence of DOPA-metal coordination bonds.[93]

4.2. Complex Formation Across Different DOPA-containing Molecules

As described in the results, the abilities of catechol, PEG-DOPA and rMfp-1-DOPA to form complexes at various pHs was explored. The consistent ability of rMfp-1-DOPA to form highly coordinated complexes under acidic pH conditions and its ability to form tris-DOPA-V^{IV} complexes when supplied with either V^{III} or VO^{IV} ions are both results that warrant further discussion.

rMfp-1-DOPA was observed to form predominantly tris DOPA-metal complexes in UV-Vis spectra with Fe^{III} at pH 6.2 (**Figure 34**), V^{III} at pH 3.7 (**Figure 35**) and VO^{IV} at pH 3.76 (**Figure 36**). At the same 3:1 DOPA : metal ion ratio, catechol can only form tris complexes with Fe^{III} once the pH surpasses 10.4 (**Figure 34**). Even stronger contrast exists when looking at the behavior of catechol with V^{III} and VO^{IV} (**Figure 35 and Figure 36**), where characteristic absorptions of tris complexes was never observed at any of the pH values examined. While the exact reasons for these behaviors are unknown, it is highly likely that the DOPA residues in rMfp-1-DOPA are endowed special properties via the amino acid sequence of their parent rMfp-1 chain. This is a very important point with regards to the majority of mussel-inspired polymeric materials that tend to take a reductionist approach in which the DOPA-catechol moiety is utilized completely out of the context of its natural protein sequence. As already shown for mussel adhesion, the presence of lysine residues that often appear adjacent to a DOPA residues in the sequences of mussel adhesive proteins mfp-3 and mfp-5 (as well as mfp-1) enhances adhesion in hydrated environment, likely by displacing the cationic double layer on metal oxide surfaces.[147] This proposed role of lysine is compelling for enhancing adhesion, but is not likely relevant for the cohesive interactions produced between mfp-1 molecules. What then is driving this enhanced metal-binding behavior?

The repeating AKPSY*PPTY*K decapeptide found in Mfp-1 (where Y* indicates DOPA) has been proposed to possess a loose secondary structure consisting of an alternating polyproline II (PPII) helices (characterized as an open coil with *trans* peptide bonds) and unstructured random coil regions.[148] While PPII domains are common in certain protein such as collagens,[149] they have not been observed in the context of metal-binding proteins that play structural roles like Mfp-1. As such, the exact impact of a possible PPII helix on such function is unknown. One hallmark of PPII helices, which are highly extended helix conformations (rise/residue = 0.29 nm),[150] is their stiffness (reported to be 9.6±0.1 GPa in a crystalline system),[151] and this stiffness could be

exploited in conjunction with DOPA to create regular, and well-defined metal binding sites linked by unstructured regions. Additionally, these PPII regions could create microenvironments that modify the pKa of amino acids like DOPA and lysine,[152] possibly modulating the pH at which rMfp-1-DOPA can form higher coordinate complexes with metals like Fe and V. However, this remains to be verified.

Besides the PPII helix motif, other amino acid constituents may influence rMfp-1-DOPAs interesting properties. Lysine likely plays a key role in this protein as it does in other mussel adhesive proteins, as mentioned.[147] A recent study found that lysine directly flanking DOPA impedes Fe^{III}-mediated cohesion via electrostatic repulsion and steric hindrance.[153] As the rMfp-1-DOPA decapeptide sequence contains one KY* pair, it begs the question as to whether spacing out these residues would result in even greater cohesion strengths than what was observed in the SFA experiments (**Figure 31**). One interesting transformation that may benefit from proximal lysines is the displacement of the vanadium-oxo bond in VO^{IV} to form tris-DOPA-V^{IV} complexes. This is achieved by rMfp-1-DOPA effortlessly in a 3:1 DOPA : VO^{IV} ratio, whereas catechol can only complete this transformation at ratios approaching 50:1.[95] Weakening of the V-O bond could conceivably occur via Lewis-acid-carbonyl-activation-type interactions between the lysine cation and the oxo oxygen, facilitating oxygen displacement for awaiting DOPA residues by protonating the oxo oxygen.[154] Other amino acids present in the sequence, such as serine and threonine may also play constructive roles in metal binding. Although they are weaker chelators than DOPA (owing to their monodenticity), serines and threonines are also able to coordinate to transition metals in enzyme active sites and other metalloproteins.[155-156] An interesting extension would be to consider the effects of hydroxyproline and dihydroxyproline, both of which make frequent appearances in the wild-type Mfp-1 sequence, but not in our recombinant variant. One known effect of these modifications is the thermodynamic stabilization of PPII helices (e.g. in collagen triple helical structures),[157] which may further amplify their positive effects on metal chelation.

Although the work in this thesis has revealed several novel properties for rMfp-1-DOPA, it has not focused on generating evidence explaining this behavior. In the future, experiments with rMfp-1-DOPA, variants created via site-directed mutagenesis, and the incorporation of

hydroxylated prolines would be beneficial to truly complete our understanding of why rMfp-1-DOPA exhibits its remarkable chelating capacities.

4.3. Stoichiometry of DOPA : Metal Complexes in the *Mytilus edulis* Cuticle

Throughout this thesis, all experiments have utilized a 3:1 "golden" ratio of DOPA to metal ions, which would be that expected for formation of tris DOPA-metal complexes. While this parameter may be manipulated when performing *in vitro* studies, in reality, the exact stoichiometry present in the byssal cuticle is unknown. Despite this, there is some knowledge of DOPA : metal ion ratios in the cuticle, and relating it to our studies raises some interesting possibilities.

At very high ratios of DOPA to metal ion (e.g. 30:1), the high DOPA concentration drives the complex equilibrium strongly towards tris-chelated complexes, such that all metals have 3 DOPA residues coordinated to them. Whilst complete conversion to tris complexes is desirable, this would result in having free DOPA present in large amounts. High DOPA levels present high levels of oxidative stress that must be managed by the mussel, as DOPA readily oxidizes to DOPA-quinone at basic pH, which is a strong Michael acceptor that can form undesirable covalent cross-links. Additionally, both DOPA and DOPA-quinone have a penchant for forming organic radicals that can also react in undesirable fashion when they are not appropriately controlled. One study probed both metal ion concentrations and DOPA concentrations in the byssal thread using ICP-OES and amino acid analysis, and found that DOPA is consistently in significant excess of all metal ions added together.[93] It also showed that the crosslinks in the cuticle are primarily metal coordination-based (as opposed to covalent bonding), as EDTA treatment reduced cuticle stiffness by 85%, and reintroduction of metal ions restored it to its original value.[93] When vanadium ions (the strongest binder to rMfp-1-DOPA based on SFA in **Figure 31**) are all chelated and excess DOPA remains, the high K_a of DOPA with other ions such as Fe^{III} (~ 10^{18} M^{-1})[66] would prompt further chelation of other metals, and we would expect to see high levels of metal colocalization in the tissue. Despite this expectation, EDX investigation have revealed that vanadium is strongly anti-correlated with iron in the granules, providing contrary evidence that supports a relatively small excess of DOPA to metal ions.[61] Interestingly, the metal removal and

reintroduction mentioned previously results in much higher metal levels than in the native thread (up to 18X for V^{III}), but no change in mechanical performance.[93] In the SFA experiments (**Figure 29**), saturation of DOPA residues with metal ions is easily achieved, resulting in a complete loss of cohesion. This suggests a mechanism for preventing saturation of the DOPA residues exists in the native cuticle system that does not exist *in vitro* for rMfp-1-DOPA.

Based on existing evidence, high DOPA concentration does not seem to produce deleterious effects on cuticle mechanics as expected. As *Mytilus edulis* consistently produces its hard, stiff, extensible and self-healing cuticle coating, it seems that may have evolved a way to prevent oversaturation of DOPA with metal ions. This control mechanism, and whether the ideal ratio of DOPA : metal ions is truly is 3:1 as is commonly used in experimentation will require further investigation.

4.4. Vanadium and its Prevalence in Biological Systems

One of the more important findings of this thesis is that not only can rMfp-1-DOPA utilize vanadium ions in its metal cross-links, but their use grants better mechanical properties than using iron, as was originally posited and believed for the last 10 years until the recent discovery of vanadium localized in the granules. [61]However, until very recently, it was completely unknown whether the metals that are present in the byssus are acquired by passive diffusion from the seawater (in which soluble Fe and V are only present in nanomolar concentrations) or by active uptake by the mussel and integration into the thread during formation. Recent studies of plaque formation in *Mytilus edulis* revealed that the organism can uptake and store vanadium and iron ions in submicron metal storage particles that are co-secreted with the plaque proteins, mixing and forming metal coordination bonds with DOPA residues in the nascent plaque.[158] The ability of *M. edulis* to hyperaccumulate vanadium places it in a very small group with only three other families of organisms that exhibit this same behavior. This section will discuss vanadium in the context of mussels and other hyperaccumulators and will show how mussels stand out by being the only species where functional benefits of vanadium accumulation and use can be ascribed.

First, it is worth mentioning there are a small number of biomolecules that coordinate vanadium in which there is a known function. However, the vanadium is not present to a degree

that would be considered hyperaccumulation. The first class are vanadium haloperoxidase enzymes (found in macroalgae, fungi and bacteria), which take halogenated compounds and oxidize them into more reactive intermediates via the use of hydrogen peroxide.[159] The other class of biomolecules containing vanadium are nitrogenases that reduce dinitrogen to ammonia in nitrogen-fixing bacteria; however, it is important to note that nitrogenases with vanadium containing centers are rare, as most known nitrogenases use molybdenum in their active sites.[159] Beyond these two small groups of enzymes, the world of vanadium is much less thoroughly understood.

Hyperaccumulators of vanadium have made scientists ask, "How and why?" for more than a century. *Amanita muscaria*, more commonly known as Fly Agaric, is a toadstool mushroom that is known to accumulate vanadium utilizing a small natural product called Amavadin.[160] Despite years of research, only guesses exist with regards to its function, such as being a toxin discouraging other organisms from consuming the mushroom.[161] In the aquatic world, ascidians (more commonly known as tunicates or sea squirts), are a class of marine invertebrate filter feeders that have gained notable interest due to the ability of some species to concentrate vanadium up to 10^7 times the concentration found in seawater.[162] Initial analyses pinned the uptake and transport on a small peptide named tunichrome, which contains DOPA-like gallic acid moieties and has vanadium chelation and reduction capabilities. However, little evidence exists today to back this up *in vivo*, as vanadium and tunichrome are stored in different body compartments.[160] A more likely candidate for vanadium uptake are the vanabins, which are also small proteins that chelate and reduce vanadium from V^V to V^{IV} in the cytoplasm. Further reduction of V^{IV} to V^{III} occurs in extremely acidic blood cell vacuoles in association with very high concentrations of sulfate ions.[163] However, in spite of the deep understanding of vanadium redox chemistry in ascidians and the role of specific proteins and subcellular compartments, the ultimate reason why ascidians concentrate vanadium remains unknown.[164] Rounding out these hyperaccumulators are the tube-dwelling Sabellidae family of polychaete worms. Certain species have been found to concentrate vanadium up to a staggering 7000 µg / g of body weight, particularly in the branchial crowns (feather-like filaments responsible for respiration and filter feeding) located at the end of their bodies.[165] The exact molecule responsible for vanadium

accumulation has never been detected, and hypotheses for the exact purpose of this vanadium accumulation range from a method of facilitating oxygen absorption to a method of making themselves "unpalatable" to predators.[165]

In our experiments with rMfp-1-DOPA, we found that vanadium coordination confers quantifiable mechanical differences over iron coordination. At pH 5.5 (a pH level relevant to the formation process of the cuticle), V^{III} produced a 2-fold increase in cohesive strength over Fe^{III} (**Figure 29**). Raising the pH to 9.0 (which is more indicative of the pH of seawater where the cuticle functions as a material), this 2-fold increase persisted, with even higher cohesive forces measured (**Figure 30**). This helped us to dispel our concerns about the difference in cohesive strength coming from different levels of bis and tris complexes in the sample, as both ions should be chelated in a tris state at pH 9 (**Figure 34 and Figure 35**). Moreover, because these experiments were performed in oxygen free conditions, we can conclude that oxidative covalent cross-linking is not a significant contributor here. We can thus posit that the differences in cohesive strength between Fe^{III} and V^{III} derives from differences in relative bond strengths. Additionally, this SFA data helps support the idea that vanadium is not only mechanically superior in the formation of the cuticle at acidic pH, but maintains these superior qualities in the cured form exposed to basic pH.

Another interesting finding stemming from SFA is that while both V^{III} and VO^{IV} produce stronger cohesion, V^{III} consistently performs better at both pH values tested, likely due to greater presence of tris vs bis coordination (**Figure 31**). While rMfp-1-DOPA does have the capacity to remove the oxo oxygen at 3:1 DOPA : VO^{IV} ratios, it does not seem to complete the conversion to V^{IV} as well when starting from VO^{IV} than when starting from V^{III}. This suggests that it may be worth the effort for a mussel to expend energy reducing vanadium from its V^{V} state in seawater to V^{III} in exchange for better cohesive qualities. However, this remains to be verified as the oxidation state of the V within the metal storage particles is currently unknown. For mussels, we have begun to form a comprehensive answer as to why vanadium might be the preferred metal of use in the cuticle. Despite this, much work is left to be done on describing the uptake, storage, and redox transformations that are performed on vanadium ions before they are mixed with cuticle proteins during the secretion of threads.

5. Conclusions and Outlook

The work in the current thesis has clear relevance for the biological and biochemical understanding of the byssus cuticle biopolymer, but it is also highly relevant in terms of bio-inspiration for novel materials. Up to this point, aside from very few studies,[94] the majority of mussel-inspired metallopolymers have been based on DOPA-Fe interactions, and have taken a reductionist approach by taking DOPA out of the context of its natural protein environment. Our findings suggest that researchers in this field should consider the potential use of vanadium as a chelating metal to augment the dynamic properties of the metallopolymeric materials they develop. Additionally, they should consider more closely mimicking the protein sequence of mfp-1 to achieve improved chelating abilities under different conditions. Such modifications to synthetic polymers may confer benefits similar to what we observed in experimentation with rMfp-1-DOPA, such as chelation at acidic pHs and tris complex formation at low stoichiometries of DOPA to metal ions. While these results offer some beneficial guidance for future material design, they will also hopefully provide inspiration for continued exploration into the rapid, sustainable assembly and mechanical mysteries of byssal threads and their remarkable cuticles.

6. References

1. Wang, Y.; Naleway, S. E.; Wang, B., Biological and bioinspired materials: Structure leading to functional and mechanical performance. *Bioactive Materials* **2020,** *5* (4), 745-757.
2. Glazko, G. V.; Nei, M., Estimation of Divergence Times for Major Lineages of Primate Species. *Molecular Biology and Evolution* **2003,** *20* (3), 424-434.
3. Ashby, M. F., *Materials selection in mechanical design*. 4th ed.; Butterworth-Heinemann: Burlington, MA, 2011; p xv, 646 p.
4. Goodall, J., Tool-using and aimed throwing in a community of free-living chimpanzees. *Nature* **1964,** *201* (4926), 1264-1266.
5. Beaune, S. A. d.; Coolidge, F. L.; Wynn, T. G., *Cognitive archaeology and human evolution*. Cambridge University Press: New York, 2009; p xii, 185 p.
6. Van Schaik, C. P.; Fox, E. A.; Sitompul, A., Manufacture and use of tools in wild Sumatran orangutans. *Naturwissenschaften* **1996,** *83* (4), 186-188.
7. Emery, N. J., Cognitive ornithology: the evolution of avian intelligence. *Philosophical Transactions of the Royal Society B: Biological Sciences* **2006,** *361* (1465), 23-43.
8. Smolker, R.; Richards, A.; Connor, R.; Mann, J.; Berggren, P., Sponge carrying by dolphins (Delphinidae, Tursiops sp.): a foraging specialization involving tool use? *Ethology* **1997,** *103* (6), 454-465.
9. Tuttle, R. H. Human Evolution. https://www.britannica.com/science/human-evolution (accessed April 9th).
10. Asimov, I., *The Solar System and Back*. 1st ed.; Doubleday: Garden City, N.Y.,, 1970; p xiv, 246 p.
11. Knauth, P., *The Metalsmiths*. Time-Life Books: New York,, 1974; p 160 p.
12. Kaufman, B., Metallurgy and ecological change in the ancient Near East. *Backdirt, Annual Review of the Cotsen Institute of Archaeology* **2011,** *2011*, 86-92.
13. Waldbaum, J. C., *From Bronze to Iron : The Transition from the Bronze Age to the Iron Age in the Eastern Mediterranean*. P. Aström: Göteborg, 1978; p 106 p.
14. Forbes, R. J., *Man The Maker: A History of Technology and Engineering*. Abelard-Schuman: London, New York,, 1958; p 365 p.
15. Mackenzie, D. S., History of quenching. *International Heat Treatment and Surface Engineering* **2008,** *2* (2), 68-73.
16. Simon, E., Ueber den flüssigen Storax (Styrax liquidus). *Ann. Pharm* **1839,** *31* (3), 265-277.
17. Painter, P. C.; Coleman, M. M., *Essentials of Polymer Science and Engineering*. DEStech Publications, Incorporated: 2008.
18. Slack, C., *Noble obsession : Charles Goodyear, Thomas Hancock, and the race to unlock the greatest industrial secret of the nineteenth century*. 1st ed.; Hyperion: New York, 2002; p 274 p.

19. Leo Hendrick Baekland and the Innovation of Bakelite. https://www.acs.org/content/acs/en/education/whatischemistry/landmarks/bakelite.html (accessed May 16th).
20. Ndiaye, P. A.; Forster, E., *Nylon and Bombs: DuPont and the March of Modern America*. Johns Hopkins University Press: 2007.
21. Smith, E. M., *A man of ideas : the biography of Dr. Waldo Lonsbury Semon, inventor of plasticized polyvinyl chloride*. Geon Co.: Cleveland, Ohio (6100 Oak Tree Blvd., Cleveland 44131), 1993; p 59 p.
22. Geyer, R.; Jambeck, J. R.; Law, K. L., Production, use, and fate of all plastics ever made. *Science advances* **2017,** *3* (7), e1700782.
23. Ashby, M. F., *Materials selection in mechanical design*. 2nd ed.; Butterworth-Heinemann: Oxford, OX ; Boston, MA, 1999; p xii, 502 p.
24. Andrady, A. L.; Neal, M. A., Applications and societal benefits of plastics. *Philosophical Transactions of the Royal Society B: Biological Sciences* **2009,** *364* (1526), 1977-1984.
25. Thompson, R. C.; Moore, C. J.; Vom Saal, F. S.; Swan, S. H., Plastics, the environment and human health: current consensus and future trends. *Philosophical Transactions of the Royal Society B: Biological Sciences* **2009,** *364* (1526), 2153-2166.
26. Hermabessiere, L.; Dehaut, A.; Paul-Pont, I.; Lacroix, C.; Jezequel, R.; Soudant, P.; Duflos, G., Occurrence and effects of plastic additives on marine environments and organisms: A review. *Chemosphere* **2017,** *182*, 781-793.
27. Cole, M.; Lindeque, P.; Halsband, C.; Galloway, T. S., Microplastics as contaminants in the marine environment: a review. *Marine pollution bulletin* **2011,** *62* (12), 2588-2597.
28. Sanchez, C.; Arribart, H.; Guille, M. M. G., Biomimetism and bioinspiration as tools for the design of innovative materials and systems. *Nature materials* **2005,** *4* (4), 277-288.
29. Whitesides, G. M., Bioinspiration: something for everyone. *Interface focus* **2015,** *5* (4), 20150031.
30. Mora, C.; Tittensor, D. P.; Adl, S.; Simpson, A. G.; Worm, B., How many species are there on Earth and in the ocean? *PLoS biology* **2011,** *9* (8).
31. Lafuma, A.; Quéré, D., Superhydrophobic states. *Nature materials* **2003,** *2* (7), 457-460.
32. Neinhuis, C.; Barthlott, W., Characterization and distribution of water-repellent, self-cleaning plant surfaces. *Annals of botany* **1997,** *79* (6), 667-677.
33. Vorobyev, A.; Guo, C., Multifunctional surfaces produced by femtosecond laser pulses. *Journal of Applied Physics* **2015,** *117* (3), 033103.
34. Chung, K.; Yu, S.; Heo, C. J.; Shim, J. W.; Yang, S. M.; Han, M. G.; Lee, H. S.; Jin, Y.; Lee, S. Y.; Park, N., Flexible, angle-independent, structural color reflectors inspired by morpho butterfly wings. *Advanced Materials* **2012,** *24* (18), 2375-2379.

35. Pan, L.; Wang, F.; Cheng, Y.; Leow, W. R.; Zhang, Y.-W.; Wang, M.; Cai, P.; Ji, B.; Li, D.; Chen, X., A supertough electro-tendon based on spider silk composites. *Nature Communications* **2020,** *11* (1), 1-9.
36. Ensikat, H. J.; Ditsche-Kuru, P.; Neinhuis, C.; Barthlott, W., Superhydrophobicity in perfection: the outstanding properties of the lotus leaf. *Beilstein journal of nanotechnology* **2011,** *2* (1), 152-161.
37. Smaal, A. C., European mussel cultivation along the Atlantic coast: production status, problems and perspectives. In *Sustainable Increase of Marine Harvesting: Fundamental Mechanisms and New Concepts: Proceedings of the 1st Maricult Conference held in Trondheim, Norway, 25–28 June 2000*, Vadstein, O.; Olsen, Y., Eds. Springer Netherlands: Dordrecht, 2002; pp 89-98.
38. Britannica, T. E. o. E., Mussel. In *Encyclopaedia Britannica*, Encyclopaedia Britannica Inc.: 2019.
39. Yonge, C., On the primitive significance of the byssus in the Bivalvia and its effects in evolution. *Journal of the Marine Biological Association of the United Kingdom* **1962,** *42* (1), 113-125.
40. WAITE, J. H., Adhesion in byssally attached bivalves. *Biological Reviews* **1983,** *58* (2), 209-231.
41. Sun, C.; Waite, J. H., Mapping chemical gradients within and along a fibrous structural tissue, mussel byssal threads. *Journal of Biological Chemistry* **2005,** *280* (47), 39332-39336.
42. Holten-Andersen, N.; Fantner, G. E.; Hohlbauch, S.; Waite, J. H.; Zok, F. W., Protective coatings on extensible biofibres. *Nature materials* **2007,** *6* (9), 669-672.
43. Benedict, C. V.; Waite, J. H., Composition and ultrastructure of the byssus of Mytilus edulis. *Journal of morphology* **1986,** *189* (3), 261-270.
44. Priemel, T.; Degtyar, E.; Dean, M. N.; Harrington, M. J., Rapid self-assembly of complex biomolecular architectures during mussel byssus biofabrication. *Nature communications* **2017,** *8* (1), 1-12.
45. Zhao, H.; Waite, J. H., Linking adhesive and structural proteins in the attachment plaque of Mytilus californianus. *Journal of Biological Chemistry* **2006,** *281* (36), 26150-26158.
46. Qin, X.; Waite, J. H., Exotic collagen gradients in the byssus of the mussel Mytilus edulis. *Journal of Experimental Biology* **1995,** *198* (3), 633-644.
47. Coyne, K. J.; Qin, X.-X.; Waite, J. H., Extensible collagen in mussel byssus: a natural block copolymer. *Science* **1997,** *277* (5333), 1830-1832.
48. Qin, X.-X.; Coyne, K. J.; Waite, J. H., Tough tendons mussel byssus has collagen with silk-like domains. *Journal of Biological Chemistry* **1997,** *272* (51), 32623-32627.
49. DeMartini, D. G.; Errico, J. M.; Sjoestroem, S.; Fenster, A.; Waite, J. H., A cohort of new adhesive proteins identified from transcriptomic analysis of mussel foot glands. *Journal of The Royal Society Interface* **2017,** *14* (131), 20170151.

50. Waite, J. H.; Qin, X.-X.; Coyne, K. J., The peculiar collagens of mussel byssus. *Matrix Biology* **1998,** *17* (2), 93-106.
51. Waite, J. H.; Vaccaro, E.; Sun, C.; Lucas, J. M., Elastomeric gradients: a hedge against stress concentration in marine holdfasts? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **2002,** *357* (1418), 143-153.
52. Waite, J. H.; Lichtenegger, H. C.; Stucky, G. D.; Hansma, P., Exploring molecular and mechanical gradients in structural bioscaffolds. *Biochemistry* **2004,** *43* (24), 7653-7662.
53. Harrington, M. J.; Waite, J. H., Holdfast heroics: comparing the molecular and mechanical properties of Mytilus californianus byssal threads. *Journal of Experimental Biology* **2007,** *210* (24), 4307-4318.
54. Sagert, J.; Waite, J. H., Hyperunstable matrix proteins in the byssus of Mytilus galloprovincialis. *Journal of Experimental Biology* **2009,** *212* (14), 2224-2236.
55. Lee, B. P.; Messersmith, P. B.; Israelachvili, J. N.; Waite, J. H., Mussel-inspired adhesives and coatings. *Annual review of materials research* **2011,** *41*, 99-132.
56. Waite, J. H., Mussel adhesion–essential footwork. *Journal of Experimental Biology* **2017,** *220* (4), 517-530.
57. Rzepecki, L. M.; Hansen, K. M.; Waite, J. H., Characterization of a cystine-rich polyphenolic protein family from the blue mussel Mytilus edulis L. *The Biological Bulletin* **1992,** *183* (1), 123-137.
58. Zhao, H.; Waite, J. H., Proteins in load-bearing junctions: the histidine-rich metal-binding protein of mussel byssus. *Biochemistry* **2006,** *45* (47), 14223-14231.
59. Waite, J. H.; Housley, T. J.; Tanzer, M. L., Peptide repeats in a mussel glue protein: theme and variations. *Biochemistry* **1985,** *24* (19), 5010-5014.
60. Valois, E.; Hoffman, C.; Demartini, D. G.; Waite, J. H., The Thiol-Rich Interlayer in the Shell/Core Architecture of Mussel Byssal Threads. *Langmuir* **2019,** *35* (48), 15985-15991.
61. Jehle, F.; Macías-Sánchez, E.; Fratzl, P.; Bertinetti, L.; Harrington, M. J., Hierarchically-structured metalloprotein composite coatings biofabricated from co-existing condensed liquid phases. *Nature communications* **2020,** *11* (1), 1-9.
62. Britannica, T. E. o. E., Levodopa. In *Encyclopaedia Britannica*, Encyclopaedia Britannica Inc: 2020.
63. Mayer, A. M., Polyphenol oxidases in plants and fungi: going places? A review. *Phytochemistry* **2006,** *67* (21), 2318-2331.
64. Lewis, E. A.; Tolman, W. B., Reactivity of dioxygen− copper systems. *Chemical reviews* **2004,** *104* (2), 1047-1076.
65. Deming, T. J., Mussel byssus and biomolecular materials. *Current opinion in chemical biology* **1999,** *3* (1), 100-105.

66. Sever, M. J.; Wilker, J. J., Absorption spectroscopy and binding constants for first-row transition metal complexes of a DOPA-containing peptide. *Dalton Transactions* **2006,** (6), 813-822.
67. Zhou, Y. Z.; Alany, R. G.; Chuang, V.; Wen, J., Studies of the rate constant of L-dopa oxidation and decarboxylation by HPLC. *Chromatographia* **2012,** *75* (11-12), 597-606.
68. Yang, J.; Saggiomo, V.; Velders, A.; Stuart, M.; Kamperman, M., Reaction Pathways in Catechol/Primary Amine Mixtures: A Window on Crosslinking Chemistry. *PLOS ONE* **2016,** *11*, e0166490.
69. Priemel, T.; Palia, R.; Babych, M.; Thibodeaux, C. J.; Bourgault, S.; Harrington, M. J., Compartmentalized processing of catechols during mussel byssus fabrication determines the destiny of DOPA. *Proceedings of the National Academy of Sciences* **2020,** *117* (14), 7613-7621.
70. Zechel, S.; Hager, M. D.; Priemel, T.; Harrington, M. J., Healing through histidine: Bioinspired pathways to self-healing polymers via imidazole–metal coordination. *Biomimetics* **2019,** *4* (1), 20.
71. Krauss, S.; Metzger, T. H.; Fratzl, P.; Harrington, M. J., Self-repair of a biological fiber guided by an ordered elastic framework. *Biomacromolecules* **2013,** *14* (5), 1520-1528.
72. Gosline, J.; Lillie, M.; Carrington, E.; Guerette, P.; Ortlepp, C.; Savage, K., Elastic proteins: biological roles and mechanical properties. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **2002,** *357* (1418), 121-132.
73. Carrington, E.; Gosline, J. M., Mechanical design of mussel byssus: load cycle and strain rate dependence. *American Malacological Bulletin* **2004,** *18* (1/2), 135-142.
74. Harrington, M. J.; Gupta, H. S.; Fratzl, P.; Waite, J. H., Collagen insulated from tensile damage by domains that unfold reversibly: In situ X-ray investigation of mechanical yield and damage repair in the mussel byssus. *Journal of structural biology* **2009,** *167* (1), 47-54.
75. Hagenau, A.; Papadopoulos, P.; Kremer, F.; Scheibel, T., Mussel collagen molecules with silk-like domains as load-bearing elements in distal byssal threads. *Journal of structural biology* **2011,** *175* (3), 339-347.
76. Schmidt, S.; Reinecke, A.; Wojcik, F.; Pussak, D.; Hartmann, L.; Harrington, M. J., Metal-mediated molecular self-healing in histidine-rich mussel peptides. *Biomacromolecules* **2014,** *15* (5), 1644-1652.
77. Suhre, M. H.; Gertz, M.; Steegborn, C.; Scheibel, T., Structural and functional features of a collagen-binding matrix protein from the mussel byssus. *Nature communications* **2014,** *5* (1), 1-11.
78. Martinez Rodriguez, N. R.; Das, S.; Kaufman, Y.; Israelachvili, J. N.; Waite, J. H., Interfacial pH during mussel adhesive plaque formation. *Biofouling* **2015,** *31* (2), 221-227.

79. Yu, J.; Wei, W.; Danner, E.; Ashley, R. K.; Israelachvili, J. N.; Waite, J. H., Mussel protein adhesion depends on interprotein thiol-mediated redox modulation. *Nature chemical biology* **2011,** *7* (9), 588-590.
80. Miller, D. R.; Spahn, J. E.; Waite, J. H., The staying power of adhesion-associated antioxidant activity in Mytilus californianus. *Journal of The Royal Society Interface* **2015,** *12* (111), 20150614.
81. Valois, E.; Mirshafian, R.; Waite, J. H., Phase-dependent redox insulation in mussel adhesion. *Science Advances* **2020,** *6* (23), eaaz6486.
82. Filippidi, E.; DeMartini, D. G.; Malo de Molina, P.; Danner, E. W.; Kim, J.; Helgeson, M. E.; Waite, J. H.; Valentine, M. T., The microscopic network structure of mussel (Mytilus) adhesive plaques. *Journal of The Royal Society Interface* **2015,** *12* (113), 20150827.
83. Raymond, K. N.; Dertz, E. A.; Kim, S. S., Enterobactin: an archetype for microbial iron transport. *Proceedings of the National Academy of Sciences* **2003,** *100* (7), 3584-3588.
84. Carrano, C. J.; Raymond, K. N., Ferric ion sequestering agents. 2. Kinetics and mechanism of iron removal from transferrin by enterobactin and synthetic tricatechols. *Journal of the American Chemical Society* **1979,** *101* (18), 5401-5404.
85. Kraemer, S. M., Iron oxide dissolution and solubility in the presence of siderophores. *Aquatic sciences* **2004,** *66* (1), 3-18.
86. Karpishin, T. B.; Dewey, T. M.; Raymond, K. N., Coordination chemistry of microbial iron transport. 49. The vanadium (IV) enterobactin complex: structural, spectroscopic, and electrochemical characterization. *Journal of the American Chemical Society* **1993,** *115* (5), 1842-1851.
87. Rohini; Baral, M.; Kanungo, B. K., Experimental and theoretical studies on structure, bonding and luminescence properties of Eu(III) and Tb(III) complexes of a new macrocyclic based 8HQ ligand. *Journal of Coordination Chemistry* **2019,** *72* (9), 1497-1523.
88. Ueki, T.; Yamaguchi, N.; Isago, Y.; Tanahashi, H., Vanadium accumulation in ascidians: a system overview. *Coordination Chemistry Reviews* **2015,** *301*, 300-308.
89. Oltz, E. M.; Bruening, R. C.; Smith, M. J.; Kustin, K.; Nakanishi, K., The tunichromes. A class of reducing blood pigments from sea squirts: isolation, structures, and vanadium chemistry. *Journal of the American Chemical Society* **1988,** *110* (18), 6162-6172.
90. Crans, D. C.; Smee, J. J.; Gaidamauskas, E.; Yang, L., The chemistry and biochemistry of vanadium and the biological activities exerted by vanadium compounds. *Chemical reviews* **2004,** *104* (2), 849-902.
91. Harrington, M. J.; Masic, A.; Holten-Andersen, N.; Waite, J. H.; Fratzl, P., Iron-clad fibers: a metal-based biological strategy for hard flexible coatings. *Science* **2010,** *328* (5975), 216-220.

92. Taylor, S. W.; Chase, D. B.; Emptage, M. H.; Nelson, M. J.; Waite, J. H., Ferric ion complexes of a DOPA-containing adhesive protein from Mytilus edulis. *Inorganic Chemistry* **1996,** *35* (26), 7572-7577.
93. Schmitt, C. N.; Winter, A.; Bertinetti, L.; Masic, A.; Strauch, P.; Harrington, M. J., Mechanical homeostasis of a DOPA-enriched biological coating from mussels in response to metal variation. *Journal of The Royal Society Interface* **2015,** *12* (110), 20150466.
94. Holten-Andersen, N.; Jaishankar, A.; Harrington, M. J.; Fullenkamp, D. E.; DiMarco, G.; He, L.; McKinley, G. H.; Messersmith, P. B.; Lee, K. Y. C., Metal-coordination: using one of nature's tricks to control soft material mechanics. *Journal of Materials Chemistry B* **2014,** *2* (17), 2467-2472.
95. Sever, M. J.; Wilker, J. J., Visible absorption spectra of metal–catecholate and metal–tironate complexes. *Dalton transactions* **2004,** (7), 1061-1072.
96. Sever, M. J.; Weisser, J. T.; Monahan, J.; Srinivasan, S.; Wilker, J. J., Metal-mediated cross-linking in the generation of a marine-mussel adhesive. *Angewandte Chemie International Edition* **2004,** *43* (4), 448-450.
97. Neilands, J., Siderophores: structure and function of microbial iron transport compounds. *Journal of Biological Chemistry* **1995,** *270* (45), 26723-26726.
98. Raymond, K. N.; Isied, S. S.; Brown, L. D.; Fronczek, F. R.; Nibert, J. H., Coordination isomers of biological iron transport compounds. VI. Models of the enterobactin coordination site. A crystal field effect in the structure of potassium tris (catecholato) chromate (III) and-ferrate (III) sesquihydrates, K3 [M (O2C6H4) 3]. cntdot. 1.5 H2O, M= chromium, iron. *Journal of the American Chemical Society* **1976,** *98* (7), 1767-1774.
99. Cooper, S. R.; Koh, Y. B.; Raymond, K. N., Synthetic, structural, and physical studies of bis (triethylammonium) tris (catecholato) vanadate (IV), potassium bis (catecholato) oxovanadate (IV), and potassium tris (catecholato) vanadate (III). *Journal of the American Chemical Society* **1982,** *104* (19), 5092-5102.
100. Branca, M.; Micera, G.; Dessi, A.; Sanna, D.; Raymond, K. N., Formation and structure of the tris (catecholato) vanadate (IV) complex in aqueous solution. *Inorganic Chemistry* **1990,** *29* (8), 1586-1589.
101. Pujol, J.-P., Formation of the byssus in the common mussel (Mytilus edulis L.). *Nature* **1967,** *214* (5084), 204-205.
102. Zuccarello, L. V., The collagen gland of Mytilus galloprovincialis: an ultrastructural and cytochemical study on secretory granules. *Journal of ultrastructure research* **1980,** *73* (2), 135-147.
103. Tamarin, A.; Lewis, P.; Askey, J., The structure and formation of the byssus attachment plaque in Mytilus. *Journal of morphology* **1976,** *149* (2), 199-221.
104. Tamarin, A.; Keller, P., An ultrastructutal study of the byssal thread forming system in Mytilus. *Journal of ultrastructure research* **1972,** *40* (3-4), 401-416.

105. Zuccarello, L. V., Ultrastructural and cytochemical study on the enzyme gland of the foot of a mollusc. *Tissue and Cell* **1981,** *13* (4), 701-713.
106. Shin, Y.; Brangwynne, C. P., Liquid phase condensation in cell physiology and disease. *Science* **2017,** *357* (6357).
107. Weber, S. C., Sequence-encoded material properties dictate the structure and function of nuclear bodies. *Current opinion in cell biology* **2017,** *46*, 62-71.
108. Astoricchio, E.; Alfano, C.; Rajendran, L.; Temussi, P. A.; Pastore, A., The Wide World of Coacervates: From the Sea to Neurodegeneration. *Trends in biochemical sciences* **2020**.
109. Wang, R.; Ozsvar, J.; Yeo, G. C.; Weiss, A. S., Hierarchical assembly of elastin materials. *Current Opinion in Chemical Engineering* **2019,** *24*, 54-60.
110. Yang, B.; Jin, S.; Park, Y.; Jung, Y. M.; Cha, H. J., Coacervation of interfacial adhesive proteins for initial mussel adhesion to a wet surface. *Small* **2018,** *14* (52), 1803377.
111. Kim, S.; Yoo, H. Y.; Huang, J.; Lee, Y.; Park, S.; Park, Y.; Jin, S.; Jung, Y. M.; Zeng, H.; Hwang, D. S., Salt triggers the simple coacervation of an underwater adhesive when cations meet aromatic π electrons in seawater. *ACS nano* **2017,** *11* (7), 6764-6772.
112. Strommen, D. P.; Nakamoto, K., Resonance raman spectroscopy. *Journal of Chemical Education* **1977,** *54* (8), 474.
113. Antosiewicz, J. M.; Shugar, D., UV–Vis spectroscopy of tyrosine side-groups in studies of protein structure. Part 1: basic principles and properties of tyrosine chromophore. *Biophysical Reviews* **2016,** *8* (2), 151-161.
114. Sonntag, M. D.; Pozzi, E. A.; Jiang, N.; Hersam, M. C.; Van Duyne, R. P., Recent advances in tip-enhanced Raman spectroscopy. *The journal of physical chemistry letters* **2014,** *5* (18), 3125-3130.
115. Cialla-May, D.; Zheng, X.-S.; Weber, K.; Popp, J., Recent progress in surface-enhanced Raman spectroscopy for biological and biomedical applications: from cells to clinics. *Chemical Society Reviews* **2017,** *46* (13), 3945-3961.
116. Meyer, R.; Yao, X.; Deckert, V., Latest instrumental developments and bioanalytical applications in tip-enhanced Raman spectroscopy. *TrAC Trends in Analytical Chemistry* **2018,** *102*, 250-258.
117. Cowcher, D. P.; Deckert-Gaudig, T.; Brewster, V. L.; Ashton, L.; Deckert, V.; Goodacre, R., Detection of protein glycosylation using tip-enhanced Raman scattering. *Analytical chemistry* **2016,** *88* (4), 2105-2112.
118. Israelachvili, J. N.; Tabor, D., The measurement of van der Waals dispersion forces in the range 1.5 to 130 nm. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* **1972,** *331* (1584), 19-38.
119. Skoog, D. A.; Holler, F. J.; Crouch, S. R., *Principles of instrumental analysis*. 6th ed.; Thomson Brooks/Cole: Belmont, CA, 2007; p xv, 1039 p.

120. Harris, D. C.; Bertolucci, M. D., *Symmetry and spectroscopy : an introduction to vibrational and electronic spectroscopy*. Oxford University Press: New York, 1978; p xii, 550 p.
121. Oh, D. X.; Shin, S.; Yoo, H. Y.; Lim, C.; Hwang, D. S., Surface forces apparatus and its applications for nanomechanics of underwater adhesives. *Korean Journal of Chemical Engineering* **2014,** *31* (8), 1306-1315.
122. Goodman, B.; Raynor, J., Electron spin resonance of transition metal complexes. In *Advances in inorganic chemistry and radiochemistry*, Elsevier: 1970; Vol. 13, pp 135-362.
123. Berliner, L. J.; Reuben, J.; Krishna, N. R., *Biological magnetic resonance*. Plenum Press: New York, 1978; Vol. 3.
124. Hwang, D. S.; Zeng, H.; Masic, A.; Harrington, M. J.; Israelachvili, J. N.; Waite, J. H., Protein- and metal-dependent interactions of a prominent protein in mussel adhesive plaques. *Journal of biological chemistry* **2010,** *285* (33), 25850-25858.
125. Stoll, S.; Schweiger, A., EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *Journal of magnetic resonance* **2006,** *178* (1), 42-55.
126. Israelachvili, J.; Min, Y.; Akbulut, M.; Alig, A.; Carver, G.; Greene, W.; Kristiansen, K.; Meyer, E.; Pesika, N.; Rosenberg, K., Recent advances in the surface forces apparatus (SFA) technique. *Reports on Progress in Physics* **2010,** *73* (3), 036601.
127. Zeng, H.; Hwang, D. S.; Israelachvili, J. N.; Waite, J. H., Strong reversible Fe3+-mediated bridging between dopa-containing protein films in water. *Proceedings of the National Academy of Sciences* **2010,** *107* (29), 12850-12853.
128. Zhang, J.; Xiang, L.; Yan, B.; Zeng, H., Nanomechanics of Anion− Π Interaction in Aqueous Solution. *Journal of the American Chemical Society* **2020,** *142* (4), 1710-1714.
129. Johnson, K. L.; Kendall, K.; Roberts, a., Surface energy and the contact of elastic solids. *Proceedings of the royal society of London. A. mathematical and physical sciences* **1971,** *324* (1558), 301-313.
130. Israelachvili, J., Thin film studies using multiple-beam interferometry. *Journal of Colloid and Interface Science* **1973,** *44* (2), 259-272.
131. Faghihnejad, A.; Zeng, H., Hydrophobic interactions between polymer surfaces: using polystyrene as a model system. *Soft Matter* **2012,** *8* (9), 2746-2759.
132. Israelachvili, J. N., *Intermolecular and surface forces*. Academic press: 2011.
133. Zeng, H., *Polymer adhesion, friction, and lubrication*. John Wiley & Sons: 2013.
134. Eilers, P. H.; Boelens, H. F., Baseline correction with asymmetric least squares smoothing. *Leiden University Medical Centre Report* **2005,** *1* (1), 5.
135. Boelens, H. F.; Eilers, P. H.; Hankemeier, T., Sign constraints improve the detection of differences between complex spectral data sets: LC− IR as an example. *Analytical chemistry* **2005,** *77* (24), 7998-8007.

136. Bonhommeau, S.; Lecomte, S., Tip-enhanced Raman spectroscopy: A tool for nanoscale chemical and structural characterization of biomolecules. *ChemPhysChem* **2018,** *19* (1), 8-18.
137. Miura, T.; Takeuchi, H.; Harada, I., Tryptophan Raman bands sensitive to hydrogen bonding and side-chain conformation. *Journal of Raman spectroscopy* **1989,** *20* (10), 667-671.
138. Rodriguez, N. R. M.; Das, S.; Kaufman, Y.; Wei, W.; Israelachvili, J. N.; Waite, J. H., Mussel adhesive protein provides cohesive matrix for collagen type-1α. *Biomaterials* **2015,** *51*, 51-57.
139. Danner, E. W.; Kan, Y.; Hammer, M. U.; Israelachvili, J. N.; Waite, J. H., Adhesion of mussel foot protein Mefp-5 to mica: an underwater superglue. *Biochemistry* **2012,** *51* (33), 6511-6518.
140. Holten-Andersen, N.; Harrington, M. J.; Birkedal, H.; Lee, B. P.; Messersmith, P. B.; Lee, K. Y. C.; Waite, J. H., pH-induced metal-ligand cross-links inspired by mussel yield self-healing polymer networks with near-covalent elastic moduli. *Proceedings of the National Academy of Sciences* **2011,** *108* (7), 2651.
141. Díaz-Cruz, C.; Alonso Nuñez, G.; Espinoza-Gómez, H.; Flores-López, L. Z., Effect of molecular weight of PEG or PVA as reducing-stabilizing agent in the green synthesis of silver-nanoparticles. *European Polymer Journal* **2016,** *83*, 265-277.
142. Michaud-Soret, I.; Andersson, K. K.; Que Jr, L.; Haavik, J., Resonance Raman studies of catecholate and phenolate complexes of recombinant human tyrosine hydroxylase. *Biochemistry* **1995,** *34* (16), 5504-5510.
143. Hardcastle, F. D.; Wachs, I. E., Determination of vanadium-oxygen bond distances and bond orders by Raman spectroscopy. *The Journal of Physical Chemistry* **1991,** *95* (13), 5031-5041.
144. Campbell, R. F.; Freed, J. H., Slow-motional ESR spectra for vanadyl complexes and their model dependence. *The Journal of Physical Chemistry* **1980,** *84* (20), 2668-2680.
145. Park, J. P.; Song, I. T.; Lee, J.; Ryu, J. H.; Lee, Y.; Lee, H., Vanadyl–catecholamine hydrogels inspired by ascidians and mussels. *Chemistry of Materials* **2015,** *27* (1), 105-111.
146. Nicklisch, S. C.; Waite, J. H., Mini-review: The role of redox in Dopa-mediated marine adhesion. *Biofouling* **2012,** *28* (8), 865-877.
147. Maier, G. P.; Rapp, M. V.; Waite, J. H.; Israelachvili, J. N.; Butler, A., Adaptive synergy between catechol and lysine promotes wet adhesion by surface salt displacement. *Science* **2015,** *349* (6248), 628-632.
148. Hwang, D. S.; Waite, J. H., Three intrinsically unstructured mussel adhesive proteins, mfp-1, mfp-2, and mfp-3: Analysis by circular dichroism. *Protein Science* **2012,** *21* (11), 1689-1695.
149. Hayashi, T.; Mizuno, K.; Bächinger, H. P., Hydroxylation-induced stabilization of the collagen triple helix: Further characterization of peptides with 4 (R)-hydroxyproline in the Xaa position. *Journal of Biological Chemistry* **2003,** *278* (34), 32373-32379.

150. Harrington, M. J.; Fratzl, P., Natural Load-Bearing Protein Materials. *Progress in Materials Science* **2020**, 100767.
151. Ruggiero, M. T.; Sibik, J.; Orlando, R.; Zeitler, J. A.; Korter, T. M., Measuring the Elasticity of Poly-l-Proline Helices with Terahertz Spectroscopy. *Angewandte Chemie* **2016,** *128* (24), 6991-6995.
152. Isom, D. G.; Castañeda, C. A.; Cannon, B. R., Large shifts in pKa values of lysine residues buried inside a protein. *Proceedings of the National Academy of Sciences* **2011,** *108* (13), 5260-5265.
153. Shin, M.; Shin, J. Y.; Kim, K.; Yang, B.; Han, J. W.; Kim, N.-K.; Cha, H. J., The position of lysine controls the catechol-mediated surface adhesion and cohesion in underwater mussel adhesion. *Journal of colloid and interface science* **2020,** *563*, 168-176.
154. Román-Leshkov, Y.; Davis, M. E., Activation of carbonyl-containing molecules with solid Lewis acids in aqueous media. *Acs Catalysis* **2011,** *1* (11), 1566-1580.
155. Degtyar, E.; Harrington, M. J.; Politi, Y.; Fratzl, P., The mechanical role of metal ions in biogenic protein-based materials. *Angewandte Chemie International Edition* **2014,** *53* (45), 12026-12044.
156. Barber-Zucker, S.; Shaanan, B.; Zarivach, R., Transition metal binding selectivity in proteins and its correlation with the phylogenomic classification of the cation diffusion facilitator protein family. *Scientific reports* **2017,** *7* (1), 1-12.
157. Holmgren, S. K.; Bretscher, L. E.; Taylor, K. M.; Raines, R. T., A hyperstable collagen mimic. *Chemistry & biology* **1999,** *6* (2), 63-70.
158. Priemel, T.; Palia, G.; Förste, F.; Jehle, F.; Mantouvalou, I.; Zaslansky, P.; Bertinetti, L.; Harrington, M., *Microfluidic-like Fabrication of a Vanadium-Cured Bioadhesive by Mussels*. 2021.
159. Winter, J. M.; Moore, B. S., Exploring the chemistry and biology of vanadium-dependent haloperoxidases. *Journal of Biological Chemistry* **2009,** *284* (28), 18577-18581.
160. Butler, A.; Carrano, C. J., Coordination chemistry of vanadium in biological systems. *Coordination chemistry reviews* **1991,** *109* (1), 61-105.
161. Garner, C. D.; Armstrong, E. M.; Berry, R. E.; Beddoes, R. L.; Collison, D.; Cooney, J. J. A.; Ertok, S. N.; Helliwell, M., Investigations of amavadin. *Journal of Inorganic Biochemistry* **2000,** *80* (1-2), 17-20.
162. Michibata, H., The mechanism of accumulation of vanadium by ascidians: some progress towards an understanding of this unusual phenomenon. *Zoological science* **1996,** *13* (4), 489-502.
163. Michibata, H.; Uyama, T.; Ueki, T.; Kanamori, K., Vanadocytes, cells hold the key to resolving the highly selective accumulation and reduction of vanadium in ascidians. *Microscopy Research and Technique* **2002,** *56* (6), 421-434.

164. Ueki, T.; Fujie, M.; Satoh, N., Symbiotic bacteria associated with ascidian vanadium accumulation identified by 16S rRNA amplicon sequencing. *Marine genomics* **2019,** *43*, 33-42.
165. Fattorini, D.; Regoli, F., Hyper-accumulation of vanadium in polychaetes. In *Vanadium*, Springer: 2012; pp 73-92.

7. Appendix

7.1. TERS Data Analysis Code (MatLab)

7.1.1. Pre-Processing: Baseline Corrections

```
function z = ASLS(y, lambda, p)
      % Estimate baseline with asymmetric least squares
      m = length(y);
      D = diff(speye(m), 2);
      w = ones(m, 1);
      for it = 1:10
          W = spdiags(w, 0, m, m);
          C = chol(W + lambda * D' * D);
          z = C \ (C' \ (w .* y));
          w = p * (y > z) + (1 - p) * (y < z);
      end

end

function primaryASLS(lambda,p,titleprefix,limit)

      opts = delimitedTextImportOptions("NumVariables", 2);

      % Specify range and delimiter
      opts.DataLines = [1, Inf];
      opts.Delimiter = "\t";

      % Specify column names and types
      opts.VariableNames = ["Wavenumber", "Intensity"];
      opts.VariableTypes = ["double", "double"];
      opts.ExtraColumnsRule = "ignore";
      opts.EmptyLineRule = "read";

      for x = 1:limit
          number = sprintf('%04d',x);
          name = titleprefix + number + "_0001.txt.csv";

          temp = readmatrix(name);
          leastsquares = ASLS(temp(:,2),lambda,p);
          smoothed = temp(:,2) - leastsquares;
          values = horzcat(temp(:,1),smoothed);
```

```
        csvwrite(name,values);
        disp(name + " smoothed and saved...")
    end

    disp("Primary ASLS successfully applied to grid!")

end
```

7.1.2. Pre-Processing: Silicon Peak Removal

```
function removeSi(titleprefix,limit)

    opts = delimitedTextImportOptions("NumVariables", 2);

    % Specify range and delimiter
    opts.DataLines = [1, Inf];
    opts.Delimiter = "\t";

    % Specify column names and types
    opts.VariableNames = ["Wavenumber", "Intensity"];
    opts.VariableTypes = ["double", "double"];
    opts.ExtraColumnsRule = "ignore";
    opts.EmptyLineRule = "read";

    for x = 1:limit
        number = sprintf('%04d',x);
        name = titleprefix + number + "_0001.txt.csv";

        temp = readmatrix(name);

        for index = 46:57
            temp(index,2) = 0;
        end

        csvwrite(name,temp);
        disp(name + " Si Peak removed...")
    end

    disp("Principal silicon peak removed from grid!")
```

```
end
```

7.1.3. Pre-Processing: Row Normalization

```
function rownormalize(titleprefix,limit)

      opts = delimitedTextImportOptions("NumVariables", 2);

      % Specify range and delimiter
      opts.DataLines = [1, Inf];
      opts.Delimiter = "\t";

      % Specify column names and types
      opts.VariableNames = ["Wavenumber", "Intensity"];
      opts.VariableTypes = ["double", "double"];
      opts.ExtraColumnsRule = "ignore";
      opts.EmptyLineRule = "read";

      for x = 1:limit
          number = sprintf('%04d',x);
          name = titleprefix + number + "_0001.txt.csv";

          temp = readmatrix(name);
          totalsignal = sum(temp(:,2));

          for index = 1:size
              temp(index,2) = temp(index,2) / totalsignal;
          end

          csvwrite(name,temp);
          disp(name + " normalized...")
      end

      disp("Row normalization of grid complete!")

end
```

7.1.4. Smoothing and Interpolation

```
function interpolator(limit,points,titleprefix)

      opts = delimitedTextImportOptions("NumVariables", 2);

       % Specify range and delimiter
       opts.DataLines = [1, Inf];
       opts.Delimiter = "\t";

       % Specify column names and types
       opts.VariableNames = ["Wavenumber", "Intensity"];
       opts.VariableTypes = ["double", "double"];
       opts.ExtraColumnsRule = "ignore";
       opts.EmptyLineRule = "read";

       data_x = readmatrix('Wavenumbers.xlsx');
       xq = linspace(330.79,2077.9,points);

       for x = 1:limit
           number = sprintf('%04d',x);
           name = titleprefix + number + "_0001.txt.csv"
           temp = readmatrix(name);

           % Savitsky Golay
           yfilt = sgolayfilt(temp(:,2),2,13);

           % Interpolation of Desired Function
           F = griddedInterpolant(data_x,yfilt,'makima');
           vq = F(xq);
           values = horzcat(transpose(xq),transpose(vq));
           csvwrite(name,values);
           disp(name + " interpolated and saved...")

       end
       writematrix(transpose(xq),"Wavenumbers.xlsx");
end
```

7.1.5. PCA

```
if exist('data_y') == 0
```

```
data_y = readmatrix('grid2_formatteddata.csv');
end

if exist('data_x') == 0
data_x = readmatrix('Wavenumbers.xlsx');
end

mu = mean(data_y);
meandata_y = data_y - mu;
V = cov(data_y);
[COEFF,latent,explained] = pcacov(V);
score = meandata_y*COEFF;
variances = cumsum(latent)./sum(latent);
explained;
```

7.1.6. Scree Plots

```
yyaxis left
plot(explained)
xlim([0 25])
title('Scree Plot of Grid 2 TERS Variances')
ylabel('Percentage of Variance Explained / Component')
xlabel('Principal Components')
yyaxis right
plot(variances*100)
ylabel('Cumulative Percentage of Variance Explained')
```

7.1.7. Area Under the Curve

```
function [regionsindex,regionswaves] = makeregions(regionswanted)

     regionsindex = zeros(regionswanted-1,2);
     regionswaves = zeros(regionswanted-1,2);
     data_x = readmatrix('Wavenumbers.xlsx');

     regionbounds = transpose(round(linspace(1,512,regionswanted),0));

        for i = 1:(regionswanted-1)
            regionsindex(i,1) = regionbounds(i);
```

```
            regionsindex(i,2) = regionbounds(i+1);
            regionswaves(i,1) = data_x(regionbounds(i));
            regionswaves(i,2) = data_x(regionbounds(i+1));
        end
 end

 function AUC = regionintegrator(regionsindex,titleprefix)

      data_x = readmatrix('Wavenumbers.xlsx');
      data_y = readmatrix(titleprefix + "formatteddata.csv");

     % AUC variable holds AUC arranged with regions of interest in
      columns (i.e. Column 1 = Region 1, and rows holding the AUC of
      the regions for every spectrum)

      AUC = zeros(size(data_y,1),size(regionsindex,1));

         for i = 1:size(data_y,1)  %% rows of AUC, aka spectra
             for j = 1:size(regionsindex,1)   %% columns of AUC, aka
regions to integrate
                 AUC(i,j) =
trapz(data_x(regionsindex(j,1):(regionsindex(j,2))),data_y(i,regionsin
dex(j,1):regionsindex(j,2)));
             end
         end
end

boxplot(AUC,'labels',{'400-480','650-700','745-806','830-850','995-
1010','1600-1630'})
xlabel('Wavenumber Ranges Examined (cm^{-1})','Interpreter','tex')
ylabel('AUC (A.U.)')
ylim([0 0.6])
set(gca,'YTick',0:0.05:0.6)
title('AUC Boxplot for Grid 4')
grid on
```

7.1.8. Peak Fitting

```
function [master, mastercfits, comments] =
applypeakfit(center,plusminus,titleprefix)
```

```
            data_x = readmatrix('Wavenumbers.xlsx');
            data_y = readmatrix(titleprefix + 'formatteddata.csv');
            [~,maxind] = min(abs(data_x - (center + plusminus)));
            [~,minind] = min(abs(data_x - (center - plusminus)));
            windowdev = std(data_x(minind:maxind));

            master = zeros(size(data_y,1),5); % C1 = RSquare, C2 = fit
type, remaining = a b c for dominant gaussian
            mastercfits = cell(size(data_y,1),2); %C1 = cFit obj of
best gaussian, C2 = cFit of the dominant gaussian
            comments = strings(size(data_y,1),1); % Note of what
spectra got fitted or why they didn't get fitt

            for i = 1:size(data_y,1)
                [peaks,~] =
findpeaks(data_y(i,minind:maxind),data_x(minind:maxind));
                istring = sprintf('%d',i);

                if isempty(peaks) == 0  % if there are peaks present,
start iterating through gaussian fits
                    coeffs = zeros(4,12);
                    r2s = zeros(4,1);
                    fitobjects = cell(4,1);

                    for j = 1:4  %Iterate through fitting 1-4
gaussians
                        number = sprintf('%d',j);
                        fititeration = "gauss" + number;

                        try   %occasionally, this can throw an error
if the fit coefficeints reach infinite values. this will skip that
fit.
                            [tempfitcoeffs,tempr2,fitobject] =
createFits(data_x(minind:maxind),data_y(i,minind:maxind),
fititeration); %each fit is contained in a row
                        catch
                            continue
                        end

                        coeffs(j,1:size(tempfitcoeffs,2)) =
tempfitcoeffs; %% adds newest coefficients
```

```
                    r2s(j,1) = tempr2; %% adds newest rsquare
value
                    fitobjects{j,1} = fitobject;
                end

                %% Cleanup Section (checks various params to
ensure good fits)

                % R2 evaluation (and choosing which fit to use
further)

                maxr2 = 0;
                maxr2ind = 0;

                for index = 1:4 % chooses first gaussian fit that
has r2 > 0.99
                    if r2s(index,1) > 0.99
                        maxr2 = r2s(index,1);
                        maxr2ind = index;
                        break
                    end
                end

                for index = 1:4 % No r2 over 0.99, lowering
criterion to 0.95
                    if r2s(index,1) > 0.95
                        maxr2 = r2s(index,1);
                        maxr2ind = index;
                        break
                    end
                end

                if maxr2 == 0
                    comments(i,1) = "All RSquare values fall below
0.95";
                    disp("Spectrum" + istring + "rejected...")
                    continue
                end

                if maxr2 < 0.99 && maxr2 > 0.95
                    comments(i,1) = "Warning: Low Rsquare (0.95 -
0.99)";
```

```
                    end

                    % b evaluation (checks to see if fitted peaks fall
within
                    % the fitting window

                    for index = 2:3:11 %iterates through any possible
b values and deletes that singular gaussian if it falls outside window
                        realb =
(coeffs(maxr2ind,index)*windowdev)+center;

                        if realb < (center - plusminus) || realb >
(center + plusminus)
                            coeffs(maxr2ind,index) = 0;
                            coeffs(maxr2ind,index-1) = 0;
                            coeffs(maxr2ind,index+1) = 0;
                        end
                    end

                    % a evaluation (checks to see that fitted gaussian
                    % meets a threshold intensity

                    threshold = 1*mean(data_y(i,minind:maxind));

                    for index = 1:3:10 % iterates through any possible
a values and deletes that singular gaussian if it doesn't meet
threshold

                        if coeffs(maxr2ind,index) < threshold ||
coeffs(maxr2ind,index) > max(data_y(i,minind:maxind))
                            coeffs(maxr2ind,index) = 0;
                            coeffs(maxr2ind,index+1) = 0;
                            coeffs(maxr2ind,index+2) = 0;
                        end
                    end

                    % check to see if there are still multiple peaks
left;
                    % pick the closest to center of range if there are
more
                    % than one!
```

```
                    peaksleft = 0;
                    counter = 0;
                    bvalues = zeros(4,1);

                    for index = 2:3:11 %checks if there is a b value,
if there is, it puts it in bvalues for later discrimination
                        counter = counter+1;
                        if coeffs(maxr2ind,index) ~= 0
                            peaksleft = 1;
                            bvalues(counter,1) =
coeffs(maxr2ind,index);
                        end
                    end

                    if peaksleft == 0
                        comments(i,1) = "Unsuitable a or b
parameters";
                        disp("Spectrum" + istring + "rejected...")
                        continue
                    end

                    if peaksleft ~= 0  % Here we pick the final
gaussian,and write it to master. Send R2, fitnumber, and a,b,c for
dominant gaussian
                        distance = inf;
                        finalindex = 0;

                        for index = 1:4
                            tempdistance = abs(bvalues(index,1));

                            if bvalues(index,1) ~= 0 && tempdistance <
distance
                                distance = tempdistance;
                                finalindex = index-1;
                            end
                        end

                        master(i,1) = maxr2; %highest r2
                        master(i,2) = maxr2ind; %how many gaussians
produced this r2
                        master(i,3) =
coeffs(maxr2ind,(3*finalindex+1)); %a param
```

```
                        master(i,4) =
coeffs(maxr2ind,(3*finalindex+2)); %b param
                        master(i,5) =
coeffs(maxr2ind,(3*finalindex+3)); %c param

                        center1 = master(i,4)*windowdev + center;
%fit a new gaussian based on best previous gaussian left
                        [~,maxind1] = min(abs(data_x - (center1 +
10)));
                        [~,minind1] = min(abs(data_x - (center1 -
10)));
                        windowdev1 = std(data_x(minind1:maxind1));

                        [~,tempr2,fitobject] =
createFits(data_x(minind1:maxind1),data_y(i,minind1:maxind1),
'gauss1'); %if it doesn't fit well, scrap it

                        if tempr2 > 0.8
                            realb1 = fitobject.b1*windowdev1 +
center1;

                            if realb1 < (center1 - 10) || realb1 >
(center + 10)
                                comments(i,1) = "Unsuitable a or b
parameters";
                                master(i,:) = 0;
                                disp("Spectrum" + istring +
"rejected...")
                                continue
                            end

                            mastercfits{i,1} = fitobjects{maxr2ind,1};
                            mastercfits{i,2} = fitobject;
                            disp("Spectrum" + istring + "sucessfully
fitted...")
                        end
                    end


                elseif isempty(peaks) ~= 0
                    comments(i,1) = "No peaks detected";
                end
```

```
            end
        end
```

7.1.9. Heatmaps

```
function [xarray,yarray] = generategrid(xdim,ydim)

xarray = [];
yarray = [];

for y = 1:ydim   %% Generates a xdim by ydim grid. Does not yet assign
colours.
    if mod(y,2) ~= 0
        for x = 1:xdim
            xarray = [xarray; x];
            yarray = [yarray; y];
        end

    else if mod(y,2) == 0
            for x = 1:xdim
                xarray = [xarray; xdim+1-x];
                yarray = [yarray; y];
            end
        end
    end
end
end

function plotheatmapTrp(xdim,ydim,xgrid,ygrid,ydata,waves,titleprefix)

palette = colormap(jet(255));  % Generate a color gradient

for j = 1:size(ydata,2)   %Find the max-min values of what is being
plotted.
    colours = [];
    [maxval,~] = max(ydata(:,j),[],'all','linear');
    [minval,~] = min(ydata(:,j),[],'all','linear');
    gradientindices = linspace(minval,maxval,255);

    for i = 1:size(ydata,1)     %Correlate signal strength to color
        [~,index]=min(abs(ydata(i,j)-gradientindices(:)));
```

```
        colours = [colours; palette(index,:)];
    end

    sz = 800;
    mkr = 's';
    scatter(xgrid,ygrid,sz,colours,mkr,'filled')
    xlabel('X Grid Position');
    set(gca,'FontSize',12);
    xlim([0 xdim+1]);
    xticks(0:5:xdim);
    ylabel('Y Grid Position');
    ylim([0 ydim+1]);
    yticks(0:5:ydim);
    figuretitle = titleprefix + " Heatmap for " + "Trp Peak";
    title(figuretitle,'Interpreter','tex');
    set(get(gca,'title'),'Position',[16 32 1],'FontSize',20);
    set(gcf,'position',[10,10,1000,950]);
    daspect([1,1,1]);
    cbar = colorbar('Position',...
        [0.876 0.110047846889952 0.0175872492442978
0.815789473684211],...
        'Ticks',[0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1],...

'TickLabels',{'0.0','','0.2','','0.4','','0.6','','0.8','','1.0'},...
        'FontWeight','bold',...
        'FontSize',12);
    cbar.Label.String = 'Peak height relative to largest Peak in
grid';

end
```

7.2. DOPA Content Calculation

First, the intensities from the deconvoluted mass spectrum are extracted, specifically those corresponding to masses corresponding to the following **Equation 1**. The intensities at these specific masses represent relative abundance of rMfp-1-DOPAs containing 1, 2, 3…24 DOPA residues generated from tyrosine.

$$Masses = 13616\frac{g}{mol} + 16x\frac{g}{mol} \quad ,0 \leq x \leq 24\ DOPAs \tag{1}$$

As each mass represents a different number of DOPA residues, the intensity at each mass must be corrected before using them to calculate the percent conversion. As an example, even if the signal at 13632 m/z and 13648 m/z have equivalent intensities, these equivalent intensities imply different amounts of DOPA (i.e. the 13648 m/z signal will contain twice the DOPA of the 13632 m/z signal). This correction is achieved via **Equation 2**.

$$IntCorr_x = Int_{13616+16x} \times \frac{x}{24} \quad , 0 \leq x \leq 24\ DOPAs \tag{2}$$

Corrected intensities in hand, the percent conversion can be calculated by dividing the summed corrected intensities by the total intensity summed across all masses and multiplying by 100 as shown in **Equation 3**.

$$Percent\ Conversion = \frac{\sum_{x=0}^{24} IntCorr_x}{Int_{Total}} \times 100 \tag{3}$$