Studies of Protein Stability by Hydrogen Exchange-Mass Spectrometry

By

Chamalee Demalgiriya Gamage

Submitted to the graduate degree program in Chemistry and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Chair: Dr. David Weis

Dr. Cindy Berrie

Dr. Teruna Siahaan

Dr. Roberto De Guzman

Dr. David Volkin

Date Defended: 10 May 2022

The dissertation committee for Chamalee Demalgiriya Gamage

certifies that this is the approved version of the following dissertation:

Studies of Protein Stability by Hydrogen Exchange-Mass Spectrometry

Chair: Dr. David Weis

Date Approved: 11 May 2022

Abstract

With the rapid growth in using proteins as therapeutics, studying the stability of therapeutic proteins has become an essential part in the manufacturing of therapeutics to ensure the quality, safety, and efficacy of a drug product. Proteins can get exposed to various stresses and undergo various modifications during their production and formulation to a final drug product. These stress conditions and modifications of therapeutic proteins can lead them to degrade causing potential risks to the safety and efficacy of a drug. Evaluating the stability of therapeutic proteins to understand product liabilities is important both in discovery and development phases of a drug. It is also important to employ analytical tools that can reliably predict stability of proteins to reduce long development times. Hydrogen Exchange-Mass Spectrometry (HX-MS) is increasingly being used in pharmaceutical industry for studying protein-protein, protein-ligand, and protein-receptor interactions. However, there is much potential for using HX-MS in a variety of applications in manufacturing and development of biotherapeutics. In this dissertation work, we explored the potential of using HX-MS in studying deamidation and aggregation of proteins which are considered as major degradation pathways of therapeutic proteins. In chapter 2 and 3, we assessed the ability of using HX-MS to predict protein deamidation and understand structural dynamics of mAbs leading to interfacial mediated agitation-induced aggregation, respectively. In chapter 4, we generated a panel of model proteins that will be used to evaluate detection limits in HX-MS measurements in characterization of protein higher-order structures such as comparability studies of biosimilars.

Estimating deamidation propensities of proteins is essential for predicting their long-term stabilities. However, predicting deamidation rates in folded proteins is challenging because higher-

order structure has a significant and unpredictable effect on deamidation. In chapter 2, we assessed the potential of using HX-MS to predict deamidation propensities of proteins because HX-MS would offer a rapid method compared to accelerated and long-term stability studies used for deamidation predictions. Therefore, we investigated the correlation between hydrogen exchange kinetics and deamidation kinetics at deamidation sites of wild-type and mutant MBP. We observed a power law correlation between hydrogen exchange and deamidation half-lives at the NG sites of MBP. This correlation demonstrates that HX-MS can be used to reliably and rapidly rank deamidation propensity in folded proteins.

Agitation-induced aggregation is a common issue that therapeutic protein formulations encounter during their manufacturing and transportation. Agitation increases exposure of proteins to the air-water interface and adsorption of proteins to the air-water interface induces aggregate formation. In chapter 3, we explored structural dynamics of three mAbs at the air-water interface using HX-MS to understand the structural changes of the mAbs that could lead to aggregate formation. Using three mAbs with variable propensities to form aggregates under agitation stress, we observed that all three mAbs underwent partial unfolding to variable. Among the three, the mAb which had the highest risk to aggregate unfolded extensively while showing cooperative unfolding in the Fab domain. The other two mAbs partially unfolded only in some regions of the molecules including some regions of the CDRs, but the unfolded regions were molecule specific. We also observe a protection against the partial unfolding of the molecules when the agitation stress was provided in the presence of polysorbate 20. This study provides direct evidence that local unfolding events resulting from interface exposure precede aggregate formation and may play a causal role in the process.

HX-MS has proven to be a reliable method for the assessment of protein higher-order structures. However, the limits of detection in the HX-MS measurements in detecting subtle differences in higher-order structures has not been tested rigorously. With the aim of fulfilling this need, as the first step, we generated a panel of model proteins with subtle structural changes and altered function in chapter 4. Here we expressed and purified a set of mutants of V_HH-F5 and a binding assay was optimized to identify the mutants with differences in relative binding affinities. The mutants with detuned binding affinities will be used to perform HX-MS experiments to identify their higher-order structural differences.

Acknowledgments

First of all, I would like to thank my advisor, Dr. David Weis, for his great mentorship and guidance throughout my graduate studies. I am forever grateful to him for his immense support to develop myself as a scientist. I am so thankful that he led me to the opportunity to do an internship at Genentech which gave me a huge experience and was very helpful for me to land a job in a pharmaceutical company. I am also thankful for the opportunity given to train in a collaborative lab to learn a new analytical technique which expanded my skill set. I highly appreciate the approach he used to help me improve my scientific writing. I also appreciate the support and encouragement he gave me to attend conferences and present my work. Dr. Weis was always very understanding and open to discussion which made it easy to work with him. Also, I am thankful for funding me for the majority of time in graduate school. I feel privileged to be trained and develop as a scientist under him. I always get inspired by his passion for science and work ethic.

I would also like to thank my dissertation committee, Dr. Cindy Berrie, Dr. Roberto De Guzman, Dr. Teruna Siahaan, Dr. David Volkin for taking time to serve on my committee. Moreover, I would also like to thank my orals comprehensive exam committee, Dr. Cindy Berrie, Dr. Michael Johnson, Dr. Robert Dunn, Dr. Roberto De Guzman, Dr. Krzysztof Kuczera, and Dr. Yong Zeng. I am immensely thankful to the Chemistry Department and the University of Kansas for accepting me to the graduate program and funding my graduate studies. My sincere thanks go to all the professors who conducted all the chemistry courses that I took.

Next, I would like to thank the past members of the Weis research group, Dr. Tyler Hageman, Dr. Mohammed Al-Naqshabandi, and Dr. Farai Rusinga, for training me on the lab protocols and instruments and supporting my work. I am also thankful to the recently graduated and current

members of the Weis group, Dr. Juan Rincon, Dr. Mihiri Weerasinghe, Yuqi Shi, Emranul Haq, Michael Wrigley, Sachini Karunaratne, and Hayley Blockinger, for their friendship and support in my graduate life. I would also like to thank the undergraduate students, Ashley Grande, Harlan Williams, and Katherine Vu, who worked in the Weis research group, for the experience I was able to gain by training and mentoring them and for the support they gave me for my work.

I am grateful to Benjamin Walters for giving me the opportunity to do an internship under him at Genentech. Thank you, Ben for being a mentor for me and providing continuous encouragement and support to successfully complete my internship. Also, I highly appreciate the enormous support you gave me in publishing my internship work and to include it my dissertation. The experience I had working with you helped me a lot to grow as a scientist. Your passion for science was an inspiration to me. I would also like to thank Genentech for the opportunity, amazing experience and funding me for the internship.

I would also like to express my gratitude to Prof. Veranja Karunaratne at SLINTEC, Sri Lanka, for helping me to apply to the University of Kansas graduate chemistry program and all my undergraduate professors at the University of Kelaniya, Sri Lanka, who made the path and guided me to graduate school.

Last, but not the least, I am forever in debt to my family, especially my parents, for providing me with a good education. Without their love, care, and constant support, I would not have achieved anything in my life. Also, I am fortunate to have a younger sister and a younger brother who always support me in achieving my goals in life. Dhanushka Weerasekara, my loving husband, I am so grateful to you than I could ever express in my words for everything you have done for me. Thank you for being so supportive and being the greatest strength behind me to get through all the hurdles

in graduate school and life. Thank you for believing in me! I am also thankful to my mother-in-law and brother-in-law for their support and encouragement for my graduate studies. Finally, I highly appreciate the support from the Sri Lankan community in Lawrence for making our time in Lawrence a happy memory.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ADCs - antibody-drug conjugates

API - active pharmaceutical ingredients

AUC - analytical ultracentrifugation

CD - circular dichroism

CDRs - complementarity-determining regions

CE - capillary electrophoresis

CID - collision-induced dissociation

DMPK - drug metabolism and pharmacokinetics

DSC - differential scanning calorimetry

DSC - differential scanning colorimetry

DTT - Dithiothreitol

ECD - electron capture dissociation

EDTA - ethylenediaminetetraacetic acid

EIC - extracted ion chromatograms

ELISA - enzyme-linked immunosorbent assay

ESI - Electrospray ionization

ETD - electron transfer dissociation

Fab - fragment antigen-binding

Fc - fragment crystallizable region

FPLC - Fast protein liquid chromatography

FTIR - fourier transform infrared spectroscopy

GdnHCl - guanidine hydrochloride

HIC - hydrophobic interaction chromatography

HMWS - high molecular weight species

HPLC - high-performance liquid chromatography

HRP - horseradish peroxidase

HX-MS – hydrogen exchange-mass spectrometry

IgG - Immunoglobulin G

IMS - ion mobility mass spectrometry

IPTG - isopropyl β-D-1-thiogalactopyranoside

ITC - isothermal titration calorimetry

K_d - dissociation constant

LB - Luria-Bertani

LC - liquid chromatography

LC-MS - liquid chromatography-mass spectrometry

LMWS – low molecular weight species

mAbs - monoclonal antibodies

MBP - maltose binding protein

MES - 2-(N-morpholino)ethanesulfonic acid

MS - mass spectrometry

MS/MS - tandem mass spectrometry

MWCO - molecular weight cut off

Ni-NTA - Ni^{2+}- nitrilotriacetic acid

NIST - National Institute of Standards and Technology

NMR - nuclear magnetic resonance spectroscopy

OD - optical density

PBS - phosphate-buffered saline

PDB - protein data bank

PS20 - polysorbate 20

PTMs - post-translational modifications

PVDF - polyvinylidene fluoride

QTOF - quadrupole time-of-flight

RF - radio frequency

RPLC - reversed-phase liquid chromatography

RTA - ricin toxin subunit A

SDS-PAGE - sodium dodecyl sulfate–polyacrylamide gel electrophoresis

SEC-MALS - size-exclusion chromatography coupled with multi-angle light scattering

SUMO - Small Ubiquitin-like Modifier

TCEP - tris (2-carboxyethyl) phosphine

TEV - tobacco etch virus

TFA - trifluoroacetic acid

TMB - 3,3’,5,5’-tetramethylbenzinidine peroxidase

TOF – time-of-flight

TPI - triosephosphate isomerase

Tris-HCl - Tris hydrochloride

UV - ultraviolet

V_HHs – variable antigen binding domain of single-chain camelid antibodies

1 CHAPTER 1: Introduction

1.1 Proteins as therapeutics

Protein-based therapeutics are a rapidly growing class of products in the pharmaceutical market since the production of first recombinant protein therapeutic, human insulin, in the 1970s. Based on the molecular type, therapeutic proteins can be categorized into groups such as antibodies and engineered antibody domains, Fc fusion proteins, enzymes, hormones, interferons, interleukins, growth factors, anticoagulants, blood factors, etc.[1] There were more than 300 biopharmaceuticals available in the United States and European Union market by 2018 including monoclonal antibodies, hormones, enzymes, clotting factors, blood related proteins, vaccines, and fusion proteins.[2-4] Out of the 53 FDA approved drugs in 2020, 13 were protein based drugs.[5] Proteins offer several advantages over small molecule drugs and these are high target specificity, less adverse effects and low toxicity, the ability to effectively replace gene therapy, and long-acting pharmacokinetics.[6, 7]

Among the recombinant protein therapeutics, monoclonal antibodies (mAbs) are increasingly becoming a dominant class of therapeutics to treat various diseases mainly in the areas of oncology, inflammation, and immune disorders.[8-10] Monoclonal antibodies accounted for 53% of the biopharmaceutical approvals in the US and EU from 2015 to 2018.[3] Antibodies are increasingly developed as therapeutics because their clinical potential can be improved by tuning their inherent properties. These properties include potential for reducing immunogenicity by developing humanized antibodies, high specificity and affinity for target antigen-binding, potential for increased biological potency and improved pharmacokinetics.[11] In addition to the mAb molecules, mAb related products such as antibody-drug conjugates (ADCs), bispecific antibodies, Fab fragments, Fc-fusion proteins, glycoengineered antibodies, antibodies fused to cytokines,

nanoantibody scaffolds are increasingly being developed and some of them are undergoing clinical trials..[12, 13]

Despite the increased and successful use of proteins as therapeutics, there are still many challenges to overcome in the development and use of therapeutic proteins. The complex structure of proteins can cause issues in therapeutic protein solubility, stability, and distribution in the body.[6] Therapeutic proteins are produced using specific host cell systems, mammalian cells being dominant, which can perform complex post-translational modifications (PTMs). PTMs such as glycosylation are required for many proteins to maintain their physiological activity, efficient secretion and stability.[14] However, some PTMs such as deamidation and oxidation can pose challenges because these modifications can affect therapeutic protein product stability and efficacy. Moreover, host systems may cause protein heterogeneity by altering the folding and aggregation tendency of proteins in large-scale production. These modifications of therapeutic proteins can provoke adverse immune responses in the body in some cases. However, adverse immune responses can be minimized by developing fully humanized antibodies.[14, 15]

1.2 Structure of proteins

Proteins are considered as the "building blocks of life" and play a central role in all biological processes. Proteins function as catalysts, transport and store other molecules, provide mechanical support, support the immune system, transmit nerve pulses, control growth and differentiation, and expression of genetic information. There is a relation between the function of a protein and its structure. Each protein has a specific function and a unique three-dimensional structure. A protein consists of a covalent backbone that contains many bonds capable of free rotation without breaking any covalent bonds. Therefore, a protein can have many possible conformations, the spatial

arrangement of the atoms of a protein. However, only a few conformations are predominant which are thermodynamically stable and critical for function. The stability of a native conformation of a protein depends on its noncovalent interactions such as hydrogen bonds, hydrophobic interactions, and ionic interactions which are involved in forming the three-dimensional structure. The structure of proteins can be described in four levels of organization: primary, secondary, tertiary, and quaternary structures.

1.2.1 Primary structure

The primary structure of a protein is the amino acid sequence of its polypeptide chain. The amino acids in a polypeptide chain are linked together by peptide bonds. Any disulfide bonds linking the amino acids are also included in the primary structure. Each protein consists of a distinctive number of amino acids and a distinctive sequence. The amino acid sequence determines how a protein will fold into its unique three-dimensional structure. The amino acid sequence of a protein is determined genetically. All proteins are made up of a set of twenty amino acids which vary in size, charge, shape, hydrophobicity, hydrogen-bonding capacity, and chemical reactivity. An amino acid contains a central carbon atom, called as α carbon, which is connected to a hydrogen, an amino group, a carboxylic group, and a R group or side chain. The R groups of the twenty amino acids are hydrogen (glycine), a methyl group (alanine), larger aliphatic side chains (valine, leucine, isoleucine, methionine, proline), aromatic side chains (phenylalanine, tyrosine, tryptophan, histidine), aliphatic hydroxyl groups (serine, threonine), aliphatic thiol group (cysteine), polar side chains (lysine, arginine), or acidic side chains (aspartic acid, glutamic acid, and their uncharged derivatives asparagine, glycine). Amino acids are linked together by the α-carboxyl group of one amino acid and the α-amino group of the next amino acid, forming a peptide

bond with the release of a water molecule. A series of amino acids linked by peptide bonds is called a polypeptide chain. Polypeptide chains of some proteins make cross links forming disulfide bonds by the oxidation of two cysteine residues.

The peptide bond has a planar geometry and characteristics of a double-bond which prevent free rotations at the peptide bond, which will provide some rigidity to protein globular structures. In addition, all peptide bonds in proteins, with few exceptions, are in *trans* configuration, in which the two α-carbon atoms are on the opposite side of the peptide bond, to avoid steric hindrance. In contrast, the bonds between the α-carbon and the amino and carbonyl groups are single, allowing rotations at these bonds. The rotations at these bonds are restricted only to sterically allowed rotations and will enable proteins to fold into thermodynamically favorable conformations.

1.2.2 Secondary structure

The secondary structure of a protein is the local special arrangement of the backbone atoms of its polypeptide chain. The secondary structures that widely occur in proteins are the α helix and the β sheet. In addition, the β turns and the omega (Ω) loops contribute to the formation of the secondary structure. In the α helix structure, polypeptide backbone is coiled around an imaginary longitudinal axis with the side chains of the amino acid residues extending outward of the helical backbone. Hydrogen bonds between the main chain CO and NH groups plays a critical role in stabilizing the α helix structure. The hydrogen bonds are formed between the NH group of an *n* th residue which is pointed along the α helix towards the CO group of the (*n*-4) th residue. In addition, van der Waals bonds are formed between the atoms in the core of the α helix. The α helices in all proteins are right-handed as the right-handed helix is more energetically favorable with less steric clashes between the backbone and the side chains. The β sheet structure is a more extended

structure composed of several β strands which are polypeptide chains arranged into a zigzag. β strands can arrange parallel (run in the same direction) or antiparallel (run in the opposite direction) to each other forming hydrogen bonds between them. The hydrogen bonds are formed between the NH groups and the CO groups of the adjacent β strands.

The secondary structural element β turns, also known as reverse turns, are observed where the polypeptide chain requires a reverse in the direction or as a connecting element between α helices or β sheets in globular proteins. A β turn involves four amino acid residues in a 180° turn where the CO group of the first amino acid residue forms hydrogen bonds with the NH group of the fourth amino acid residue. In addition, the Ω loops are also formed when reversals of the polypeptide chain occur. The loop structures have rigid, well defined more elongated structures. β turns and loops are often observed on the surface of a protein and are involved in forming hydrogen bonds with water.

1.2.3 Tertiary structure

The tertiary structure of a protein is referred to its three-dimensional arrangement of a polypeptide chain or how the folding of its secondary structure occurs. A polypeptide chain folds into its three-dimensional structure in a way that its nonpolar residues valine, leucine, isoleucine, methionine, and phenylalanine are mostly located in the interior of the protein while its charged polar residues arginine, histidine, lysine, aspartate, and glutamate are located on the surface of the protein. The uncharged polar groups serine, threonine, asparagine, glutamine, tyrosine, and tryptophan can occur either on the surface or interior of a protein. The tertiary structure of a protein is stabilized by hydrogen bonding between CO and NH groups of the residues in the interior of the protein and van der Waals interactions between tightly packed hydrophobic side chains.

Some large polypeptide chains fold into more than one compact globular units which are known as domains. A domain usually consists of 100-200 amino acid residues and the domains of a protein are connected together by flexible polypeptide segments. Domains can work as independent structural units and each domain can perform a specific function.

1.2.4 Quaternary structure

Some proteins are composed of more than one polypeptide chain which are known as subunits of a protein. The special arrangement of the subunits of a protein is referred to as its quaternary structure. The subunits are associated with each other through noncovalent interactions or disulfide bonds. Proteins can have two or more subunits and the simplest form is a dimer that consists of two subunits. A multi subunit protein is called as a multimer and a multi subunit protein with small number of subunits is called as an oligomer. Subunits of a protein can be identical or nonidentical. Association of multi subunits is important for a protein to perform its functions. Many multi subunit proteins perform regulatory functions where binding of small molecules alters the interactions between the subunits and changes the activity of the protein. Also, multi subunit proteins perform separate functions which are related, such as catalysis and regulation. Further, multi subunit proteins serve as reaction sites for complex multistep reactions such as protein synthesis that occurs at ribosomes which consists of many subunits.

1.2.5 Protein structure and function

Proteins are dynamic molecules whose functions are dependent on their three-dimensional structures and interactions with other molecules. Proteins are flexible and can change their conformation either subtly or dramatically. Subtle conformational changes occur due to molecular vibrations and small movements of amino acid residues. The dramatic conformational changes

occur when a large segment of a protein moves. Some of these conformational changes are essential for a protein to function. Many proteins function through binding with other molecules. When a protein binds with another molecule, the protein often changes its conformation for better fit and tight binding.

Proteins undergo structural changes when the conditions in their environment differs from the natural cellular environment. Changes in three-dimensional structure of proteins can cause the loss of their function. Proteins unfold/denature upon exposure to heat and extreme pH through destabilization of the weak interactions such as hydrogen bonds and electrostatic interactions. Organic solvents and some solutes such as urea and guanidine hydrochloride also denature proteins by disrupting the hydrophobic interactions in the interior of proteins. Denaturation of proteins exposes the hydrophobic interior of proteins to outside and forms aggregates. Proteins can also misfold in the cells and these misfolded proteins contribute to various diseases such as Alzheimer's disease and Parkinson's disease.

1.3 Stability of protein therapeutics

Therapeutic proteins are susceptible to degradation during their manufacturing, storage, and clinical use. Therapeutic protein manufacturing processes consists of many steps that can expose proteins to various types of stresses. In upstream processing, during cell culture and cell separation, proteins are exposed to stirring, air-water interfaces, elevated temperatures, and unfavorable pHs, and ionic strength variations. In addition, downstream processing (purification) exposes proteins to temperature and pH fluctuations, a wide range of ion and protein concentrations, various surfaces in containers and filtration membranes, freezing and thawing, and mechanical stresses. These stresses can destabilize proteins and cause both chemical and physical modifications in the

proteins.[16-18] Moreover, formulation, fill, and finish processes which converts a drug substance to a final dosage form can affect the stability of a protein. Operations such as freeze−thaw, mixing, filtration, and exposure to light and various manufacturing equipment surfaces can cause physical and chemical instability in therapeutic proteins.[19-21] Therefore, understanding protein degradation pathways and mechanisms thoroughly is essential in developing safe and efficacious drug products. Further, this will help develop formulations which can maintain acceptable chemical and physical stability of protein drug product during shipping, storage, and clinical use.

1.3.1 Chemical degradation of proteins

Chemical degradation of proteins occurs as covalent modifications to protein structure such as oxidation, deamidation, aspartate isomerization, hydrolysis, proteolysis, and disulfide bond shuffling.[22] Among these, oxidation is a common degradation pathway of proteins which leads to modifications in the primary and higher-order structure of proteins.[23] Oxidation can be induced by exposure of proteins to light, oxygen, redox-active metals, and some impurities (e.g., peroxide in the excipients used for formulations) during production, purification, formulation, transportation, storage, and handling.[21] Amino acid residues that are prone to oxidation are methionine, cysteine, histidine, phenylalanine, tryptophan, and tyrosine. Among these, methionine is the most susceptible residue to oxidation in the presence of oxygen. Oxidation of methionine increases the mass of a protein by 16 Da by producing methionine sulfoxide as the main oxidation product while increasing the polarity of the methionine side chain.[24, 25] Oxidation rate of methionine is dependent on the degree of exposure to the solvent. Therefore, methionine residues that are buried in the core of proteins have slow oxidation rates.[24] Cysteine is also a highly oxidation prone residue in protein which produces sulfinic acid, sulfonic acid, and di- and trisulfides as oxidation products. Oxidation

of histidine produces oxo-histidine and cross-linked adducts as photo-oxidation products. In addition, Tryptophan, tyrosine, and phenylalanine undergoes oxidation when a protein is exposed to light and oxidation of these residues gives rise to a complex mixture of products. Oxidation of tryptophan forms hydroxy-Trp, kynurenine (Kyn), 3-hydroxy-Kyn, N-formyl-kynurenine (NFK), 3-hydroxy-NFK. Oxidation products of tyrosine includes 3,4-Dihydroxyphenylalanine (DOPA) and 2-amino-3-(3,4-dioxocyclohexa-1,5-dien-1-yl) propanoicacid (DOCH).[26] Photodegradation by photooxidation is also a common degradation pathway for many proteins.[27] The structural changes caused by oxidation can in turn induce protein aggregation and thereby affect biological activity, pharmacokinetics, and immunogenicity of protein drugs.[22, 23, 26, 28]

Deamidation is another major chemical degradation reaction that occurs in proteins. Deamidation converts asparagine to aspartic acid and glutamine to glutamic acid. Deamidation of glutamine is less prevalent and only rarely observed. Asparagine deamidation forms a mixture of aspartic/*iso*-aspartic acid products at a ratio of 1:3 through the cyclic succinimide intermediate at basic pH. Similarly, glutamine deamidation forms a glutarimide cyclic intermediate yielding glutamic and *iso*-glutamic acids.[29, 30] Deamidation rates of proteins are primarily dependent on pH, temperature, and primary sequence and higher-order protein structure. Deamidation is accelerated at basic pH and higher temperatures. The primary sequence of a protein has a significant effect on deamidation propensity mainly the residue C-terminal to asparagine. The presence of an amino acid with a smaller side chain C-terminal to asparagine favors the cyclic intermediate formation due to lack of steric hindrance.[31] Therefore, asparagine-glycine is the most susceptible site for deamidation. In addition, asparagine followed by amino acid residues such as serine, histidine, and alanine can also have relatively higher propensities for deamidation. Further, higher order structure of proteins also plays an important role in deamidation rates. Asparagine sites that are located in

more ordered secondary structures as α-helices and β-sheets may have slower deamidation rates. The major impacts of deamidation on protein are changes in protein charge, lowering of pI, changes in primary sequence, heterogeneity due to isomer formation, and changes in higher-order protein conformation. All of these changes can destabilize proteins and thereby affect protein activity, half-life, aggregate formation propensity and immunogenicity.[22, 32] Deamidation of proteins is further discussed in chapter 2 of this dissertation.

Aspartate isomerization to *iso*-aspartate is another important chemical degradation pathway of proteins that affect protein structure and function.[33] *Iso*-aspartate can be formed directly via isomerization of aspartate or hydrolysis of succinimide intermediate from deamidation, following similar reaction mechanisms.[34] Aspartate isomerization is favored when an aspartate residue is followed by glycine, serine, or histidine residues, mostly in conformationally flexible regions of proteins.[35] The isomerization reaction is acid-catalyzed between pH 4-6.[36] Aspartate isomerization increases the heterogeneity of a protein sample by adding charge variants. [33]

1.3.2 Physical degradation of proteins

Physical degradation of proteins occurs as non-covalent modifications to proteins. The most common types of physical degradation include protein unfolding/denaturation, aggregation, and undesirable adsorption to surfaces.[37, 38] Physical degradation can be accelerated by the exposure of proteins to interfaces (e.g., air-liquid), light, and mechanical stress such as freeze-thaw, mixing, pumping, and shipping during biotherapeutic manufacturing processes.[37] During physical degradation, only the physical state of a protein undergoes changes without any chemical modifications. Denaturation can result in loss of three-dimensional or secondary structure of a protein. Denaturation of proteins can be caused by several factors including thermal stress, freezing

temperatures, and pressure. Among these, thermal stress is the most common stress that proteins undergo resulting in denaturation. Elevated temperatures may unfold proteins irreversibly leading to form aggregates.[16, 22]

Aggregation is the most prominent physical degradation pathway of proteins. Protein aggregates can have a potential risk of increasing immunogenicity of a drug.[39, 40] Moreover, aggregate formation can reduce the efficacy of a drug.[41] Aggregates can be formed at any stage of therapeutic protein production and formulation, long-term storage, shipping, and clinical use.[42] Protein aggregates can be formed in various forms such as soluble or insoluble, reversible or irreversible, or native/denatured.[43] Further, aggregates in solution can be present as oligomers, high molecular weight species, sub-visible and visible particles.[44, 45] There are many factors that induce aggregate formation during bioprocessing including high temperature, low pH, high protein concentrations, shear forces, formulation buffer composition (i.e. salts, excipients), stresses such as mechanical stress, freezing and thawing, pumping, long-term storage, exposure to interfaces (e.g., air-liquid, liquid-liquid, liquid-solid), and interactions with metal surfaces. [46-48] Concentration-dependent aggregation is considered as most challenging when developing highly concentrated drug formulations which is often the case in developing monoclonal antibody drugs for subcutaneous injection.[49]

1.4 Predicting protein stability

With the advancement of the techniques for recombinant protein generation, there can be more than one protein molecule identified having potential for a target therapy. Among these molecules, one lead candidate has to be identified in the drug discovery phase based on the success in technical development, long-term stability, biophysical properties, and DMPK properties.[50] In

addition, identifying product liabilities of candidate molecules such as stability and product heterogeneity will enable re-engineering the molecules or introducing changes to bioprocessing and formulation development processes to develop an ideal protein drug. An ideal therapeutic protein candidate is expected to have properties such as low product heterogeneity, acceptable levels of chemical and physical degradation, long-term stability, and consistent manufacturability. Therefore, a developability assessment is performed to identify potential liabilities of a protein candidate. First, *in silico* methods are used to predict potential degradation sites of candidate protein molecules. Then, stress studies are conducted to identify degradation sites and pathways and thereby identify stable candidates.[51]

1.4.1 Accelerated and forced degradation studies

As therapeutic proteins are susceptible to degradation from manufacture through clinical use, it is essential to gain a better understanding of various degradation pathways and identify degradation products to assess the intrinsic stability of protein molecules. The biopharmaceutical industry conducts forced degradation studies to investigate the degradation pathways of candidate molecules to help formulation development and identify critical quality attributes of a protein drug. Further, these studies are conducted at the early stages of drug development to select candidate molecules with most desirable properties and select the formulation conditions that maintains the drug product stability during the life cycle of the product. Accelerated and forced degradation studies are commonly used to investigate deamidation and oxidation and thermal, shaking, freeze-thaw, light, and pH stresses.[52, 53] To conduct a forced degradation study, a drug substance or a drug product is typically stored under extreme stress conditions for between less than a week to 3 months and data are collected for 3-5 time points. The stress conditions used for forced degradation

are relatively harsh compared to actual storage conditions and are able to generate degradation products in a short period of time.[54] Forced degradation studies are conducted to understand the degradation mechanisms of a protein candidate. On the other hand, accelerated stability studies are conducted at exaggerated storage conditions such as elevated temperature and humidity over a minimum of 6 months to increase degradation rates of a drug product. The data obtained from accelerated studies can be used to predict long-term effects at label storage conditions and to assess short-term changes that may occur during shipping.[55]

1.4.2 Long-term stability studies

Long-term stability studies are performed to establish the stability of a drug product over its proposed shelf-life. For active pharmaceutical ingredients (API) that have a proposed shelf-life of more than 12 months, it is recommended to evaluate the stability of the API at its long-term storage conditions. The time duration of long-term stability studies can typically vary from 6 months to several years and the frequency of testing typically occur every three months or six months.[56] The conditions that are commonly studied in long-term stability studies are effects of temperature, humidity, oxidation, photostability and susceptibility to hydrolysis of a drug product. The intention of performing long-term stability studies is to identify any degradation products during the proposed shelf-life of a drug. The information obtained from these studies are required to obtain the approval for a drug product from the regulatory agencies.[55, 56]

1.5 Therapeutic protein characterization

1.5.1 Characterization of higher-order structure of therapeutic proteins

With the increased growth of using proteins as therapeutics, there has been a need to employ variety of analytical tools to characterize proteins in the biopharmaceutical industry. Unlike small molecules, proteins have complex higher-order structures with unique characteristics. The higher-order structure and conformational dynamics of proteins are related directly to their biological function and physiochemical properties as therapeutics and thereby function, efficacy, and safety of a drug. Therefore, the higher-order structure of a therapeutic protein may be considered a critical quality attribute and critical characterization of higher-order structure of proteins may be required to ensure the quality of a therapeutic product.[57] Biotherapeutic manufacturing process starting from production in living cell systems to formulation, storage and clinical administration can introduce microheterogeneity to the protein product as various structural modifications, conformational changes, and non-covalent interactions. Therefore, it is important to employ robust and sensitive analytical techniques that can reliably characterize protein higher-order structure and conformational dynamics.[58] However, characterizing protein-higher order structure has been challenging due to the complex nature and behavior of proteins and limitations of the available biophysical techniques.[57]

There are various techniques available for characterizing protein higher-order structure. The biophysical techniques often used to assess the three-dimensional structure of proteins include circular dichroism (CD), fluorescence, ultraviolet (UV), differential scanning colorimetry (DSC), analytical ultracentrifugation (AUC), isothermal titration calorimetry (ITC), size-exclusion chromatography coupled with multi-angle light scattering (SEC-MALS), and Fourier transform

infrared spectroscopy (FTIR). These methods provide information on protein structure only at a global level, thus only low-resolution information, and are not capable of detecting small yet significant differences in protein structure.[59] However, more advanced techniques such as nuclear magnetic resonance spectroscopy (NMR) and X-ray crystallography provide high-resolution information on protein structure, but there are limitations associated with these techniques. X-ray crystallography requires the samples to be optimally crystalized, which can be difficult for some proteins, and provides only a static view of protein structure. Therefore, x-ray crystallography data may not well reflect protein dynamics in formulated drugs. NMR requires large sample amounts and has size limitations to the sample which is not suitable for large size therapeutic antibodies even though this method provides information on dynamic states of proteins. In addition, NMR analysis requires proteins to be isotopically labeled, which will cause spectral crowding for large antibodies and proteins in formulations. [60, 61] Moreover, both techniques require a pure protein sample for analysis, so that formulated protein drug products in the presence of excipients cannot be used.[60, 62] In addition to the above techniques, mass spectrometry is widely used for characterization of proteins because of the high sensitivity, specificity, and selectivity of the technique.

Mass spectrometry is used routinely for analyses of covalent structure and post-translational modifications (PTMs) of recombinant therapeutic proteins in biopharma. Moreover, mass spectrometry is used as an emerging reliable platform for protein higher-order structure analysis in the biopharmaceutical industry and also included in regulatory filings.[63] Among the mass spectrometry techniques, native mass spectrometry, ion mobility mass spectrometry (IMS), HX-MS, and chemical labeling and cross-linking methods are increasingly being employed for

characterization of protein higher-order structure with the capabilities of overcoming the challenges in classical biophysical techniques.[57, 58, 63]

1.5.2 Characterization of therapeutic proteins using mass spectrometry

Mass spectrometry (MS) is a technique that is vital in characterizing therapeutic proteins in biopharmaceutical industry. MS can be used to obtain a variety of information on proteins such as molecular weight, amino acid sequence, post-translational modifications, sequence variants, drug conjugation sites etc. Quadrupole time-of-flight (QTOF) and hybrid Orbitrap mass spectrometers are widely used for high-resolution and high-accuracy mass measurements of biopharmaceuticals. Electrospray ionization (ESI) is the major type of ionization source that is used for biopharmaceuticals.[13, 64] The fragmentation techniques that are used for performing tandem mass spectrometry (MS/MS) of proteins include collision-induced dissociation (CID), electron transfer dissociation (ETD), and electron capture dissociation (ECD). MS is typically coupled with liquid chromatography (LC) for biopharmaceutical analysis and reversed phase LC and hydrophobic interaction LC are the preferable modes.[64, 65]

Intact protein and middle-up MS measurements are performed to determine the molecular weight of proteins and large protein fragments such as light chain and heavy chain fragments of an antibody.[64] These measurements can be used to confirm sequence of a purified protein, presence of post-translational modifications, presence of glycoforms, identify mutations, study heterodimeric purity of bispecific antibodies, characterize antibody drug conjugates etc.[66-69] Moreover, top-down and middle-down measurements are also employed to obtain information on protein structure, sequence, and modifications. These approaches use ETD and ECD gas-phase

fragmentation modes to fragment proteins (top-down) and large protein fragments (middle-down).[13, 64]

Bottom-up MS measurements provide information at site specific peptide level and perform MS and MS/MS analysis of peptides. Peptides are generated by digesting proteins using enzymes such as trypsin, chymotrypsin, GluC, and LysC. In this approach, protein digests are separated typically by RPLC and analyzed using a high-resolution mass spectrometer like QTOF of Orbitrap. MS/MS analysis is performed by fragmenting the peptides commonly by CID to identify peptides and locate protein modification sites. In addition, ETD and ECD fragmentation techniques are also used to improve the sequence coverage.[13, 64, 65]

1.5.2.1 RP-HPLC of proteins and peptides

RPLC is commonly coupled with MS for characterization of proteins at intact and peptide level. In RPLC, proteins and peptides are separated based on their hydrophobicity using an organic solvent gradient on columns with a stationary phase of porous particles containing hydrophobic ligands such as C3-, C4-, C8-, C18-alkyl chains or phenyl and cyanopropyl groups.[65] Protein or peptide mixtures in aqueous buffer bind with the hydrophobic stationary phase of RP columns. The bound proteins and peptides are separated and eluted by passing a mobile phase containing an organic solvent such as acetonitrile. The elution can be performed as isocratic elution using constant percentage of organic solvent or as gradient elution by gradually increasing the strength of organic solvent over time. Therefore, proteins and peptides are eluted in the order of their increasing hydrophobicity. RP-HPLC has several advantages in the analysis of proteins and peptides including high resolution that can be achieved by changing chromatographic conditions, high selectivity achieved by changing the mobile phase, high recoveries and high reproducibility.[70]

Another advantage of RP-HPLC in protein analysis is that its compatibility with ESI-MS because of the volatility of the solvents used for the mobile phase.[65]

1.5.2.2 ESI of proteins and peptides

ESI is most commonly used for proteins because it is a soft ionization technique which can maintain the protein polypeptide backbone structure intact.[71] In addition, ESI is capable of producing multiply charged ions from large molecules with ionizable sites or even make ions from the molecules without any ionizable sites by forming adducts such as sodium, ammonium, and chloride.[72] Proteins and peptides are typically separated by HPLC prior to their MS analysis. Therefore, ESI transfers the solution phase protein and peptide ions into gas phase ions and directs towards the mass analyzer.[73] In ESI, gas phase ions are produced by passing the sample solution through a capillary tube with a strong applied electric field under atmospheric pressure. The electric field is generated by applying a 3-6 kV potential difference between the capillary and a counter electrode, which is at a distance of 0.3-2 cm. The electric field induces charges on the surface of the liquid at the end of the capillary and due to the large potential difference, the liquid is dispersed as a mist with highly charged droplets. Then the charged droplets pass a through a stream of heated inert gas, often nitrogen, where the evaporation of the solvent occurs. When the desolvation of the droplets occur, the charge per unit volume increases causing deformation of the droplets and breakage into smaller droplets. The solvent in these small droplets further evaporates and the resulting large electric field leads to desorption of ions/charged molecules from the surface.[72] The emitted ions in the gas phase are accelerated to the mass analyzer using skimmer cones.[73]

1.5.2.3 QTOF mass spectrometer

Quadrupole-time-of -flight mass spectrometers coupled with HPLC separation and ion fragmentation method CID are widely being used for protein characterization with high resolution, enhanced accuracy, and throughput.[13] QTOF is a hybrid mass analyzer which combines quadrupole and time-of-flight mass analyzers. The configuration of QTOF is set as a mass-resolving quadrupole, a collision cell and a TOF mass spectrometer. **Figure 1.1** shows a schematic diagram of a QTOF mass spectrometer. The advantages of this configuration are the capability of performing tandem mass analysis with high sensitivity, resolution, and mass accuracy for both precursor (MS) and product ions (MS/MS). The quadrupole can operate in two modes. The first is in radio frequency (RF) only mode which transmit all ions through the quadrupole to the TOF, where spectra are recorded, to obtain single MS measurements. Here, the collision cell also operates in the RF only mode. The second mode is as a mass filter which selects specific ions based on their mass-to-charge ratio to obtain MS/MS measurements.[74] In MS/MS mode, the collision cell fragments ions by collision induced dissociation (CID) which collides ions with a neutral gas, usually nitrogen or argon. The resulting fragment ions along with any unfragmented precursor ions are transmitted to the TOF analyzer. The ions leaving the collision cell are focused into a parallel beam using ion optics and accelerated into the ion modulator region of the TOF. In the ion modulator region, an electric field is applied at a frequency of several kHz which accelerates the ions orthogonal to their original direction. The ions then enter the flight tube which is a field free drift region where mass separation takes place. The ions travelling in the flight tube are reversed back in the direction of the ion modulator by the ion mirror at the opposite end of the drift tube and reaches the detector. The ion mirror will correct the energy and special spread of the ions in the flight tube.[75]

In the acceleration region, all the ions are accelerated by applying a potential of Vs, providing the ions an electric potential energy of E_{el}, which is converted to kinetic energy of E_k. Therefore, the energy of an ion with a mass of m and total charge of q=ze (z: charge, e: charge of an electron) can be expressed as:

$$E_k = \frac{1}{2}mv^2 = E_{el} = qV_s = zeV_s$$

Rearranging the above equation for velocity (v) of an ion gives:

$$v = \sqrt{\frac{2zeV_s}{m}}$$

As all the ions acquire the same kinetic energy by acceleration, smaller masses will have larger velocities while the larger masses will have smaller velocities.

The time required for an ion to travel the distance L for reaching the detector is at velocity v is given by:

$$t = \frac{L}{v}$$

Replacing v with its value in the previous equation and rearranging gives:

$$t^2 = \frac{m}{z}\left(\frac{L^2}{2eV_s}\right)$$

Therefore, mass-to-charge (m/z) of an ion can be determined by measuring the flight time (the time it takes to reach the detector) of that ion.[72, 76]

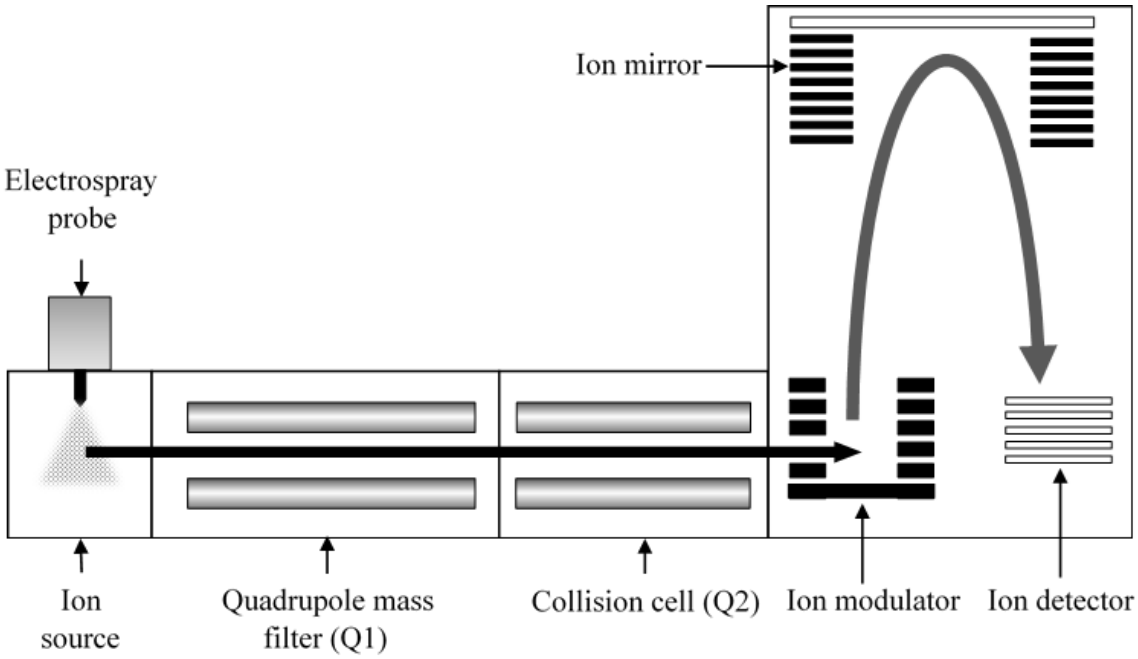


Figure 1.1: Schematic diagram of a QTOF mass spectrometer (adapted from references 75 and 76).

1.5.2.4 Tandem mass spectrometry of proteins and peptides

Tandem mass spectrometry (MS/MS) is a method that is used to acquire structural information, mainly sequence specific information, on proteins and peptides. The most common way of performing MS/MS is using the bottom-up approach of protein analysis. In this approach, proteins are digested using proteolytic enzymes and the resulting peptides are separated typically by HPLC. The separated peptides are ionized and mass-to-charge (m/z) values of the peptides are obtained, which is the initial MS analysis step. Then, a set of peptides are selected, known as precursor ions, and activated to produce fragments. The produced fragment ions, known as product ions, are identified by the second MS analysis step. Fragment/product ions are assigned to peptides

and matched to proteins using various algorithms.[71, 77] Both data acquisition and data interpretation in tandem mass spectrometry are automated due to large volumes of data produced.[78]

Collision induced dissociation (CID) is the most common ion activation method used for MS/MS of proteins. CID is performed by accelerating ions allowing collisions with a non-reactive gas such as helium, nitrogen, or argon in a collision chamber. The accumulation of internal energy in the ions facilitates their dissociation into fragments. In CID, fragment ions are formed by the cleavage at the peptide bonds between the C-N amide bonds. The fragment ions that retain the N-terminus are called b ions while the fragment ions that retain C-terminus are called y ions. In addition, fragment ions are numbered based on the number of N-terminal and C-terminal amino acids contained by each ion.[77, 78]

1.5.3 Characterization of therapeutic protein modifications: deamidation, oxidation

Studying the deamidation profile of a pharmaceutical product is required for its regulatory acceptance. Various techniques are used for identification and quantification of deamidation of proteins such as isoelectric focusing, capillary electrophoresis, ion exchange chromatography, liquid chromatography-mass spectrometry (LC-MS). Among these techniques, isoelectric focusing has disadvantages for quantification of deamidation due to its low sensitivity and specificity.[32] Capillary electrophoresis, particularly capillary zone electrophoresis, has been used to study deamidation, but the reproducibility of the technique has been an issue.[79, 80] Ion exchange-LC is also used to study deamidation. But this technique is also not suitable for quantification because of its higher detection limits.[81] CE and LC techniques can only be used to identify the charge variant produced in deamidation reaction. Therefore, the widely used technique for characterization of deamidation is MS coupled with LC.[32]

Mass spectrometry is capable of analyzing deamidation at site-specific level with high sensitivity and specificity. High-resolution MS can detect the mass increase of 0.98 Da caused by deamidation in proteins and identify deamidation sites with tandem mass spectrometry. Both top-down and bottom-up approaches of MS are used for studying deamidation of proteins.[82] In top-down approach, deamidated protein form is identified at intact protein level.[83] This approach has limitations with larger proteins and mostly suitable only for purified proteins or simple protein mixtures.[84] In the bottom-up approach, which is more commonly used, enzymatic digests of proteins are used to analyze deamidation at site specific level. The enzymes typically used for protein digestion are trypsin and Lys-C. Site specific deamidation is commonly analyzed using RPLC coupled to a high-resolution mass spectrometer with ESI. Deamidation sites can be identified by performing RPLC-MS/MS.[82, 85] Iso-aspartate containing peptides resulting from deamidation are identified by tandem MS based on specific fragments generated by ECD and ETD.[82]

Oxidation of methionine and tryptophan are commonly observed in therapeutic proteins and the oxidation product mostly observed such as methionine sulfoxide is considered as a critical quality attribute of therapeutic proteins. Therefore, tight monitoring of oxidation products in proteins is critical in biotherapeutic development.[86] oxidized products of therapeutic proteins are mainly characterized by peptide mapping and chromatographic approaches including hydrophobic interaction chromatography (HIC) and RP-HPLC coupled with MS.[87] Further, ion exchange chromatography has also been employed to study oxidative variants of therapeutic antibodies .[13] Hinterholzer et al. has shown the utility of NMR spectroscopy for identification and quantification of oxidation products in antibodies.[86] In addition to the classical LC-MS peptide mapping

approach, 18O labelling is also used for characterizing oxidation products with the advantage of avoiding artefactual oxidation that occurs in sample preparation for oxidation analysis.[88]

1.5.4 Hydrogen exchange mass spectrometry (HX-MS) in therapeutic protein characterization

HX-MS is one of the most promising methods to study protein structure, conformation, and dynamics.[89-93] HX-MS has gained increased popularity in the biopharmaceutical industry for therapeutic protein structural characterization.[94, 95] This method monitors the exchange of protein backbone amide hydrogens with deuterium in solution. Amides that are involved in hydrogen bonding and located in rigid structural regions in proteins exchange slowly while amides that are not protected from hydrogen bonding, are exposed to solvent, and are located in flexible regions of proteins exchange faster. Therefore, the differential exchange rates of backbone amide hydrogens provide information on protein conformation and dynamics.[96] The advantages of HX-MS over other protein structural characterization techniques are no size limitations to proteins to be analyzed, low sample requirement, ability to study proteins in their native state in formulated buffers, and advantages in sensitivity and sample preparation.[59, 61]

There are many applications of HX-MS in drug discovery and drug development. In drug discovery phase, HX-MS is used to study higher-order structural and conformational dynamics of protein targets and antibody and non-antibody therapeutic proteins in relation to their function, protein-ligand interactions to understand binding mechanisms, and protein-protein interactions to identify binding epitopes.[97, 98] In drug development, HX-MS is used to for comparability studies of biosimilars and reference products, studying effects of structural modifications such as post translational modifications (PTMs) on therapeutic protein structure, understanding the effects of

pharmaceutical excipients on conformational stability in formulation development, and for characterizing minor changes in high-order structure as a quality control step to manufacturing process changes.[97-100]

1.6 Hydrogen Exchange Mass Spectrometry (HX-MS)

1.6.1 Fundamentals of HX-MS

There are different types of hydrogens present in proteins including backbone amide hydrogens, carbon-bound hydrogens, and side-chain hydrogens. The heteroatom-bound hydrogens (O-H, N-H, and S-H) in proteins are exchange labile and exchange with the solvent due to the electronegativity of the heteroatoms. The hydrogens bound to carbon do not exchange in a measurable time scale.[101] The exchange labile hydrogens can be replaced with deuterium when a protein is exposed to D_2O-containing solvent. The replacement of hydrogen with deuterium results in an increase of the mass of the protein which can be monitored by MS. The exchangeable hydrogens in the side chains exchange rapidly, so that when replaced with deuterium, they back-exchange to hydrogens at a rapid rate under quench conditions and during LC separation of an HX-MS experiment. Therefore, exchange rates of side chain exchangeable hydrogens cannot be easily measured by MS. The backbone amide hydrogens which form the hydrogen-bond network of the protein secondary structure exchange at a rate which can be monitored by MS. Every amino acid residue except proline contains an amide hydrogen. Therefore, exchange rates of backbone amide hydrogens can be used to probe structural and dynamic features of proteins.[102, 103]

The exchange rate of amide hydrogens is primarily dependent on pH, temperature, solvent accessibility, and intramolecular hydrogen bonding.[102] Amide hydrogen exchange can be acid or base catalyzed. The exchange rate of an amide in an unstructured peptide is given by Equation 1:

$$k_{\mathrm{ch}} = k_{\mathrm{int,H}}[\mathrm{H^+}] + k_{\mathrm{int,OH}}[\mathrm{OH^-}] \quad (1)$$

where k_{ch} is the chemical exchange rate constant, $k_{\mathrm{int,H}}$ is the rate constant for acid-catalyzed reaction and $k_{\mathrm{int,OH}}$ is the rate constant for base-catalyzed reaction. The exchange reaction is primarily base-catalyzed above pH 4 and acid-catalyzed below pH 2. Between pH 2.5-3.0, the exchange rates of both acid- and base- catalyzed reactions are equal and the exchange rate goes to a minimum.[103] Most labeling experiments are performed at physiological or formulation pH which is suitable for proteins. The exchange reaction is quenched by reducing the pH to 2.5 where the exchange rate is minimum ($\mathrm{pH_{min}}$). Reducing the pH to $\mathrm{pH_{min}}$ decreases the exchange rate by four orders of magnitude which allows sufficient time for LC-MS analysis to be performed. Temperature has an exponential effect on chemical exchange rate constant. Temperature alters the water ionization constant and thereby the $\mathrm{OH^-}$ concentration and affects the chemical exchange rate. Therefore, deuterium labeling is carried out typically at 25 °C and then the temperature is decreased to 0 °C to quench the exchange reaction, which in turn slows down the chemical exchange rate by about 14-fold.[102] Decreasing the pH to $\mathrm{pH_{min}}$ and temperature to 0 °C enable reducing the back-exchange of deuterium labeled amide hydrogens to a minimum and retain the labeling until MS analysis.

In a folded protein, amide exchange rate is mostly governed by intramolecular hydrogen bonds and solvent accessibility of amide hydrogens.[96] The amide hydrogens that are located on the surface and in the flexible regions of a protein exchange rapidly while the amide hydrogens

located in more structured regions in the core of the protein exchange much slowly.[92] The exchange rates of the protected amides in the core of a protein are mediated by transient structural fluctuations of the protein which disrupt the intermolecular hydrogen bonds and provide access to the solvent.[96, 102, 104]

The Linderstrøm-Lang model explains how the exchange of amides in a folded protein occur through transient structural unfolding which is also known as breathing motions of a protein. The transient unfolding events disrupts the hydrogen bonding network of a protein and opens the protected amides to the solvent making them exchange competent as expressed by Equation 2:

$$\text{cl(H)} \underset{K_{cl}}{\overset{k_{op}}{\rightleftharpoons}} \text{op(H} \xrightarrow[D_2O]{k_{ch}} \text{op(D} \underset{k_{op}}{\overset{k_{cl}}{\rightleftharpoons}} \text{cl(D)} \qquad (2)$$

where cl represents the closed or folded form of an amide which is exchange-incompetent, and op represents the open or unfolded form which is exchange-competent. H and D refer to protiated and deuterated forms, respectively. k_{op} and k_{cl} are the unfolding/opening and refolding/closing rate constants, and k_{ch} is the chemical exchange rate constant of an amide hydrogen.[103] A protein in solution interconverts between the open and close states. Therefore, the isotopic exchange of a protected amide is governed by structural unfolding and refolding rates of a folded protein and the intrinsic exchange rate of an amide hydrogen.[102] Thus, the isotopic exchange rate constant (k_{HX}) of a protected amide can be given by Equation 3:

$$k_{\mathrm{HX}} = \frac{k_{\mathrm{op}} \times k_{\mathrm{ch}}}{k_{\mathrm{op}} + k_{\mathrm{cl}} + k_{\mathrm{ch}}} \qquad (3)$$

Under native conditions, proteins favor the folded state and refold rapidly to the folded state following a transient unfolding event making k_{cl} much faster than k_{op}.[102] Therefore, Equation 3 can be approximated to Equation 4:

$$k_{HX} = \frac{k_{op} \times k_{ch}}{k_{cl} + k_{ch}} \quad (4)$$

Two kinetic limits of exchange, referred to as EX1 and EX2 kinetics, can be defined considering the relative rates of the opening and closing events.[102] When the closing rate (k_{cl}) is much slower than the chemical exchange rate of an amide hydrogen (k_{ch}), it is defined as the EX1 kinetic limit. Here, the Equation 4 simplifies to Equation 5:

$$k_{HX} = k_{op}\ (k_{cl} \ll k_{ch}) \quad (5)$$

In this case, a single unfolding event allows for multiple amide hydrogens to exchange at the same time and complete the exchange in an unfolded region. However, under physiological conditions, the closing rate of most of the stable proteins is much faster relative to the chemical exchange rate of an amide hydrogen (k_{ch}). This kinetic regime is defined as EX2 kinetics and expressed by the Equation 6:

$$k_{HX} = \frac{k_{op}}{k_{cl}} \times k_{ch}\ (k_{cl} \gg k_{ch}) \quad (6)$$

Under EX2 kinetic regime, a protein undergoes many transient unfolding events before isotopic exchange is complete. EX1 and EX2 kinetics can be distinguished based on the mass spectral distributions of a protein as shown in **Figure 1.2**. EX1 kinetics result in bimodal mass distributions where low mass distribution represents the population of molecules that did not unfold, and the

high mass distribution represents the population of molecules that unfolded. EX2 kinetics results in binomial mass distribution which gradually increase in mass with time.[101, 102, 105-109]

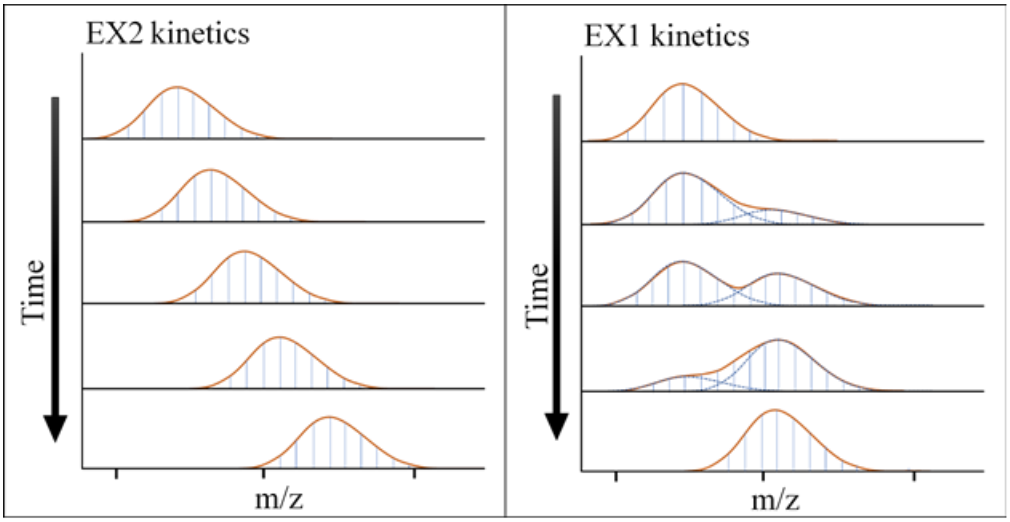


Figure 1.2: Mass spectra characteristic to EX2 kinetics (a) and EX1 kinetics (b).

1.6.2 HX-MS Experimental workflow

For an HX-MS experiment, the sample or the protein of interest is prepared in a buffer system. The buffer system used for sample preparation should be compatible with the mass spectrometer. The initial protein concentration of the sample should be high enough for subsequent dilution steps and the final concentration be sufficient to give a good signal for MS analysis. The first step in an HX experiment is to dilute the protein sample which is in H_2O buffer to a deuterated buffer which has the same composition as the aqueous buffer. A dilution of 10-20-fold is typically performed to have a >95% of deuterium concentration in the labeling solution.[110] The pD of the deuterated buffer has to be adjusted to match the pH of the aqueous buffer because glass pH electrodes read pH of a deuterated solution 0.4 units lower than an aqueous solution (pD =

$pH_{read} + 0.4$).[111] If the protein sample is lyophilized, the sample is directly reconstituted in deuterated buffer.

The most common way of carrying out deuterium labeling is the continuous labeling mode. **Figure 1.3** shows the experimental workflow of an HX-MS experiment following the continuous labeling mode. In this mode, the protein sample in aqueous buffer is diluted into an excess of identical deuterated buffer for specified time durations.[101] In a typical workflow, at least five time points are performed which can range from seconds to hours and three technical replicates are performed for each time point. Continuous labeling approach uses reaction conditions at which a protein maintain its native state, usually physiological pH. Therefore, the exchange rate monitored at a specific time point in a continuous labeling experiment is the average exchange of all dynamic sates of a protein under equilibrium conditions. At the completion of each labeling time, hydrogen exchange reaction is quenched by lowering the pH of the solution to pH 2.5 with the addition of a chilled acidic quench solution containing an acid such as formic acid or trifluoroacetic acid in an aqueous buffer or aqueous buffers such as phosphate or glycine. Reducing the pH of the reaction to pH 2.5 slows down the hydrogen exchange rates to a minimum and this will enable the sample to retain the exchanged deuterium for reproducible mass spectrometric analysis. The quenched sample is either analyzed immediately by mass spectrometry or stored at -80 °C for later analysis.[110]

Pulse labeling is the other mode of deuterium labeling, but this approach is not commonly applied. In pulse labeling, deuterium exchange is carried out for fixed time durations, usually a short period of time, and this time duration is constant over the experiment.[101, 112] A pulsed labeling experiment is initiated by adding a perturbant such as chaotropic agents or binding ligands or

changing temperature or pH. An aliquot of the reaction mixture is then taken out at predetermined time points and exposed to deuterium for the specified period of time, quenched, and analyzed by mass spectrometry. The addition of a perturbant may establish a new equilibrium population of molecules and the deuterium labeling will provide information on this new population.[110] Pulsed labeling is mostly used for protein folding studies and to study the changes takes place in a protein when a reaction proceeds.[101]

After labeling a protein, mass spectrometry analysis of deuterium incorporation can be carried out as global or local level analysis. At global level analysis, total amount of deuterium exchanged across the whole protein at a particular time point is measured. In local analysis, the amount of deuterium exchanged is measured at peptide level and this provides more localized information on hydrogen exchange with an increased resolution. For local analysis, the protein is digested using an acid protease which is optimally active at quench conditions, pH 2.5. Pepsin is the most commonly used acid protease for hydrogen exchange experiments.[101] Pepsin is a non-specific protease which cleaves at hydrophobic residues with a reproducible cleavage pattern, but the cleavage sites cannot be predicted based alone on protein sequence ahead of time.[113] Pepsin digestion can be carried out offline in solution phase or online using an immobilized pepsin column that is integrated to the LC system, under quench conditions, pH 2.5 and 0 °C.[114]

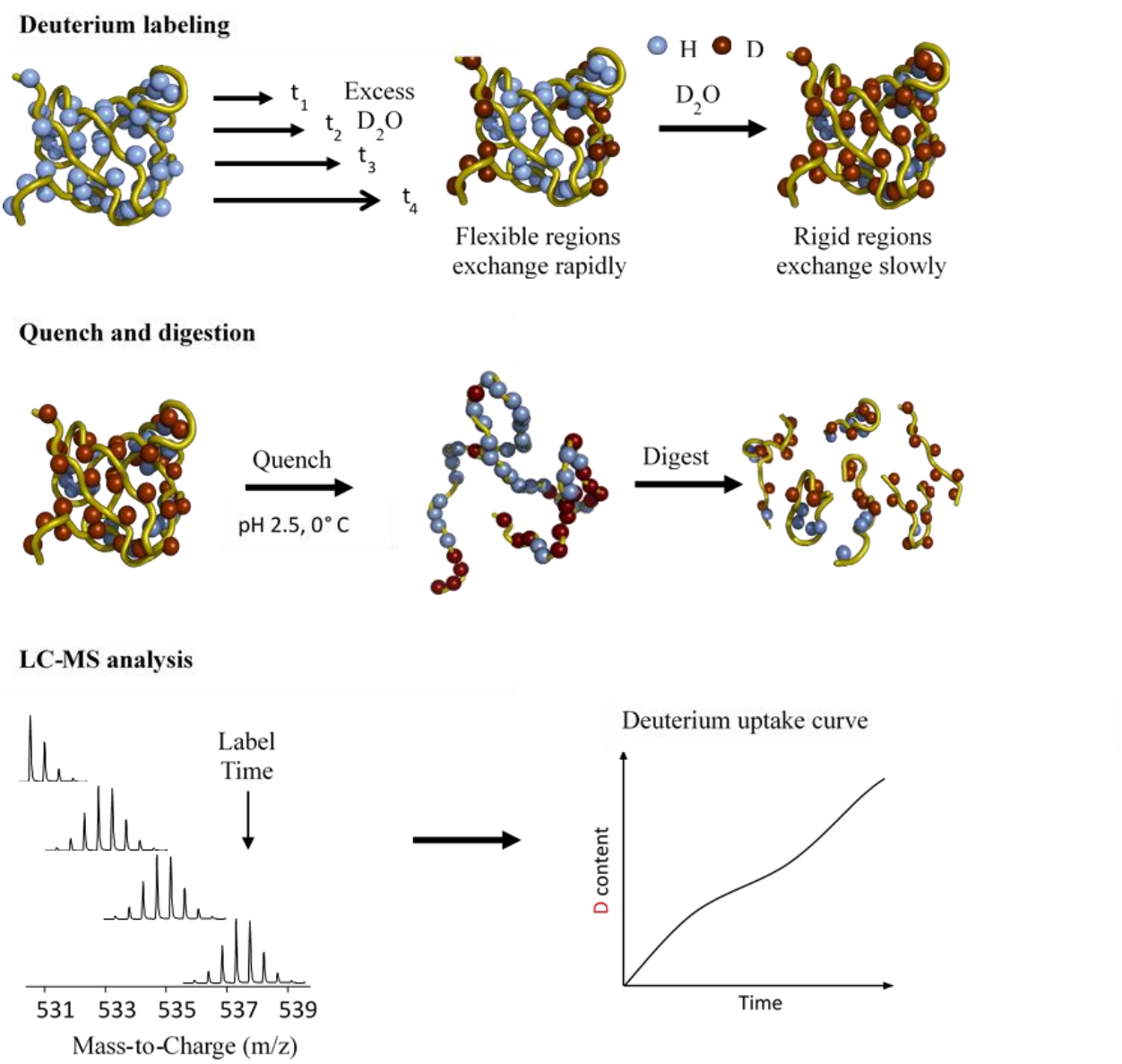


Figure 1.3: Schematic of HX-MS experiment workflow in the continuous labeling mode.

The next step is the chromatographic separation of the peptic peptides. Aqueous mobile phase of the LC solvent system can increase the exchange of the deuterium labeled amides back to hydrogen at a 100% H_2O environment. In order to prevent back-exchange, LC separation has to be carried out at quench conditions where the mobile phase is maintained at pH 2.5, typically

using formic acid and temperature at 0 °C using a refrigerated column compartment. Even at the quench conditions where the hydrogen exchange rate is minimum, the half-life of back-exchange ranges between 30-120 min.[110] Therefore, it is critical to complete the LC separation as short as possible while maintaining an optimum separation as well as minimum back-exchange. An acetonitrile gradient is usually used for elution of peptic peptides. Finally, the eluted peptides are analyzed by electrospray ionization-mass spectrometry (ESI-MS).[101, 110]

For data analysis, first, the peptides that result from pepsin digestion for a particular protein are identified by tandem (MS/MS) spectrometry separately.[71] These MS/MS identified peptides are used for HX-MS data analysis. The average mass of each peptide at each time point of HX labeling is determined based on the centroid masses of the MS spectral peak envelope of each peptide. Then the deuterium incorporation of each peptide is calculated by subtracting the average mass of undeuterated peptides from their corresponding deuterated peptides.[110] As HX-MS experiments typically generate a large amount of data, the data analysis is performed using software packages developed for HX-MS data analysis. For the interpretation and visualization of HX-MS data, deuterium uptake curves are used where the deuterium uptake of each peptide is plotted as a function of time. Here, typically the x-axis is time in log scale and the y-axis is average numbers of deuterium. In addition, deuterium uptake of all the peptides of a protein can be visualized on a plot where the x-axis displays each peptide from N- to C- terminal and y-axis displays average number of deuterium exchanged by each peptide. Further, heat maps are also used for data visualization which display relative exchange of deuterium on protein primary sequence using color codes to represent different deuteration levels. Also, differential exchange of deuterium can be mapped onto the homology model of the protein using similar color coding to a heat map.[115]

1.7 References

1. Dimitrov, D. S., Therapeutic proteins. *Methods Mol Biol* **2012,** *899*, 1-26.

2. Walsh, G., Biopharmaceutical benchmarks 2014. *Nat Biotechnol* **2014,** *32* (10), 992-1000.

3. Walsh, G., Biopharmaceutical benchmarks 2018. *Nat Biotechnol* **2018,** *36* (12), 1136-1145.

4. Lagassé, H. A.; Alexaki, A.; Simhadri, V. L.; Katagiri, N. H.; Jankowski, W.; Sauna, Z. E.; Kimchi-Sarfaty, C., Recent advances in (therapeutic protein) drug development. *F1000Res* **2017,** *6*, 113.

5. *Nature Reviews Drug Discovery 20* **2021**, 85-90.

6. Leader, B.; Baca, Q. J.; Golan, D. E., Protein therapeutics: a summary and pharmacological classification. *Nature Reviews Drug Discovery* **2008,** *7* (1), 21-39.

7. Imai, K.; Takaoka, A., Comparing antibody and small-molecule therapies for cancer. *Nature Reviews Cancer* **2006,** *6* (9), 714-727.

8. Elvin, J. G.; Couston, R. G.; van der Walle, C. F., Therapeutic antibodies: market considerations, disease targets and bioprocessing. *Int J Pharm* **2013,** *440* (1), 83-98.

9. Reichert, J. M.; Rosensweig, C. J.; Faden, L. B.; Dewitz, M. C., Monoclonal antibody successes in the clinic. *Nat Biotechnol* **2005,** *23* (9), 1073-8.

10. Kaplon, H.; Muralidharan, M.; Schneider, Z.; Reichert, J. M., Antibodies to watch in 2020. *MAbs* **2020,** *12* (1), 1703531.

11. Carter, P. J., Potent antibody therapeutics by design. *Nature Reviews Immunology* **2006,** *6* (5), 343-357.

12. Ueda, T., Next-generation optimized biotherapeutics — A review and preclinical study. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2014,** *1844* (11), 2053-2057.

13. Beck, A.; Wagner-Rousset, E.; Ayoub, D.; Van Dorsselaer, A.; Sanglier-Cianférani, S., Characterization of Therapeutic Antibodies and Related Products. *Analytical Chemistry* **2013,** *85* (2), 715-736.

14. Jenkins, N.; Murphy, L.; Tyther, R., Post-translational Modifications of Recombinant Proteins: Significance for Biopharmaceuticals. *Molecular Biotechnology* **2008,** *39* (2), 113-118.

15. Jenkins, N., Modifications of therapeutic proteins: challenges and prospects. *Cytotechnology* **2007,** *53* (1-3), 121-5.

16. Krause, M. E.; Sahin, E., Chemical and physical instabilities in manufacturing and storage of therapeutic proteins. *Curr Opin Biotechnol* **2019,** *60*, 159-167.

17. Wang, W.; Singh, S.; Zeng, D. L.; King, K.; Nema, S., Antibody structure, instability, and formulation. *J Pharm Sci* **2007,** *96* (1), 1-26.

18. Dengl, S.; Wehmer, M.; Hesse, F.; Lipsmeier, F.; Popp, O.; Lang, K., Aggregation and chemical modification of monoclonal antibodies under upstream processing conditions. *Pharm Res* **2013,** *30* (5), 1380-99.

19. Rathore, N.; Rajan, R. S., Current Perspectives on Stability of Protein Drug Products during Formulation, Fill and Finish Operations. *Biotechnology Progress* **2008,** *24* (3), 504-514.

20. Butreddy, A.; Janga, K. Y.; Ajjarapu, S.; Sarabu, S.; Dudhipala, N., Instability of therapeutic proteins — An overview of stresses, stabilization mechanisms and analytical techniques involved in lyophilized proteins. *International Journal of Biological Macromolecules* **2021,** *167*, 309-325.

21. Wang, W., Instability, stabilization, and formulation of liquid protein pharmaceuticals. *International Journal of Pharmaceutics* **1999,** *185* (2), 129-188.

22. Manning, M. C.; Chou, D. K.; Murphy, B. M.; Payne, R. W.; Katayama, D. S., Stability of protein pharmaceuticals: an update. *Pharm Res* **2010,** *27* (4), 544-75.

23. Torosantucci, R.; Schöneich, C.; Jiskoot, W., Oxidation of therapeutic proteins and peptides: structural and biological consequences. *Pharm Res* **2014,** *31* (3), 541-53.

24. Davies, M. J., The oxidative environment and protein damage. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2005,** *1703* (2), 93-109.

25. Vogt, W., Oxidation of methionyl residues in proteins: Tools, targets, and reversal. *Free Radical Biology and Medicine* **1995,** *18* (1), 93-105.

26. Grassi, L.; Cabrele, C., Susceptibility of protein therapeutics to spontaneous chemical modifications by oxidation, cyclization, and elimination reactions. *Amino Acids* **2019,** *51* (10), 1409-1431.

27. Schöneich, C., Photo-Degradation of Therapeutic Proteins: Mechanistic Aspects. *Pharmaceutical Research* **2020,** *37* (3), 45.

28. Liu, H.; Gaza-Bulseco, G.; Faldu, D.; Chumsae, C.; Sun, J., Heterogeneity of monoclonal antibodies. *J Pharm Sci* **2008,** *97* (7), 2426-47.

29. Riggs, D. L.; Silzel, J. W.; Lyon, Y. A.; Kang, A. S.; Julian, R. R., Analysis of Glutamine Deamidation: Products, Pathways, and Kinetics. *Analytical Chemistry* **2019,** *91* (20), 13032-13038.

30. Schroeter, E. R.; Cleland, T. P., Glutamine deamidation: an indicator of antiquity, or preservational quality? *Rapid Commun Mass Spectrom* **2016,** *30* (2), 251-5.

31. Robinson, N. E.; Robinson, A. B., *Molecular clocks: deamidation of asparaginyl and glutaminyl residues in peptides and proteins*. Althouse press: Cave Junction, OR, 2004.

32. Gervais, D., Protein deamidation in biopharmaceutical manufacture: understanding, control and impact. *Journal of Chemical Technology & Biotechnology* **2016,** *91* (3), 569-575.

33. Wakankar, A. A.; Borchardt, R. T., Formulation considerations for proteins susceptible to asparagine deamidation and aspartate isomerization. *Journal of Pharmaceutical Sciences* **2006,** *95* (11), 2321-2336.

34. Sreedhara, A.; Cordoba, A.; Zhu, Q.; Kwong, J.; Liu, J., Characterization of the Isomerization Products of Aspartate Residues at Two Different Sites in a Monoclonal Antibody. *Pharmaceutical Research* **2012,** *29* (1), 187-197.

35. Yang, H.; Zubarev, R. A., Mass spectrometric analysis of asparagine deamidation and aspartate isomerization in polypeptides. *ELECTROPHORESIS* **2010,** *31* (11), 1764-1772.

36. Oliyai, C.; Borchardt, R. T., Chemical pathways of peptide degradation. IV. Pathways, kinetics, and mechanism of degradation of an aspartyl residue in a model hexapeptide. *Pharm Res* **1993,** *10* (1), 95-102.

37. Rathore, N.; Rajan, R. S., Current perspectives on stability of protein drug products during formulation, fill and finish operations. *Biotechnol Prog* **2008,** *24* (3), 504-14.

38. Chi, E. Y.; Krishnan, S.; Randolph, T. W.; Carpenter, J. F., Physical Stability of Proteins in Aqueous Solution: Mechanism and Driving Forces in Nonnative Protein Aggregation. *Pharmaceutical Research* **2003,** *20* (9), 1325-1336.

39. Moussa, E. M.; Panchal, J. P.; Moorthy, B. S.; Blum, J. S.; Joubert, M. K.; Narhi, L. O.; Topp, E. M., Immunogenicity of Therapeutic Protein Aggregates. *J Pharm Sci* **2016,** *105* (2), 417-430.

40. Hermeling, S.; Crommelin, D. J. A.; Schellekens, H.; Jiskoot, W., Structure-Immunogenicity Relationships of Therapeutic Proteins. *Pharmaceutical Research* **2004,** *21* (6), 897-903.

41. Telikepalli, S. N.; Kumru, O. S.; Kalonia, C.; Esfandiary, R.; Joshi, S. B.; Middaugh, C. R.; Volkin, D. B., Structural Characterization of IgG1 mAb Aggregates and Particles Generated Under Various Stress Conditions. *Journal of Pharmaceutical Sciences* **2014,** *103* (3), 796-809.

42. Rosenberg, A. S., Effects of protein aggregates: An immunologic perspective. *The AAPS Journal* **2006,** *8* (3), E501-E507.

43. Cromwell, M. E. M.; Hilario, E.; Jacobson, F., Protein aggregation and bioprocessing. *The AAPS Journal* **2006,** *8* (3), E572-E579.

44. Vázquez-Rey, M.; Lang, D. A., Aggregates in monoclonal antibody manufacturing processes. *Biotechnology and Bioengineering* **2011,** *108* (7), 1494-1508.

45. Roberts, C. J.; Das, T. K.; Sahin, E., Predicting solution aggregation rates for therapeutic proteins: Approaches and challenges. *International Journal of Pharmaceutics* **2011,** *418* (2), 318-333.

46. Vázquez-Rey, M.; Lang, D. A., Aggregates in monoclonal antibody manufacturing processes. *Biotechnol Bioeng* **2011,** *108* (7), 1494-508.

47. Shah, M., Commentary: New perspectives on protein aggregation during Biopharmaceutical development. *Int J Pharm* **2018,** *552* (1-2), 1-6.

48. Mahler, H. C.; Friess, W.; Grauschopf, U.; Kiese, S., Protein aggregation: pathways, induction factors and analysis. *J Pharm Sci* **2009,** *98* (9), 2909-34.

49. Courtois, F.; Agrawal, N. J.; Lauer, T. M.; Trout, B. L., Rational design of therapeutic mAbs against aggregation through protein engineering and incorporation of glycosylation motifs applied to bevacizumab. *MAbs* **2016,** *8* (1), 99-112.

50. Hughes, J. P.; Rees, S.; Kalindjian, S. B.; Philpott, K. L., Principles of early drug discovery. *Br J Pharmacol* **2011,** *162* (6), 1239-1249.

51. Jarasch, A.; Koll, H.; Regula, J. T.; Bader, M.; Papadimitriou, A.; Kettenberger, H., Developability assessment during the selection of novel therapeutic antibodies. *J Pharm Sci* **2015,** *104* (6), 1885-1898.

52. Nowak, C.; J, K. C.; S, M. D.; Katiyar, A.; Bhat, R.; Sun, J.; Ponniah, G.; Neill, A.; Mason, B.; Beck, A.; Liu, H., Forced degradation of recombinant monoclonal antibodies: A practical guide. *MAbs* **2017,** *9* (8), 1217-1230.

53. Alsenaidy, M. A.; Jain, N. K.; Kim, J. H.; Middaugh, C. R.; Volkin, D. B., Protein comparability assessments and potential applicability of high throughput biophysical methods and data visualization tools to compare physical stability profiles. *Front Pharmacol* **2014,** *5*, 39.

54. Halley, J.; Chou, Y. R.; Cicchino, C.; Huang, M.; Sharma, V.; Tan, N. C.; Thakkar, S.; Zhou, L. L.; Al-Azzam, W.; Cornen, S.; Gauden, M.; Gu, Z.; Kar, S.; Lazar, A. C.; Mehndiratta, P.; Smith, J.; Sosic, Z.; Weisbach, P.; Stokes, E. S. E., An Industry Perspective on Forced

Degradation Studies of Biopharmaceuticals: Survey Outcome and Recommendations. *Journal of Pharmaceutical Sciences* **2020,** *109* (1), 6-21.

55. Q1A(R2) Stability Testing of New Drug Substances and Products. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/q1ar2-stability-testing-new-drug-substances-and-products (accessed 12.12.2021).

56. Stability testing of active pharmaceutical ingredients and finished pharmaceutical products. https://database.ich.org/sites/default/files/Q1F_Stability_Guideline_WHO_2018.pdf (accessed 12.12.2021).

57. Kaltashov, I. A.; Bobst, C. E.; Pawlowski, J.; Wang, G., Mass spectrometry-based methods in characterization of the higher order structure of protein therapeutics. *Journal of Pharmaceutical and Biomedical Analysis* **2020,** *184*, 113169.

58. Chen, G.; Warrack, B. M.; Goodenough, A. K.; Wei, H.; Wang-Iverson, D. B.; Tymiak, A. A., Characterization of protein therapeutics by mass spectrometry: recent developments and future directions. *Drug Discovery Today* **2011,** *16* (1), 58-64.

59. Houde, D.; Berkowitz, S. A.; Engen, J. R., The utility of hydrogen/deuterium exchange mass spectrometry in biopharmaceutical comparability studies. *J Pharm Sci* **2011,** *100* (6), 2071-86.

60. Gulerez, I. E.; Gehring, K., X-ray crystallography and NMR as tools for the study of protein tyrosine phosphatases. *Methods* **2014,** *65* (2), 175-183.

61. Burkitt, W.; Domann, P.; O'Connor, G., Conformational changes in oxidatively stressed monoclonal antibodies studied by hydrogen exchange mass spectrometry. *Protein Sci* **2010,** *19* (4), 826-835.

62. Wang, D.; Zhuo, Y.; Karfunkle, M.; Patil, S. M.; Smith, C. J.; Keire, D. A.; Chen, K., NMR Spectroscopy for Protein Higher Order Structure Similarity Assessment in Formulated Drug Products. *Molecules* **2021,** *26* (14), 4251.

63. Leurs, U.; Mistarz, U. H.; Rand, K. D., Getting to the core of protein pharmaceuticals – Comprehensive structure analysis by mass spectrometry. *European Journal of Pharmaceutics and Biopharmaceutics* **2015,** *93*, 95-109.

64. Fekete, S.; Guillarme, D.; Sandra, P.; Sandra, K., Chromatographic, Electrophoretic, and Mass Spectrometric Methods for the Analytical Characterization of Protein Biopharmaceuticals. *Analytical Chemistry* **2016,** *88* (1), 480-507.

65. Sandra, K.; Vandenheede, I.; Sandra, P., Modern chromatographic and mass spectrometric techniques for protein biopharmaceutical characterization. *Journal of Chromatography A* **2014,** *1335*, 81-103.

66. Reusch, D.; Haberger, M.; Maier, B.; Maier, M.; Kloseck, R.; Zimmermann, B.; Hook, M.; Szabo, Z.; Tep, S.; Wegstein, J.; Alt, N.; Bulau, P.; Wuhrer, M., Comparison of methods for the analysis of therapeutic immunoglobulin G Fc-glycosylation profiles—Part 1: Separation-based methods. *mAbs* **2015,** *7* (1), 167-179.

67. Zhang, T.; Huang, Y.; Chamberlain, S.; Romeo, T.; Zhu-Shimoni, J.; Hewitt, D.; Zhu, M.; Katta, V.; Mauger, B.; Kao, Y.-H. In *Identification of a single base-pair mutation of TAA (Stop codon)→ GAA (Glu) that causes light chain extension in a CHO cell derived IgG1*, MAbs, Taylor & Francis: 2012; pp 694-700.

68. Woods, R. J.; Xie, M. H.; Von Kreudenstein, T. S.; Ng, G. Y.; Dixit, S. B. In *LC-MS characterization and purity assessment of a prototype bispecific antibody*, MAbs, Taylor & Francis: 2013; pp 711-722.

69. Valliere-Douglass, J. F.; McFee, W. A.; Salas-Solano, O., Native intact mass determination of antibodies conjugated with monomethyl Auristatin E and F at interchain cysteine residues. *Analytical chemistry* **2012,** *84* (6), 2843-2849.

70. Aguilar, M.-I., Reversed-Phase High-Performance Liquid Chromatography. In *HPLC of Peptides and Proteins: Methods and Protocols*, Aguilar, M.-I., Ed. Springer New York: Totowa, NJ, 2004; pp 9-22.

71. Leurs, U.; Mistarz, U. H.; Rand, K. D., Getting to the core of protein pharmaceuticals--Comprehensive structure analysis by mass spectrometry. *Eur J Pharm Biopharm* **2015,** *93*, 95-109.

72. Hoffmann, E. d. S. V., *Mass spectrometry : principles and applications*. J. Wiley: Chichester, West Sussex, England; Hoboken, NJ, 2007.

73. Ho, C. S.; Lam, C. W. K.; Chan, M. H. M.; Cheung, R. C. K.; Law, L. K.; Lit, L. C. W.; Ng, K. F.; Suen, M. W. M.; Tai, H. L., Electrospray ionisation mass spectrometry: principles and clinical applications. *Clin Biochem Rev* **2003,** *24* (1), 3-12.

74. Chernushevich, I.; Loboda, A.; Thomson, B., An Introduction to quadrupole-time-of-flight mass spectrometry. *Journal of mass spectrometry : JMS* **2001,** *36*, 849-65.

75. Allen, D. R.; McWhinney, B. C., Quadrupole Time-of-Flight Mass Spectrometry: A Paradigm Shift in Toxicology Screening Applications. *Clin Biochem Rev* **2019,** *40* (3), 135-146.

76. Time-of-Flight Mass Spectrometry: Technical Overview. https://www.agilent.com/cs/library/technicaloverviews/Public/5990-9207EN.pdf (accessed 1-15-2022).

77. Brodbelt, J. S., Ion Activation Methods for Peptides and Proteins. *Analytical Chemistry* **2016,** *88* (1), 30-51.

78. Medzihradszky, K. F.; Chalkley, R. J., Lessons in de novo peptide sequencing by tandem mass spectrometry. *Mass Spectrom Rev* **2015,** *34* (1), 43-63.

79. He, Y.; Isele, C.; Hou, W.; Ruesch, M., Rapid analysis of charge variants of monoclonal antibodies with capillary zone electrophoresis in dynamically coated fused-silica capillary. *J Sep Sci* **2011,** *34* (5), 548-55.

80. Moritz, B.; Schnaible, V.; Kiessig, S.; Heyne, A.; Wild, M.; Finkler, C.; Christians, S.; Mueller, K.; Zhang, L.; Furuya, K.; Hassel, M.; Hamm, M.; Rustandi, R.; He, Y.; Solano, O. S.; Whitmore, C.; Park, S. A.; Hansen, D.; Santos, M.; Lies, M., Evaluation of capillary zone

electrophoresis for charge heterogeneity testing of monoclonal antibodies. *J Chromatogr B Analyt Technol Biomed Life Sci* **2015,** *983-984*, 101-10.

81. Farnan, D.; Moreno, G. T., Multiproduct High-Resolution Monoclonal Antibody Charge Variant Separations by pH Gradient Ion-Exchange Chromatography. *Analytical Chemistry* **2009,** *81* (21), 8846-8857.

82. Hao, P.; Adav, S. S.; Gallart-Palau, X.; Sze, S. K., Recent advances in mass spectrometric analysis of protein deamidation. *Mass Spectrom Rev* **2017,** *36* (6), 677-692.

83. Li, X.; Yu, X.; Costello, C. E.; Lin, C.; O'Connor, P. B., Top-down study of β2-microglobulin deamidation. *Analytical chemistry* **2012,** *84* (14), 6150-6157.

84. Zabrouskov, V.; Han, X.; Welker, E.; Zhai, H.; Lin, C.; van Wijk, K. J.; Scheraga, H. A.; McLafferty, F. W., Stepwise deamidation of ribonuclease A at five sites determined by top down mass spectrometry. *Biochemistry* **2006,** *45* (3), 987-992.

85. Chelius, D.; Rehder, D. S.; Bondarenko, P. V., Identification and characterization of deamidation sites in the conserved regions of human immunoglobulin gamma antibodies. *Anal Chem* **2005,** *77* (18), 6004-11.

86. Hinterholzer, A.; Stanojlovic, V.; Regl, C.; Huber, C. G.; Cabrele, C.; Schubert, M., Identification and Quantification of Oxidation Products in Full-Length Biotherapeutic Antibodies by NMR Spectroscopy. *Analytical Chemistry* **2020,** *92* (14), 9666-9673.

87. Hawkins, C. L.; Davies, M. J., Detection, identification, and quantification of oxidative protein modifications. *The Journal of biological chemistry* **2019,** *294* (51), 19683-19708.

88. Liu, H.; Ponniah, G.; Neill, A.; Patel, R.; Andrien, B., Accurate Determination of Protein Methionine Oxidation by Stable Isotope Labeling and LC-MS Analysis. *Analytical Chemistry* **2013,** *85* (24), 11705-11709.

89. Hamuro, Y.; Coales, S. J.; Southern, M. R.; Nemeth-Cawley, J. F.; Stranz, D. D.; Griffin, P. R., Rapid analysis of protein structure and dynamics by hydrogen/deuterium exchange mass spectrometry. *J Biomol Tech* **2003,** *14* (3), 171-82.

90. Zhang, Z.; Smith, D. L., Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein Sci* **1993,** *2* (4), 522-531.

91. Houde, D.; Peng, Y.; Berkowitz, S. A.; Engen, J. R., Post-translational modifications differentially affect IgG1 conformation and receptor binding. *Mol Cell Proteomics* **2010,** *9* (8), 1716-28.

92. Engen, J. R., Analysis of Protein Conformation and Dynamics by Hydrogen/Deuterium Exchange MS. *Analytical Chemistry* **2009,** *81* (19), 7870-7875.

93. Benhaim, M.; Lee, K. K.; Guttman, M., Tracking Higher Order Protein Structure by Hydrogen-Deuterium Exchange Mass Spectrometry. *Protein Pept Lett* **2019,** *26* (1), 16-26.

94. Houde, D.; Arndt, J.; Domeier, W.; Berkowitz, S.; Engen, J. R., Characterization of IgG1 conformation and conformational dynamics by hydrogen/deuterium exchange mass spectrometry. *Analytical chemistry* **2009,** *81* (7), 2644-2651.

95. Bobst, C. E.; Abzalimov, R. R.; Houde, D.; Kloczewiak, M.; Mhatre, R.; Berkowitz, S. A.; Kaltashov, I. A., Detection and characterization of altered conformations of protein pharmaceuticals using complementary mass spectrometry-based approaches. *Anal Chem* **2008,** *80* (19), 7473-81.

96. Skinner, J. J.; Lim, W. K.; Bédard, S.; Black, B. E.; Englander, S. W., Protein dynamics viewed by hydrogen exchange. *Protein Sci* **2012,** *21* (7), 996-1005.

97. Deng, B.; Lento, C.; Wilson, D. J., Hydrogen deuterium exchange mass spectrometry in biopharmaceutical discovery and development - A review. *Anal Chim Acta* **2016,** *940*, 8-20.

98. Wei, H.; Mo, J.; Tao, L.; Russell, R. J.; Tymiak, A. A.; Chen, G.; Iacob, R. E.; Engen, J. R., Hydrogen/deuterium exchange mass spectrometry for probing higher order structure of protein therapeutics: methodology and applications. *Drug Discovery Today* **2014,** *19* (1), 95-102.

99. Huang, R. Y.; Chen, G., Higher order structure characterization of protein therapeutics by hydrogen/deuterium exchange mass spectrometry. *Anal Bioanal Chem* **2014,** *406* (26), 6541-58.

100. Hageman, T.; Wrigley, M.; Weis, D., Statistical Equivalence Testing of Higher-Order Protein Structures with Differential Hydrogen Exchange-Mass Spectrometry (HX-MS). *Analytical Chemistry* **2021,** *93*.

101. Morgan, C. R.; Engen, J. R., Investigating solution-phase protein structure and dynamics by hydrogen exchange mass spectrometry. *Curr Protoc Protein Sci* **2009,** *Chapter 17*, Unit 17.6.1-17.

102. Jensen, P. F.; Rand, K. D., Hydrogen Exchange: A Sensitive Analytical Window into Protein Conformation and Dynamics. In *Hydrogen Exchange Mass Spectrometry of Proteins*, 2016; pp 1-17.

103. Konermann, L.; Pan, J.; Liu, Y.-H., Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chemical Society Reviews* **2011,** *40* (3), 1224-1234.

104. Skinner, J. J.; Lim, W. K.; Bédard, S.; Black, B. E.; Englander, S. W., Protein hydrogen exchange: testing current models. *Protein Sci* **2012,** *21* (7), 987-95.

105. Konermann, L.; Tong, X.; Pan, L., Protein Structure and Dynamics Studied by Mass Spectrometry: H/D Exchange, Hydroxyl Radical Labeling, and Related Approaches. *Journal of mass spectrometry : JMS* **2008,** *43*, 1021-36.

106. Engen, J. R.; Wales, T. E.; Chen, S.; Marzluff, E. M.; Hassell, K. M.; Weis, D. D.; Smithgall, T. E., Partial cooperative unfolding in proteins as observed by hydrogen exchange mass spectrometry. *Int Rev Phys Chem* **2013,** *32* (1), 96-127.

107. Weis, D. D.; Wales, T. E.; Engen, J. R.; Hotchko, M.; Ten Eyck, L. F., Identification and Characterization of EX1 Kinetics in H/D Exchange Mass Spectrometry by Peak Width Analysis. *Journal of the American Society for Mass Spectrometry* **2006,** *17* (11), 1498-1509.

108. Ferraro, D. M.; Lazo, N.; Robertson, A. D., EX1 hydrogen exchange and protein folding. *Biochemistry* **2004,** *43* (3), 587-94.

109. Weis, D. D.; Kjellen, P.; Sefton, B. M.; Engen, J. R., Altered dynamics in Lck SH3 upon binding to the LBD1 domain of Herpesvirus saimiri Tip. *Protein Sci* **2006,** *15* (10), 2402-2410.

110. Wang, L. C.; Krishnamurthy, S.; Anand, G. S., Hydrogen Exchange Mass Spectrometry Experimental Design. In *Hydrogen Exchange Mass Spectrometry of Proteins*, 2016; pp 19-35.

111. Glasoe, P. K.; Long, F. A., USE OF GLASS ELECTRODES TO MEASURE ACIDITIES IN DEUTERIUM OXIDE1,2. *The Journal of Physical Chemistry* **1960,** *64* (1), 188-190.

112. Deng, Y.; Zhang, Z.; Smith, D. L., Comparison of continuous and pulsed labeling amide hydrogen exchange/mass spectrometry for studies of protein dynamics. *Journal of the American Society for Mass Spectrometry* **1999,** *10* (8), 675-684.

113. Hamuro, Y.; Coales, S. J.; Molnar, K. S.; Tuske, S. J.; Morrow, J. A., Specificity of immobilized porcine pepsin in H/D exchange compatible conditions. *Rapid Commun Mass Spectrom* **2008,** *22* (7), 1041-6.

114. Wang, L.; Pan, H.; Smith, D. L., Hydrogen exchange-mass spectrometry: optimization of digestion conditions. *Mol Cell Proteomics* **2002,** *1* (2), 132-8.

115. Sarpe, V.; Schriemer, D. C., Data Processing in Bottom-Up Hydrogen Exchange Mass Spectrometry. In *Hydrogen Exchange Mass Spectrometry of Proteins: Fundamentals, Methods, and applications*, First ed.; Weis, D. D., Ed. John Wiley & Sons, Ltd.: 2016; pp 37-53.

2 CHAPTER 2: Rapid prediction of deamidation rates of proteins to assess their long-term stability using hydrogen exchange-mass spectrometry (HX-MS)

2.1 Introduction

Proteins are increasingly used as therapeutics to treat various diseases because of their safety and efficacy.[1-7] However, proteins are susceptible to a variety of modifications during expression, purification, formulation, and long-term storage, potentially leading to loss of efficacy or a decrease in safety.[8,9] Among various protein modifications, deamidation is one of the most common degradation reactions, occurring during production and long-term storage. Deamidation (see **Figure 2.1**) is a spontaneous, post-translational modification in which asparagine converts to aspartate or isoaspartate and glutamine coverts to glutamate.[10-13] This conversion introduces a negatively-charged side chain and increases the mass of the protein by 0.984 Da. Furthermore, formation of isoaspartate alters the protein backbone by inserting a methylene group.[14] These modifications cause charge, size, and structural heterogeneity in proteins that can alter protein conformation thereby potentially affecting function, stability, efficacy, purity, and immunogenicity.[15-23]

Deamidation is a non-enzymatic, spontaneous chemical reaction that occurs via acid- and base- catalyzed pathways. In the base-catalyzed pathway, the backbone amide nitrogen on the C-terminal side of asparagine is deprotonated by OH^- in the solvent, as shown in **Figure 2.1**. The nucleophilic attack of the resulting amidate anion on the carbonyl carbon of the asparagine side-chain forms the succinimide cyclic intermediate.[24-27] The formation of the succinimide intermediate is the rate-determining step of the deamidation reaction.[24,26] The hydrolysis of the succinimide intermediate forms a mixture of isoaspartate and aspartate products in a ratio of approximately 3:1, respectively.[24,28] In addition, the resulting isoaspartate and aspartate products can be present in both L and D isoforms.[12,29,30]. Both asparagine and glutamine reactions follow

the same mechanism, but the rate of glutamine deamidation is much slower the asparagine deamidation because of the instability of the six-membered intermediate formed in glutamine deamidation.[31] The deamidation rate is dependent on solution pH, temperature, solvent ionic strength, protein sequence, and protein higher-order structure.[32-35] Protein sequence has a stronger effect on deamidation rate than the other rate-determining factors, mostly the amino acid residue C-terminal to the asparagine residue. It has been thought that sidechains with little steric hindrance enable conformational flexibility required for the intermediate formation and residues with low dipole moment stabilize the intermediate formation, thereby accelerating deamidation. However, it is now recognized that the acidity of the amide hydrogen can be strongly influenced by backbone conformation.[36] Overall, deamidation is fastest at asparagine-glycine sites followed by asparagine-histidine and asparagine-serine sites.[14,32] In unstructured peptides at neutral pH, deamidation half-life can vary from one day to one year depending on the C-terminal side neighboring residue. Deamidation half-lives of asparagine followed by glycine, histidine, serine, alanine, and isoleucine in model pentapeptides have been reported as 1 day, 10 days, 16 days, 25 days, and 318 days, respectively, at pH 7.4, 37.0°C in 0.15 M Tris-HCl. For glutamine, half-lives of unstructured peptides can extend to more than ten years.[14,37]

It is essential to identify deamidation-prone sites early in the drug development process in order to eliminate candidates with deamidation liabilities, to develop stable formulations, and to re-engineer candidates with high deamidation propensity.[38] However, accurate prediction of deamidation rates in folded proteins is challenging. Compared to unstructured peptides, deamidation rates in folded proteins are much slower, ranging from days to years, because higher-order structure of proteins plays a significant role in determining deamidation rates. Higher-order

structure stabilizes asparagine residues due to conformational rigidity and decreased nucleophilic reactivity of the H-bonded backbone amides.[39-44]

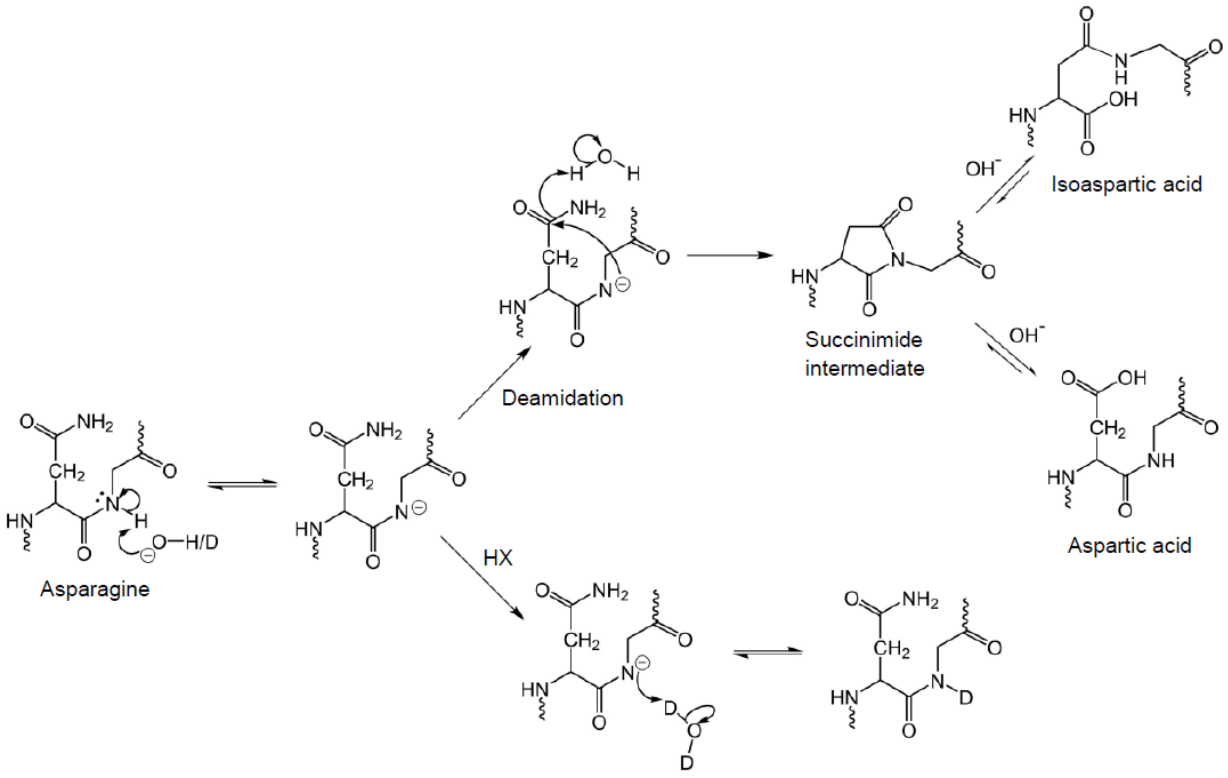


Figure 2.1: Base-catalyzed deamidation and HX mechanisms at an NG motif.

In order to determine the extent of deamidation under protein storage conditions, deamidation needs to be measured beyond the projected shelf-life, often as much as two years. Because such long-term stability studies are time-consuming, a common approach to identify potential deamidation sites in protein therapeutics is an accelerated stability study under stress conditions using elevated pH and temperature.[45,46] The stress conditions can increase the frequency of local and global transient unfolding, which in turn can accelerate deamidation. However, the deamidation propensities under stress conditions are different from the deamidation propensities

under actual protein storage conditions.[46-48] It usually takes 3-4 weeks to complete accelerated stability-based predictions, but it is very important to limit the time required for these studies to minimize drug development time.[46] Therefore, there is a large unmet need for analytical tools that can rapidly and reliably predict deamidation rates in folded proteins under long-term storage and realistic formulation conditions. Robinson et al. developed a computational method for the estimation of deamidation rates in proteins based on the sequence-controlled asparagine deamidation rates in pentapeptide models and the three-dimensional structure of proteins with well characterized deamidations.[42,49] Sydow et al. developed a computational method to predict degradation propensity of asparagine residues in the variable region of mAbs.[50] However, these methods require either that the three-dimensional structure of protein is known or that the structure can be accurately modelled.

Accurate prediction of deamidation rates in folded proteins requires knowledge about the protein conformation and the dynamics of amide H-bonding because the rate-determining step in the deamidation reaction mechanism involves deprotonation of a backbone amide leading to succinimide intermediate formation (see **Figure 2.1**). These amide hydrogens contribute to the formation of the fundamental H-bonding network of protein higher-order structure.[51] Hydrogen exchange monitored by mass spectrometry (HX-MS) is a technique that measures isotopic exchange of the protein backbone amide hydrogens. In the base-catalyzed mechanism of hydrogen exchange, an OD^- ion in the solvent abstracts the amide proton through nucleophilic attack producing an amidate ion, which is the rate-determining step of the reaction. Deprotonation is followed by the deuteration of the amidate ion (see **Figure 2.1**) by the D_2O solvent.[52] HX reveals the conformation and the dynamics of the backbone amide hydrogens in folded proteins; the deamidation rate in folded proteins also depends on the conformation and dynamics of the amide

hydrogen. Furthermore, when the reaction mechanisms of HX and deamidation are compared, the rates of both the reactions are determined by the rate of amidate ion formation, leading to the hypothesis that there is a correlation between the deamidation rates and HX rates in folded proteins. In fact, there is experimental support for this hypothesis: Philllips et al. recently found that the HX rates at the two adjacent amides in a mAb correlated with their deamidation rates.[53] In this study, we investigated the correlation between deamidation during long-term storage and HX to assess the potential for bottom-up HX-MS to be used as an analytical tool to rapidly predict localized long-term deamidation propensities.

2.2 Experimental

Wild-type and mutant (W169G) maltose binding protein (MBP) were expressed in *E.* coli and purified using nickel affinity chromatography as described in the Appendix A.

2.2.1 Long-term and accelerated deamidation studies

Samples of 20 µM wild-type and mutant MBP were prepared in 20 mM sodium phosphate, 100 mM sodium chloride, 100 µM EDTA at pH 7.0. The protein samples were sterile filtered through a 0.22 µM PVDF membrane (Millipore) under sterile conditions and stored in autoclaved microcentrifuge tubes at room temperature (23±2°C) in the dark. Wild-type and mutant MBP samples were collected after 1, 3, 6, 9, and 12 and 1, 3, 5, 7, 9, and 12 calendar months, respectively. For method development, deamidation was also accelerated in wild-type MBP by applying stress conditions to identify all the sites susceptible to deamidation. Wild-type MBP was buffer-exchanged by dialysis at 4°C into 50 mM tris at pH 8.5. The buffer-exchanged protein was diluted in 50 mM tris and 100 µM EDTA at pH 8.5 to prepare 50 µM protein samples that were

incubated at 40°C for 4 weeks. The collected samples were flash frozen on liquid N_2 and stored at −80°C until analyzed.

2.2.2 Quantification of deamidation

The extent of deamidation at asparagine sites was determined by LC-MS analysis of the tryptic peptides of MBP, as described in more detail in the Appendix A. The relative deamidation percentage at each site was quantified using the integrated peak areas of extracted ion chromatograms (EIC) of unmodified and deamidated peptides. It has been previously shown that the peak areas of the EICs can be used for relative quantification as ionization efficiencies of deamidated and unmodified peptides are similar.[54] Deamidation percentage at each asparagine site was calculated as follows:

$$\%\ \text{deamidation} = \frac{A_{\mathrm{d}}}{A_{\mathrm{u}} + A_{\mathrm{d}}} \times 100\%$$

where *A* denotes the integrated EIC peak area and the subscripts u and d denote the unmodified and deamidated peptides, respectively. The tryptic peptides selected for the quantification each contained only one deamidation site. EICs of the deamidated peptides for some asparagine sites consisted of two peaks resulting from the chromatographic resolution of aspartate and isoaspartate products of deamidation. For those sites, the area of both peaks was combined as deamidated products when obtaining the deamidated peptide peak area. In some cases, co-elution of the unmodified and deamidated peptides was observed. In these cases, the overlap between the peaks was determined and corrected as illustrated in the Appendix A **Figure 2.7.A-2** for the peptide 100-LYPFTWDAVRYNGK-113. The co-elution of the unmodified and the deamidated peptides resulted in the overlap of the M+1 peak (m/z = 577.628) of the unmodified peptide with the

monoisotopic peak (M) of the deamidated peptide (m/z = 577.625, see Appendix A **Figure 2.7.A-2B and 2.7.A-2D**). The area corresponding to the overlapping M+1 peak of the unmodified peptide was calculated by multiplying the unmodified peptide EIC peak area by the theoretical isotopic abundance of the unmodified peptide M+1 peak (99.25%) obtained using MassHunter Isotope Distribution Calculator software (Agilent, Santa Clara, CA) as shown in Appendix A **Figure 2.7.A-2E**. To determine the peak area of the deamidated peptide alone, the area of the overlapping M+1 peak of the unmodified peptide was subtracted from the total peak area of the EIC of the deamidated peptide.

2.2.3 Deamidation kinetics

Deamidation half-life ($t_{\frac{1}{2},\mathrm{d}}$) at 23±2°C and pH 7.0 at each deamidation site was determined assuming first-order kinetics.[14] The deamidation rate constant at each deamidation site was obtained by linear regression of the natural log of fraction unmodified vs time for each deamidation site. Deamidation half-lives were determined from the first-order rate constant values. The error in deamidation half-life was determined by propagation of random error using the standard deviation of the slope obtained using the LINEST function in Excel.

2.2.4 Hydrogen exchange–mass spectrometry

HX-MS measurements used robotic sample preparation followed by online pepsin digestion and LC-MS analysis of the deuterated peptic peptides, as described in more detail in the Appendix A. Samples were labeled with D_2O buffer (20 mM sodium phosphate, 100 mM sodium chloride, pD 7.0 corrected for the isotope effect[55]) at 25°C. Deuteration was normalized to 100% by labeling MBP peptic peptides obtained from offline digestion. The rate of HX was determined

in terms of peptide averaged $t_{\frac{1}{2},\mathrm{HX}}$; the time required to reach 50% deuteration. $t_{\frac{1}{2},\mathrm{HX}}$ values were determined by linear interpolation between the two HX times that spanned 50% deuteration. The error associated with the $t_{\frac{1}{2},\mathrm{HX}}$ values was determined by propagation of random errors in the slope and the intercept of the linear fit. The errors in the slope and the intercept were estimated using the standard deviations of the slope and the intercept obtained using the LINEST function in Excel.

2.3 Results

MBP was selected as a model protein for this study because it contains 21 asparagine residues; five of them are followed by glycine. The NG motif is the most susceptible to deamidation.[14] MBP is a 41 kDa monomeric protein with two distinct globular domains separated by a groove.[56] Two forms of MBP; wild-type and a mutant form (W169G) were used in the study. The mutant form is less stable than wild-type MBP, based on a lower thermal transition temperature ($\Delta T = -4.8$°C) measured by differential scanning calorimetry (see Appendix A **Table 2.3.A-1**). To investigate the potential of using HX-MS as a tool to rapidly predict deamidation rates in proteins, deamidation rates at NG and NA sites in MBP were correlated with HX rates using the shortest peptic peptides containing those asparagine sites.

2.3.1 Deamidation kinetics

A long-term stability study was performed to determine deamidation rates at the asparagine sites of MBP under realistic storage conditions without temperature or pH stress. For the stability study, wild-type and mutant MBP were stored in the dark at 23±2°C and pH 7.0 over 12 months under sterile conditions. (A preliminary study indicated that deamidation was too slow to measure at 4°C, results not shown.) The intact masses of the two proteins were measured over the course

of the study to determine if the proteins underwent any other modifications because other protein modifications can alter protein conformation affecting protein deamidation rate. The deconvoluted masses of the two proteins at zero and 12 months are shown in Appendix A **Figure 2.7.A-3**. The mass of mutant MBP increased by 1.9 Da over the 12 months which can be attributed to the mass increase of 0.98 Da due to deamidation. For wild-type MBP, a mass increment was not observed over the 12 months. Low abundance peaks that appeared at 41258 Da and 41129 Da for wild-type and mutant MBP, respectively, indicate a mass loss of 208 Da. Both MBP forms contain an N-terminal alanine preceding a hexa-histidine tag; the loss of 208 Da is consistent with the hydrolysis of the first two N-terminal residues. In addition, two other peaks at 39269 Da and 38841 Da were observed in the mutant MBP sample after 12 months indicating some fragmentation in the protein which could not be assigned. These two peaks were also detected after 7 months of storage, but at lower abundance. The deconvoluted mass spectra did not reveal other common modifications of proteins such as oxidation that increase protein mass. Although some degradation occurred over one year of storage, unmodified protein remained the major form. Thus, the observed deamidation kinetics represent deamidation by the unmodified forms of the proteins. There was no substantial decrease in the soluble protein concentration in any of the protein samples over the course of storage, as determined by UV absorbance (results not shown), indicating that MBP did not form insoluble aggregates to any significant extent.

The extent of deamidation was measured by LC-MS analysis of tryptic peptides containing only one deamidation site per peptide (see **Table 2.1**). Deamidation percentage at each deamidation site was quantified using peak areas of the EICs of deamidated and unmodified tryptic peptides as described in the Experimental section. Deamidation was detected only at NG and NA sites in both protein forms during the 12 month period; there was no deamidation detected at other

asparagine sites. The asparagine residues with NG sites in MBP are N23G24, N29G30, N111G112, N184G185, N238G239 and the asparagine residues with NA sites are N196A197, N216A217, N278A279, N334A335 and N360A361. Among these sites, deamidation was only detected at N23G24, N29G30, N111G112, N184G185, N196A197 and N360A361.

Table 2.1: Tryptic peptides used for the deamidation analysis and peptic peptides used for the HX analysis of the NG and NA sites in both wild-type and mutant MBP.

Asparagine sites	Tryptic peptides	Peptic peptides
N23G24	LVIWINGDK (18-26)	VIWINGDKGY (19-28)
N29G30	GYNGLAEVGK (27-36)	VIWINGDKGYNGL (19-31)
N111G112	LYPFTWDAVRYNGK (100-113)	VRYNGKL (108-114)
N184G185	YENGKYDIK (182-190)	FKYENGKY (180-187)
N238G239	GETAMTINGPWAWSNIDTSK (231-250)	TINGPWAWSN (236-245)
N196A197	DVGVDNAGAK (191-200)	VGVDNAGAKAGLTF (192-205)
N360A361	TAVINAASGR (356-365)	YAVRTAVINAASGRQTVDE (352-370)
N216A217	HMNADTDYSIAEAAFNK (214-230)	VDLIKNKHMNADTDY (207-221)
N278A279	PFVGVLSAGINAASPNK (268-284)	SAGINAASPNKEL (274-286)
N334A335	IAATMENAQK (328-337)	AKDPRIAATMENAQKGEIM (322-341)

Figure 2.2 shows the extent of deamidation at each of the deamidation sites in wild-type and mutant MBP over the 12 month period. Among the NG sites in wild-type MBP (**Figure 2.2A**), deamidation was fastest at N184G185 and N111G112, reaching 55% and 42%, respectively, after 12 months. There was much less deamidation at N23G24 and N29G30, which were only 6% and 4%, respectively. There was no detectable deamidation at N238G239. There was only modest deamidation at NA sites N196A197 and N360A361 (**Figure 2.2C**) in the wild-type, reaching 7%

and 1.5%, respectively. Overall, deamidation was faster in the mutant than the wild-type (**Figure 2.2B and 2.2D**). In the mutant, there was extensive deamidation at both N184G185, which reached 100% deamidation, and at N111G112, which reached 55%. Deamidation was much slower at N23G24 and N29G30, which were 7.5% and 5.6%, respectively. Deamidation was not detected at N238G239 in the mutant. Deamidation was moderate at NA sites N196A197 and N360A361, reaching 17% and 7%, respectively. The deamidation half-lives ($t_{\frac{1}{2},\mathrm{d}}$), based on first-order kinetics,[14] obtained for each deamidation site in wild-type and mutant MBP are shown in **Table 2.2**. Faster deamidation rates in the mutant demonstrates the increased susceptibility of the mutant to deamidation due to its lower stability relative to the wild-type.

According to the X-ray crystal structure of wild-type MBP[56], shown in **Figure 2.3**, N184G185 and N111G112, which deamidated the fastest ($t_{\frac{1}{2},\mathrm{d}}$ of 12±1 and 18±2 months, respectively), are located in solvent-exposed loop regions. N23G24 which deamidated at a moderate rate ($t_{\frac{1}{2},\mathrm{d}}$of 400±100 months), is located in a partially buried loop region. N29G30 which deamidated the slowest ($t_{\frac{1}{2},\mathrm{d}}$ of 1100±700 months), is located in an α-helical region. N238G239 which did not have detectable amount of deamidation is located in a β-sheet in the core of the protein. The results of the deamidation kinetics of the NG sites demonstrate that asparagine residues in highly structured regions undergo deamidation slowly while asparagine residues in less structured regions undergo deamidation faster. Similar observations have been reported previously. Chelius et al. reported, based on a study on an IgG1 mAb, that the deamidation site located in a buried β-sheet structure had a slower deamidation rate compared to the deamidation sites located on the surface of the antibody because of reduced flexibility and inaccessibility to react with water.[57] Fukuda et al. reported that between two NG sites located on the surface of β_2-

microglobulin, the NG site with asparagine whose side chain was involved in H-bonding deamidated much slower than another asparagine residue which was not H-bonded. This non hydrogen-bonded asparagine was also located in an environment enhancing the electrophilicity of the asparagine side chain and the nucleophilicity of the amide nitrogen.[58]

Considering the NA sites, the NA sites which are located in highly flexible regions from HX data (**Table 2.2**) underwent deamidation while the NA sites in less flexible regions, did not measurably deamidate during the time of the study. As shown in **Figure 2.3**, N196A197 and N360A361, the NA sites which underwent deamidation, are located in α-helix termini with surface-exposed asparagine side chains. N216A217 and N278A279 sites, which did not undergo deamidation, are located in loop regions. N216A217 has a partially surface exposed asparagine side chain while N278A279 site is not on the surface, thus the apparent absence of deamidation at N216A217 is not consistent with a simple structural interpretation. N334A335, another site that did not undergo deamidation, is located in the middle of an α-helix in the core of the protein. This is consistent with previous work where between two unstable asparagine residues in the enzyme triosephosphate isomerase (TPI), the more solvent-accessible asparagine deamidated faster.[59]

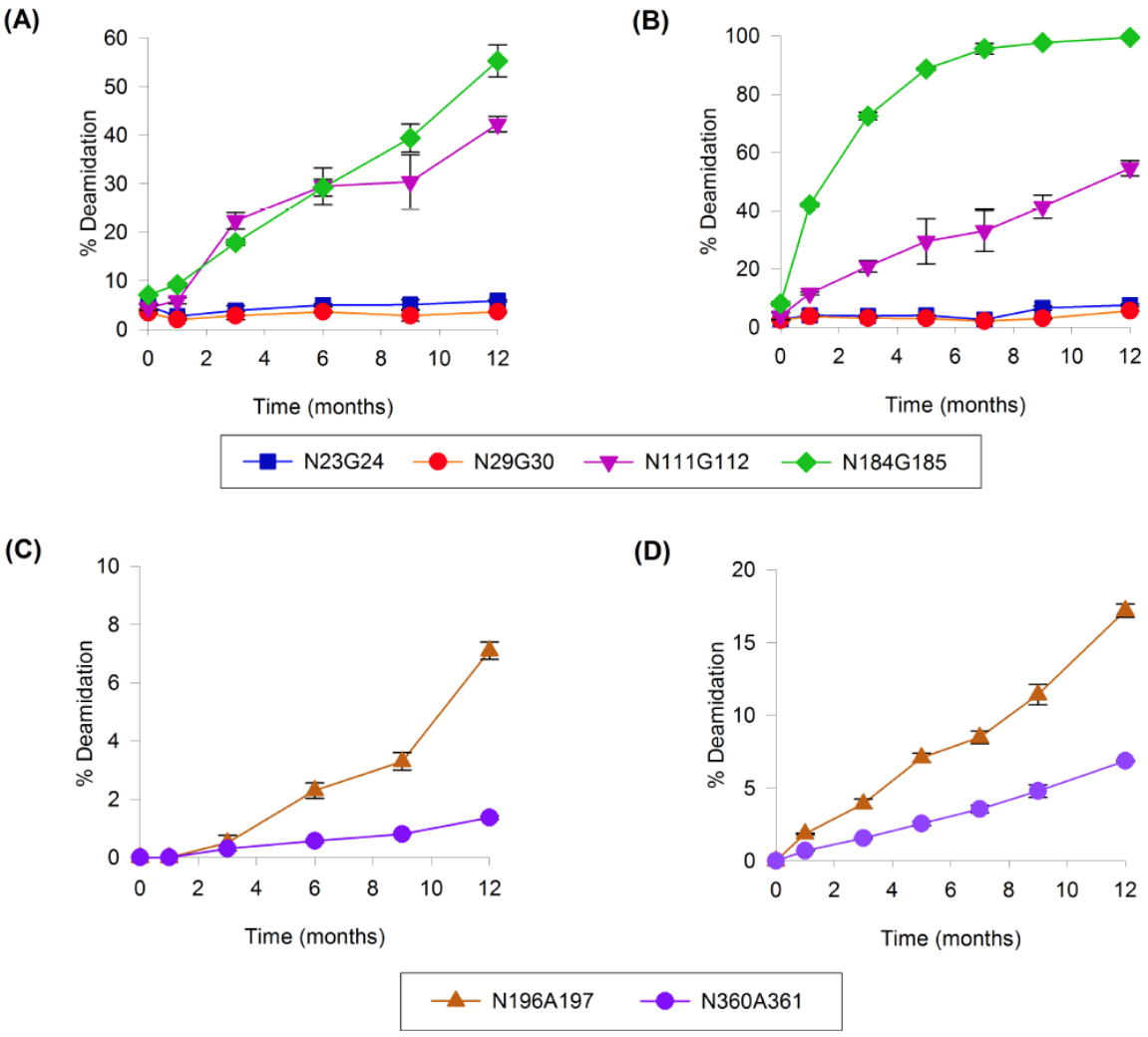


Figure 2.2: Extent of deamidation at each deamidation site in maltose binding protein (MBP) in 20 mM sodium phosphate, 100 mM sodium chloride, 100 µM EDTA, at pH 7.0 at 23±2°C in the dark. (A) NG sites in wild-type MBP, (B) NG sites in mutant MBP, (C) NA sites in wild-type MBP, and (D) NA sites in mutant MBP.

Table 2.2: Deamidation half-lives[1] at 23±2°C and HX half-lives at pD 7.0 corrected for the isotope effect and 25°C obtained for the NG and NA sites in wild-type and mutant MBP.

Asparagine sites	**Wild-type MBP**		**Mutant MBP**	
	Deamidation half-life (months) $t_{\frac{1}{2},\mathrm{d}}$	HX half-life (s) $t_{\frac{1}{2},\mathrm{HX}}$	Deamidation half-life (months) $t_{\frac{1}{2},\mathrm{d}}$	HX half-life (s) $t_{\frac{1}{2},\mathrm{HX}}$
N23G24	400 ± 100	980 ± 60	200 ± 40	700 ± 60
N29G30	1100 ± 700	1300 ± 100	500 ± 200	600 ± 20
N111G112	18 ± 2	150 ± 10	12 ± 5	130 ± 5
N184G185	12 ± 1	< 30	2.0 ± 0.5	< 30
N238G239	Not detected	10900 ± 400	Not detected	7800 ± 100
N196A197	120 ± 10	< 30	50 ± 2	< 30
N360A361	630 ± 30	< 30	120 ± 4	< 30
N216A217	Not detected	270 ± 30	Not detected	< 30
N278A279	Not detected	3800 ± 400	Not detected	2600 ± 100
N334A335	Not detected	35000 ± 900	Not detected	16700 ± 20

[1]Deamidation half-lives were calculated from experimentally determined first-order rate constants using the data shown in Figure 1.

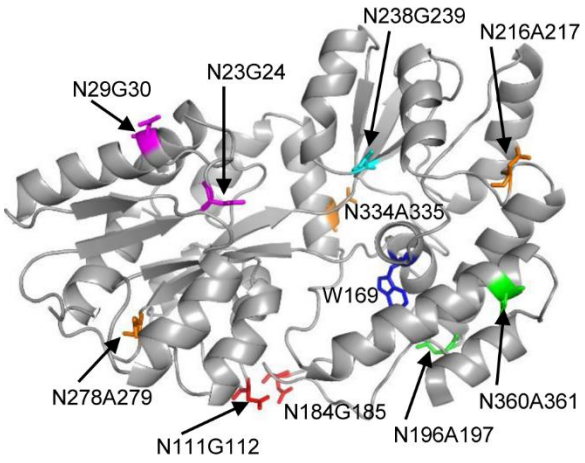


Figure 2.3: Locations of the NG and NA sites and the point mutation W169G on the X-ray crystal structure of wild-type MBP (PDB entry 1OMP).

2.3.2 HX kinetics of deamidation Sites

With peptide-level resolution, the HX rate at a specific amide cannot be determined directly. Therefore, we have used empirical $t_{\frac{1}{2},\mathrm{HX}}$ values, the time required to reach 50% deuteration, of the shortest peptic peptides containing each deamidation site to estimate the HX rate at the deamidation sites of interest (see **Table 2.1**). The HX experiments were performed only on protein samples prior to storage with the goal of predicting the outcome of the long-term stability study. The wild-type and mutant MBP samples were labeled with deuterium for times ranging from 30 s to 12 hr at 25°C at pD 7.0. **Figure 2.4** shows deuterium uptake curves for the shortest peptic peptides containing the deamidation sites. **Table 2.2** lists the $t_{\frac{1}{2},\mathrm{HX}}$ values calculated for those peptides, obtained by linear interpolation. $t_{\frac{1}{2},\mathrm{HX}}$ values for the peptides which reached 50% deuteration before the shortest exchange time, 30 s, cannot be estimated accurately, therefore

$t_{\frac{1}{2},\mathrm{HX}}$ values for those peptides are reported as less than 30 s. Rapid HX in these regions indicates that the protein backbone is highly dynamic.

In wild-type MBP, the peptic peptide containing the N238G239 site, which had no detectable deamidation, had the slowest HX, $t_{\frac{1}{2},HX} = 10900 \pm 400$ s. At the other extreme, the peptic peptide with N184G185, the NG site with the fastest deamidation rate, also had the fastest HX rate, resulting in $t_{\frac{1}{2},HX} < 30$ s. Sites with intermediate rates of deamidation also had intermediate HX rates. In the mutant, HX was generally faster at all sites, consistent with its decreased stability. Overall, the trend in HX rates is similar to the wild-type.

The two NA sites, N196A197 and N360A361, that deamidated at moderate rates had faster HX rates, $t_{\frac{1}{2},HX}$ values less than 30 s, in both wild-type and mutant MBP. In comparison to the deamidation rate at the NG site which had similar $t_{\frac{1}{2},HX}$ of less than 30 s, the deamidation rates of these NA sites are much slower. According to the deamidation rates of peptides reported by Robinson et al.,[14] deamidation rates at NA sites are 20-fold slower than NG sites. For the NA sites in MBP, deamidation was only detected in the regions where HX was too fast to measure and no deamidation was detected in the regions where HX rates were measurable. In other words, measurable deamidation was only observed at the most dynamic NA sites. This observation is consistent with the trends observed at NG sites among the deamidation rates and HX rates.

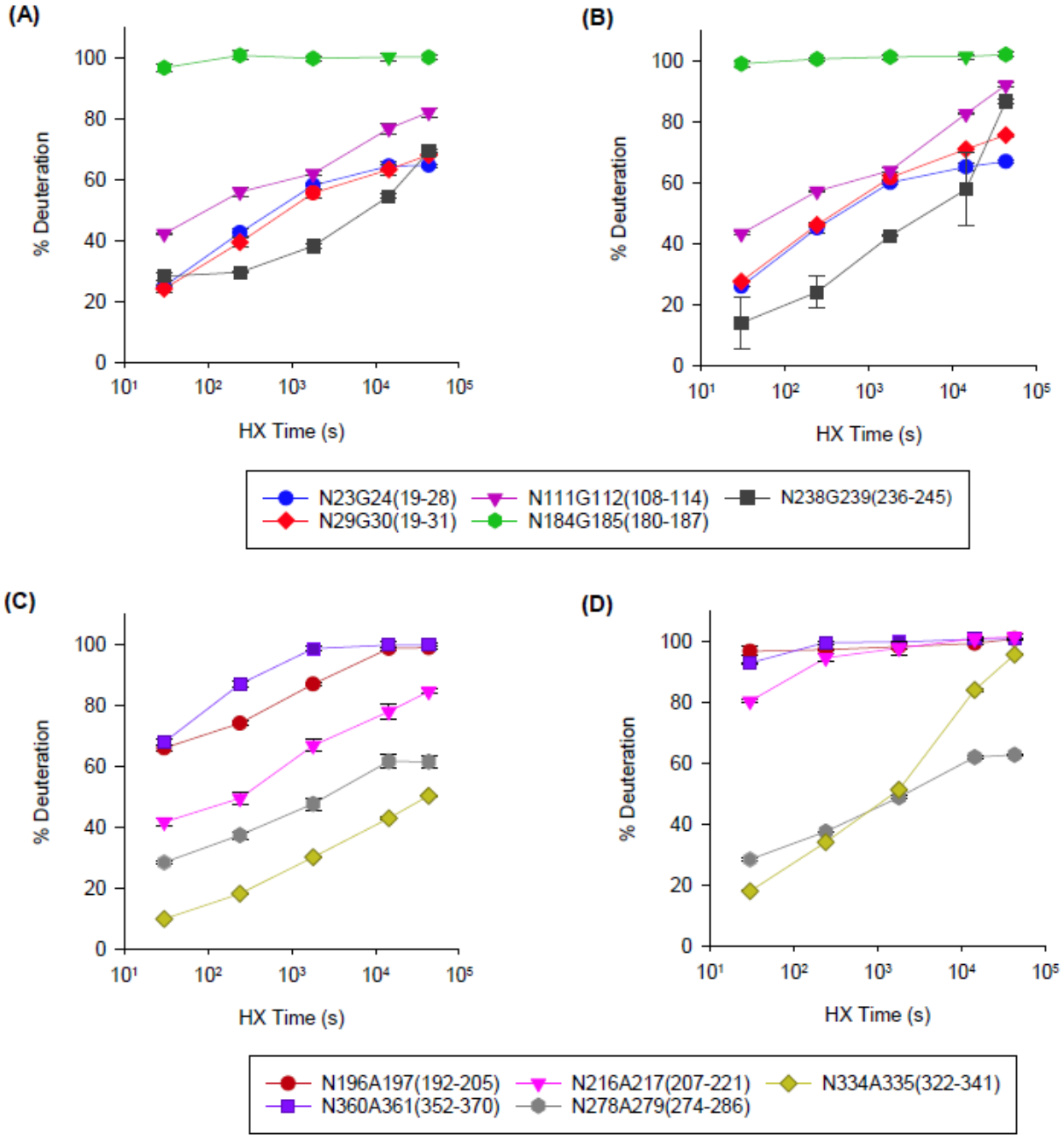


Figure 2.4: Hydrogen exchange kinetics at 25°C at pD 7.0 corrected for the isotope effect for the shortest peptic peptides containing each deamidation-prone site in MBP. (A) NG sites in wild-type MBP, (B) NG sites in mutant MBP, (C) NA sites in wild-type MBP, and (D) NA sites in mutant MBP.

2.3.3 Correlation between the deamidation rates and the HX rates

Overall, when the deamidation kinetics ($t_{\frac{1}{2},\mathrm{d}}$) and HX kinetics ($t_{\frac{1}{2},\mathrm{HX}}$) at the NG sites are compared in both protein forms, NG sites with faster deamidation rates (N184G185 and N111G112) had faster HX and NG sites with moderate and slower deamidation rates (N23G24 and N29G30) had moderate HX rates. N238G239, with no detectable deamidation, had the slowest HX. This demonstrates that the NG sites with faster deamidation rates are located in more flexible regions in the protein (indicated by faster HX rates) and the asparagine sites with slower deamidation rates are located in less flexible or rigid regions (indicated by slower HX rates). This observation, in turn, suggests that $t_{\frac{1}{2},\mathrm{HX}}$ values can be used to differentiate between the deamidation rates at different deamidation sites within a class of deamidation sites in a protein (e.g, NG motifs or NA motifs). Moreover, as observed between the two proteins states, $t_{\frac{1}{2},\mathrm{HX}}$ values can also differentiate between the deamidation rates at a particular asparagine residue in two protein states with different stabilities.

To examine the correlation between deamidation and HX at the NG sites in MBP, the half-lives were correlated on a log-log plot (see **Figure 2.5**). N184G185 and N238G239 sites were not included in the correlation analysis because HX rates of N184G185 was too fast and no deamidation was detected at N238G239. The measurable half-lives in these cases are indicated as the horizontal and vertical dashed lines in **Figure 2.5**. The results from the NA sites were not included in the correlation because either only deamidation kinetics or HX kinetics were measurable for the NA sites and also because NA sites have inherently slower deamidation rates than NG sites.[14] The square of the correlation coefficient was 0.94 demonstrating a power law

relationship between deamidation rates and HX rates in MBP. The correlation can be represented as

$$t_{\frac{1}{2},\mathrm{d}} = \frac{\left(t_{\frac{1}{2},\mathrm{HX}}\right)^{1.8\pm0.2}}{10^{2.7\pm0.7}}.$$

This correlation indicates that HX rate, based on $t_{\frac{1}{2},\mathrm{HX}}$ values, is predictive of the deamidation rates at NG sites in MBP that have measurable HX kinetics. Thus, an initial HX analysis of MBP data yields reliable ranking of deamidation kinetics in a one year stability study.

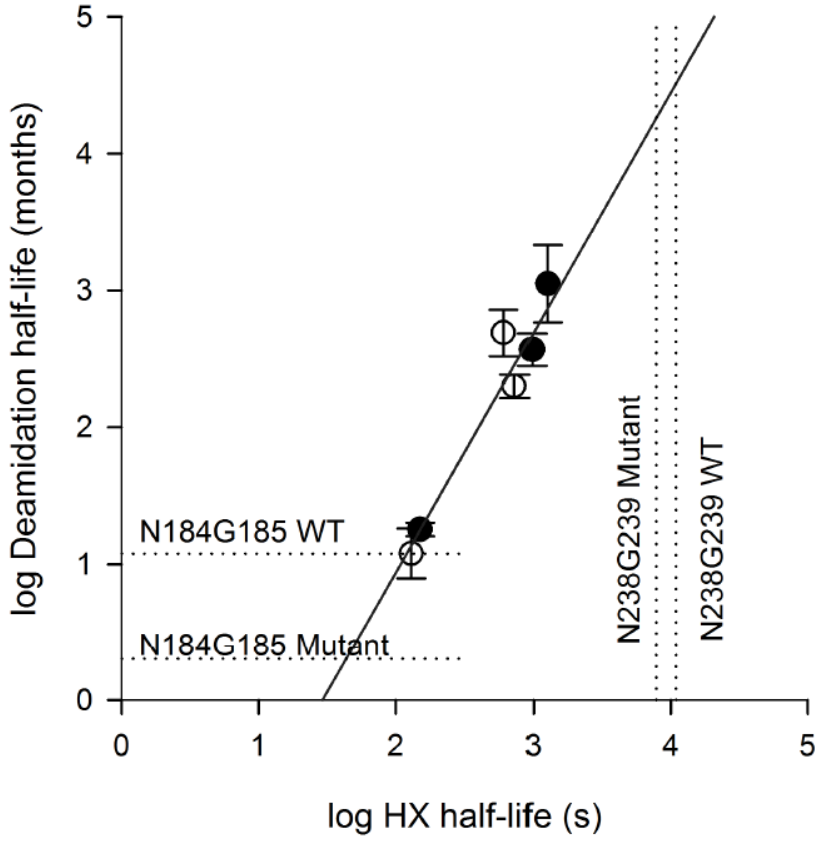


Figure 2.5: Correlation between HX half-life ($\boldsymbol{t}_{\frac{1}{2},\mathbf{HX}}$ in s) at 25°C at pD 7.0 corrected for the isotope effect and deamidation half-life ($\boldsymbol{t}_{\frac{1}{2},\mathbf{d}}$ in months) at 23±2°C at pH 7.0 obtained for deamidation-prone NG sites of wild-type (closed circles) and mutant (open circles) MBP, shown on a log-log plot. The R^2 value for the correlation is 0.94 and the linear fit is $y = 1.85x - 2.78$ where y denotes the base 10 log of deamidation half-life in months and χ denotes the base 10 log of HX half-life in seconds. The vertical dotted lines represent the HX $\boldsymbol{t}_{\frac{1}{2},\mathbf{HX}}$ values at the N238G239 sites where deamidation was too slow to measure over 12 months. The horizontal dotted lines represent the $\boldsymbol{t}_{\frac{1}{2},\mathbf{d}}$ of the N184G185 sites where HX was too fast for determination of $\boldsymbol{t}_{\frac{1}{2},\mathbf{HX}}$.

2.4 Discussion

Deamidation in therapeutic proteins can potentially lead to loss of efficacy or compromise safety.[13] Therefore, it is necessary to assess deamidation propensity and measure its extent in therapeutic protein candidates in order to select a stable protein drug candidate and to identify stabilizing conditions. Accelerated stability studies under stress conditions cannot accurately predict deamidation rates under actual drug storage conditions, thus they tend to represent a worst-case scenario. In addition, even accelerated stability studies take several weeks to complete. This makes it important and useful to identify an analytical technique that could rapidly and accurately predict deamidation rates in proteins under realistic storage conditions.

Even though the asparagine sites that undergo rapid deamidation in unstructured peptides can be predicated based on amino acid sequence,[14] predicting the rates at which those sites undergo deamidation in a folded protein is not possible without understanding how higher-order structure and dynamics influences the rate. For example, in a study carried out by Gasa-Bulseco et al., NG sites located in β-sheet structure had less susceptibility to deamidation than a NN site located in a loop region, despite the fact that the NN motif has less propensity for deamidation than NG. This observation highlights the greater influence of three-dimensional structure rather than sequence in determining deamidation rates in folded proteins.[60] While there are obvious connections between structure and deamidation in MBP, it is often the case that the precise structure of a protein of interest is not known or cannot be accurately modelled such as in the variable domains of antibodies. Therefore, gaining a better understanding of how protein conformation and backbone dynamics affect deamidation will enable more accurate predictions of protein deamidation rates based on empirical measurements. HX data reveal experimental information about how protein

conformation and backbone flexibility influence the amide hydrogen bonding; the mechanism of deamidation shares the same characteristics of hydrogen exchange, leading to the hypothesis that HX could potentially be used to predict deamidation rates in proteins based on experimental observations.

Our results demonstrate a power law correlation between deamidation and HX kinetics at the NG sites in the two forms of MBP. This correlation suggests that HX could be used for ranking of the deamidation sites in order of their relative deamidation rates, based on peptide-averaged HX rates. Consistent with our results here, Phillips et al. reported a similar observation that HX rate at the two amides adjacent to asparagine (residues *i*+1 and *i*+2 where *i* represents the deamidation site) correlated with the rate of deamidation.[53] In our work, the large deviation of the exponent of power law correlation between HX and deamidation clearly indicates that there is not a linear correspondence between HX and deamidation rates. Although the mechanisms of deamidation and HX share some similar features (see **Figure 2.1**), both reactions have complex mechanisms that defy simplified kinetic analysis. It is also likely that there is cooperativity between deamidation at different sites and other degradation pathways. In the case of NA sites, only slow deamidation was observed at the NA sites that are located in the rapidly exchanging regions. Although these sites, based on HX kinetics, are located in highly flexible regions, their deamidation occurred only at moderate rates. This can be explained by the deamidation rate dependence on protein sequence: deamidation at NA sites is approximately 20-fold slower than deamidation at NG sites.[14] At three other NA sites with slow HX, deamidation was undetectable even after 12 months.

Using HX for the prediction of deamidation rates in proteins would offer a rapid method compared to both accelerated stability studies, which take several weeks, and long-term stability

studies, which can take up to two years. HX-based experiments and data analysis for deamidation predictions can easily be completed within 2-3 days. Even though HX-MS is often considered time-consuming because of the data analysis,[61] deamidation predictions require analysis of data only for the peptides containing suspected deamidation sites. Once the peptides with deamidation sites are assigned by peptide mapping, an optimized HX-MS experiment can be completed in one day and HX-MS data analysis of the selected peptides can be completed in only a few hours using HX-MS data analysis software. Furthermore, HX predictions of deamidation rates can be made under realistic storage conditions rather than basing predictions on results obtained under stress conditions. In addition, as observed for wild-type and mutant MBP in this study, peptide-averaged HX rates can be used to compare relative deamidation propensities at deamidation-prone sites among similar forms of proteins such as mutants or engineered protein forms and also the same protein in different formulations in drug discovery and development. While HX-MS cannot be a replacement for a long-term stability study, it clearly has the potential to serve as a reliable predictor of deamidation propensity that is commensurate with pharmaceutical timelines required in late discovery and early development.

The results obtained in this study are based only on two forms of the model protein MBP, a mostly of α-helical protein. The specific correlation observed for MBP between its HX rates and deamidation rates cannot be directly applied to other proteins. Therefore, it is necessary to further explore the correlation between HX rates and deamidation rates in proteins containing other secondary structures and folds, especially the immunoglobulin fold in antibodies, and also to determine whether the correlation can be generalized or is protein-specific. Usually, the time window for HX measurements ranges from tens of seconds to approximately one day due to experimental limitations. This time window will not capture very fast exchange requiring less than

seconds and very slow exchange that can span months. This is a limitation to the application of HX that could potentially be extended using msec HX[62] and altered pH and temperature.[63-65] An additional limitation is that conventional HX-MS data is limited to peptide-level resolution where the exchange rate is averaged across the peptide and thus is not capable of achieving residue-level resolution.

The results from this study could be extended to examine the correlation between deamidation and HX rates in a wider variety of proteins including mAbs. Furthermore, correlation between HX and other protein modifications that involve backbone dynamic changes could also be investigated to identify the potential of using HX-MS as a prediction tool for such modifications in therapeutic protein development. To improve the spatial resolution of HX data obtained in the conventional HX method, electron transfer dissociation (ETD) can be coupled with HX which provides residue-level resolution.[66,67] With residue-level resolution, HX kinetics of the deamidation prone residue or its neighboring residue can be obtained which is more accurate than using peptide-averaged HX kinetics to predict deamidation propensity of a single residue or its adjacent residue as demonstrated recently with peptide-based HX-MS coupled with ETD.[53] HX coupled to ETD could be used to extend our study to investigate how the HX rate at an asparagine residue correlates with its deamidation rate and thereby to use HX as a potential tool to predict deamidation rates in proteins.

2.5 References

1. Walsh G 2014. Biopharmaceutical benchmarks 2014. Nat Biotechnol 32(10):992-1000.

2. Lagassé HAD, Alexaki A, Simhadri VL, Katagiri NH, Jankowski W, Sauna ZE, Kimchi-Sarfaty C 2017. Recent advances in (therapeutic protein) drug development. F1000Research 6:113.

3. Carter PJ 2006. Potent antibody therapeutics by design. Nat Rev Immunol 6(5):343-357.

4. Imai K, Takaoka A 2006. Comparing antibody and small-molecule therapies for cancer. Nat Rev Cancer 6(9):714-727.

5. Elvin JG, Couston RG, van der Walle CF 2013. Therapeutic antibodies: Market considerations, disease targets and bioprocessing. Int J Pharm 440(1):83-98.

6. Pavlou AK, Belsey MJ 2005. The therapeutic antibodies market to 2008. Eur J Pharm Biopharm 59(3):389-396.

7. Aggarwal RS 2014. What's fueling the biotech engine-2012 to 2013. Nat Biotechnol 32(1):32-39.

8. Liu H, Gaza-Bulseco G, Faldu D, Chumsae C, Sun J 2008. Heterogeneity of monoclonal antibodies. J Pharm Sci 97(7):2426-2447.

9. Walsh G, Jefferis R 2006. Post-translational modifications in the context of therapeutic proteins. Nat Biotechnol 24(10):1241-1252.

10. Gervais D 2016. Protein deamidation in biopharmaceutical manufacture: understanding, control and impact. J Chem Technol Biotechnol 91(3):569-575.

11. Wang W, Singh S, Zeng DL, King K, Nema S 2007. Antibody structure, instability, and formulation. J Pharm Sci 96(1):1-26.

12. Geiger T, Clarke S 1987. Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. J Biol Chem 262(2):785-794.

13. Jenkins N, Murphy L, Tyther R 2008. Post-translational modifications of recombinant proteins: significance for biopharmaceuticals. Mol Biotechnol 39(2):113-118.

14. Robinson NE, Robinson AB 2001. Molecular Clocks. Proceedings of the National Academy of Sciences of the United States of America 98(3):944-949.

15. Noguchi S 2010. Structural changes induced by the deamidation and isomerization of asparagine revealed by the crystal structure of Ustilago sphaerogena ribonuclease U2B. Biopolymers 93(11):1003-1010.

16. Curnis F, Longhi R, Crippa L, Cattaneo A, Dondossola E, Bachi A, Corti A 2006. Spontaneous formation of L-isoaspartate and gain of function in fibronectin. J Biol Chem 281(47):36466-36476.

17. Charache S, Fox J, McCurdy P, Kazazian H, Jr., Winslow R, Hathaway P, van Beneden R, Jessop M 1977. Postsynthetic deamidation of hemoglobin Providence (beta 82 Lys replaced by Asn, Asp) and its effect on oxygen transport. J Clin Invest 59(4):652-658.

18. Friedman AR, Ichhpurani AK, Brown DM, Hillman RM, Krabill LF, Martin RA, Zurcher-Neely HA, Guido DM 1991. Degradation of growth hormone releasing factor analogs in neutral aqueous solution is related to deamidation of asparagine residues. Replacement of asparagine residues by serine stabilizes. Int J Pept Protein Res 37(1):14-20.

19. Solstad T, Flatmark T 2000. Microheterogeneity of recombinant human phenylalanine hydroxylase as a result of nonenzymatic deamidations of labile amide containing amino acids. Effects on catalytic and stability properties. Eur J Biochem 267(20):6302-6310.

20. Mamula MJ, Gee RJ, Elliott JI, Sette A, Southwood S, Jones PJ, Blier PR 1999. Isoaspartyl post-translational modification triggers autoimmune responses to self-proteins. J Biol Chem 274(32):22321-22327.

21. Doyle HA, Gee RJ, Mamula MJ 2007. Altered immunogenicity of isoaspartate containing proteins. Autoimmunity 40(2):131-137.

22. Moss CX, Matthews SP, Lamont DJ, Watts C 2005. Asparagine deamidation perturbs antigen presentation on class II major histocompatibility complex molecules. J Biol Chem 280(18):18498-18503.

23. Shi Y, Rhodes NR, Abdolvahabi A, Kohn T, Cook NP, Marti AA, Shaw BF 2013. Deamidation of Asparagine to Aspartate Destabilizes Cu, Zn Superoxide Dismutase, Accelerates Fibrillization, and Mirrors ALS-Linked Mutations. J Am Chem Soc 135(42):15897-15908.

24. Robinson NE, Robinson A. 2004. Molecular clocks: deamidation of asparaginyl and glutaminyl residues in peptides and proteins. ed., Cave Junction, OR: Althouse press.

25. Catak S, Monard G, Aviyente V, Ruiz-López MF 2006. Reaction Mechanism of Deamidation of Asparaginyl Residues in Peptides: Effect of Solvent Molecules. J Phys Chem A 110(27):8354-8365.

26. Capasso S, Mazzarella L, Sica F, Zagari A, Salvadori S 1993. Kinetics and mechanism of succinimide ring formation in the deamidation process of asparagine residues. J Chem Soc, Perkin Trans 2 (4):679-682.

27. Clarke S 1987. Propensity for spontaneous succinimide formation from aspartyl and asparaginyl residues in cellular proteins. Int J Pept Protein Res 30(6):808-821.

28. Reissner KJ, Aswad DW 2003. Deamidation and isoaspartate formation in proteins: unwanted alterations or surreptitious signals? Cell Mol Life Sci 60(7):1281-1295.

29. Riggs DL, Gomez SV, Julian RR 2017. Sequence and Solution Effects on the Prevalence of d-Isomers Produced by Deamidation. ACS Chem Biol 12(11):2875-2882.

30. Catak S, Monard G, Aviyente V, Ruiz-Lopez MF 2009. Deamidation of asparagine residues: direct hydrolysis versus succinimide-mediated deamidation mechanisms. J Phys Chem A 113(6):1111-1120.

31. Wright HT 1991. Nonenzymatic deamidation of asparaginyl and glutaminyl residues in proteins. Crit Rev Biochem Mol Biol 26(1):1-52.

32. Wakankar AA, Borchardt RT 2006. Formulation considerations for proteins susceptible to asparagine deamidation and aspartate isomerization. J Pharm Sci 95(11):2321-2336.

33. Pace AL, Wong RL, Zhang YT, Kao Y-H, Wang YJ 2013. Asparagine deamidation dependence on buffer type, pH, and temperature. J Pharm Sci 102(6):1712-1723.

34. Athmer L, Kindrachuk J, Georges F, Napper S 2002. The Influence of Protein Structure on the Products Emerging from Succinimide Hydrolysis. J Biol Chem 277(34):30502-30507.

35. Scotchler JW, Robinson AB 1974. Deamidation of glutaminyl residues: dependence on pH, temperature, and ionic strength. Anal Biochem 59(1):319-322.

36. Radkiewicz JL, Zipse H, Clarke S, Houk KN 1996. Accelerated Racemization of Aspartic Acid and Asparagine Residues via Succinimide Intermediates: An ab Initio Theoretical Exploration of Mechanism. J Am Chem Soc 118(38):9148-9155.

37. Robinson NE, Robinson ZW, Robinson BR, Robinson AL, Robinson JA, Robinson ML, Robinson AB 2004. Structure-dependent nonenzymatic deamidation of glutaminyl and asparaginyl pentapeptides. J Pept Res 63(5):426-436.

38. Yang X, Xu W, Dukleska S, Benchaar S, Mengisen S, Antochshuk V, Cheung J, Mann L, Babadjanova Z, Rowand J, Gunawan R, McCampbell A, Beaumont M, Meininger D, Richardson

D, Ambrogelly A 2013. Developability studies before initiation of process development: Improving manufacturability of monoclonal antibodies. mAbs 5(5):787-794.

39. Xie M Secondary structure and protein deamidation. J Pharm Sci 88(1):8-13.

40. Capasso S, Salvadori S 1999. Effect of the three-dimensional structure on the deamidation reaction of ribonuclease A. J Pept Res 54(5):377-382.

41. Sinha S, Zhang L, Duan S, Williams TD, Vlasak J, Ionescu R, Topp EM 2009. Effect of protein structure on deamidation rate in the Fc fragment of an IgG1 monoclonal antibody. Protein Sci 18(8):1573-1584.

42. Robinson NE, Robinson AB 2001. Prediction of protein deamidation rates from primary and three-dimensional structure. Proc Natl Acad Sci U S A 98(8):4367-4372.

43. Kosky AA, Razzaq UO, Treuheit MJ, Brems DN 1999. The effects of alpha-helix on the stability of Asn residues: deamidation rates in peptides of varying helicity. Protein Sci 8(11):2519-2523.

44. Wearne SJ, Creighton TE 1989. Effect of protein conformation on rate of deamidation: ribonuclease A. Proteins 5(1):8-12.

45. Haberger M 2014. Assessment of chemical modifications of sites in the CDRs of recombinant antibodies: Susceptibility vs. functionality of critical quality attributes. mAbs (Online) 6:327-339.

46. Jarasch A, Koll H, Regula JT, Bader M, Papadimitriou A, Kettenberger H 2015. Developability Assessment During the Selection of Novel Therapeutic Antibodies. J Pharm Sci 104(6):1885-1898.

47. Thiagarajan G, Semple A, James JK, Cheung JK, Shameem M 2016. A comparison of biophysical characterization techniques in predicting monoclonal antibody stability. mAbs 8(6):1088-1097.

48. Nowak C, J KC, S MD, Katiyar A, Bhat R, Sun J, Ponniah G, Neill A, Mason B, Beck A, Liu H 2017. Forced degradation of recombinant monoclonal antibodies: A practical guide. mAbs 9(8):1217-1230.

49. Robinson NE 2002. Protein Deamidation. Proc Natl Acad Sci U S A 99(8):5283-5288.

50. Sydow JF, Lipsmeier F, Larraillet V, Hilger M, Mautz B, Molhoj M, Kuentzer J, Klostermann S, Schoch J, Voelger HR, Regula JT, Cramer P, Papadimitriou A, Kettenberger H 2014. Structure-based prediction of asparagine and aspartate degradation sites in antibody variable regions. PLoS One 9(6):e100736.

51. Engen JR 2009. Analysis of Protein Conformation and Dynamics by Hydrogen/Deuterium Exchange MS. Anal Chem 81(19):7870-7875.

52. Berger A, Loewenstein A, Meiboom S 1959. Nuclear Magnetic Resonance Study of the Protolysis and Ionization of N-Methylacetamide1. J Am Chem Soc 81(1):62-67.

53. Phillips JJ, Buchanan A, Andrews J, Chodorge M, Sridharan S, Mitchell L, Burmeister N, Kippen AD, Vaughan TJ, Higazi DR, Lowe D 2017. Rate of Asparagine Deamidation in a Monoclonal Antibody Correlating with Hydrogen Exchange Rate at Adjacent Downstream Residues. Anal Chem 89(4):2361-2368.

54. Wang W, Meeler AR, Bergerud LT, Hesselberg M, Byrne M, Wu Z 2012. Quantification and characterization of antibody deamidation by peptide mapping with mass spectrometry. Int J Mass spectrom 312:107-113.

55. Glasoe PK, Long FA 1960. Use of glass electrodes to measure acidities in deuterium oxide1,2. J Phys Chem 64(1):188-190.

56. Sharff AJ, Rodseth LE, Spurlino JC, Quiocho FA 1992. Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. Biochemistry 31(44):10657-10663.

57. Chelius D, Rehder DS, Bondarenko PV 2005. Identification and characterization of deamidation sites in the conserved regions of human immunoglobulin gamma antibodies. Anal Chem 77(18):6004-6011.

58. Fukuda M, Takao T 2012. Quantitative Analysis of Deamidation and Isomerization in β2-Microglobulin by 18O Labeling. Anal Chem 84(23):10388-10394.

59. Ugur I, Marion A, Aviyente V, Monard G 2015. Why does Asn71 deamidate faster than Asn15 in the enzyme triosephosphate isomerase? Answers from microsecond molecular dynamics simulation and QM/MM free energy calculations. Biochemistry 54(6):1429-1439.

60. Gaza-Bulseco G, Li B, Bulseco A, Liu H 2008. Method to Differentiate Asn Deamidation That Occurred Prior to and during Sample Preparation of a Monoclonal Antibody. Anal Chem 80(24):9491-9498.

61. Iacob RE, Engen JR 2012. Hydrogen exchange mass spectrometry: are we out of the quicksand? J Am Soc Mass Spectrom 23(6):1003-1010.

62. Wilson DJ, Millisecond hydrogen exchange. In Hydrogen exchange mass spectrometry of proteins: Fundamentals, methods, and applications, Weis, D. D., Ed. Wiley: Chichester, 2016; pp 73-92.

63. Coales SJ, E SY, Lee JE, Ma A, Morrow JA, Hamuro Y 2010. Expansion of time window for mass spectrometric measurement of amide hydrogen/deuterium exchange reactions. Rapid Commun Mass Spectrom 24(24):3585-3592.

64. Walters BT, Jensen PF, Larraillet V, Lin K, Patapoff T, Schlothauer T, Rand KD, Zhang J 2016. Conformational Destabilization of Immunoglobulin G Increases the Low pH Binding Affinity with the Neonatal Fc Receptor. J Biol Chem 291(4):1817-1825.

65. Hamuro Y 2017. Determination of Equine Cytochrome c Backbone Amide Hydrogen/Deuterium Exchange Rates by Mass Spectrometry Using a Wider Time Window and Isotope Envelope. J Am Soc Mass Spectrom 28(3):486-497.

66. Pan J, Zhang S, Chou A, Hardie DB, Borchers CH 2015. Fast Comparative Structural Characterization of Intact Therapeutic Antibodies Using Hydrogen–Deuterium Exchange and Electron Transfer Dissociation. Anal Chem 87(12):5884-5890.

67. Rand KD, Zehl M, Jensen ON, Jørgensen TJD 2009. Protein Hydrogen Exchange Measured at Single-Residue Resolution by Electron Transfer Dissociation Mass Spectrometry. Anal Chem 81(14):5577-5584.

2.6 Appendix A

2.6.1 Protein expression and purification

A plasmid containing the gene for *E. coli* maltose-binding protein (MBP) was obtained as a gift from Ray Owens (plasmid #41134, Addgene, Cambridge, MA). The plasmid was modified by GenScript (Piscataway, NJ) to express His-tagged MBP by inserting a stop codon at end of the MBP gene. Primers were designed to substitute the codon for tryptophan (W) 169 to a codon for glycine (G). A Q5 Site-Directed Mutagenesis Kit (New England Biolabs, Ipswich, MA) along with designed primers was used for mutagenesis. The mutated plasmid was enriched in NEB 5-alpha competent *E. coli* (New England Biolabs) and then extracted and purified by a QIAprep Spin Miniprep Kit (Qiagen, Germantown, MD). The purified mutated plasmid sequence was confirmed by DNA sequencing (GENEWIZ, South Plainfield, NJ).

Purified plasmids containing MBP wild-type gene and MBP mutant (W169G) gene were used separately to transform One Shot BL21 Star (DE3) competent *E. coli* (Thermo Fisher Scientific, Waltham, MA). Transformed cells were cultured in LB medium supplemented with plasmid-selective spectinomycin at 37°C to an optical density of 0.6 and then chemically induced with IPTG at 1 mM for protein expression. After 4 hours, the resulting cells were pelleted, washed, and lysed. Proteins were purified from centrifuge-clarified lysate by nickel affinity chromatography on an ÄKTAprime FPLC (GE Healthcare, Piscataway, NJ) using an imidazole gradient for elution. Purified proteins were confirmed by intact mass analysis and SDS-PAGE. Purified proteins were buffer exchanged into protein storage buffer (20 mM phosphate, 100 mM sodium chloride, pH 7.0) at 4°C using 10 kDa MWCO centrifugal filters (EMD Millipore, Billerica, MA) then aliquoted, flash frozen on liquid N_2, and stored at −80°C for later use.

2.6.2 Proteolysis for peptide mapping

The protein samples were denatured by adding solid guanidine hydrochloride (GdnHCl) to 6 M concentration and incubating at 37°C for 1 hr prior to their digestion. The denatured samples were then diluted with 50 mM tris at pH 8.0 to decrease the GdnHCl concentration to less than 1 M. The digestion was performed at 37°C for 3 hr using both trypsin (Promega, Madison, WI) and Lys-C (Promega, Madison, WI) enzymes at an enzyme-to-protein ratio of 1:10. A short digestion time was used to minimize artifactual deamidation under digestion conditions[1-3] and Lys-C was used to enhance the digestion efficiency. The samples were flash frozen on liquid N_2 to stop the digestion reaction and stored at −80°C until analyzed.

2.6.3 LC-MS and LC-MS/MS analysis

LC-MS and LC-MS/MS analyses were performed on a LC system (Agilent 1260, Santa Clara, CA) coupled with a quadrupole time-of-flight (QTOF) mass analyzer (Agilent 6530, Santa Clara, CA). Tryptic peptides were desalted on a reversed-phase trap (Poroshell 120 EC-C18, 2.1× 5 mm, 2.7 μm particle diameter, Agilent, Santa Clara CA) and subsequently separated on a reversed-phase C8 column (Zorbax 300SB 2.1 × 50 mm, 1.8 μm particle diameter, Agilent, Santa Clara, CA). Mobile phases A and B consisted of 0.1% formic acid in water and 0.1% formic acid in acetonitrile, respectively. Unmodified and deamidated peptides were separated at room temperature using this gradient: 1-35% B in 60 min, 35-45% B from 60-65 min, 45-85% B from 65-75 min, 85-1% from 75-80 min, at a flow rate of 200 μL min^{-1}. For MS/MS experiments, fragmentation was carried out by collision induced dissociation (CID). All MS and MS/MS spectra were analyzed using MassHunter Qualitative Analysis (version B.07.00) software. Intact mass analysis of the proteins was carried out using a LC system (Agilent 1200 series, Santa Clara, CA)

coupled with a time-of-flight (TOF) mass analyzer (Agilent 6220, Santa Clara, CA) and deconvoluted masses were obtained using MassHunter Qualitative Analysis (version B.07.00) software.

2.6.4 Mapping deamidation sites

Deamidation-susceptible sites of MBP were assigned by LC-MS/MS analysis of the tryptic peptides of the stressed wild-type MBP sample. Deamidated peptides were identified by a mass increase of 0.98 Da compared to their unmodified peptides. The identity of each deamidated peptide and location of the deamidation site were confirmed using b and y ion fragments in the MS/MS spectra. **Figure 2.6.A-1** in the Appendix A shows, as an example, MS/MS spectra of the tryptic peptide 18-LVIWI**N**GDK-26 containing the deamidation-susceptible site (shown in bold). For the tryptic peptides 18-26 and 182-190, each of the corresponding deamidated peptides produced two resolved chromatographic peaks less than one minute apart with 0.98 Da mass increases for both peaks, consistent with the presence of both aspartate and isoaspartate products of deamidation. For the peptide 18-26, both the peaks were confirmed as deamidated products by MS/MS data. But for the peptide 182-190, MS/MS data were available only for the higher abundance peak of the two peaks. However, the isotopic distributions of the MS spectra from both peaks were the same, therefore both peaks were assigned as deamidated products. The two products of deamidation, aspartate and isoaspartate cannot be differentiated by the methods used in this study,[4] thus only total deamidation is reported. Furthermore, the succinimide intermediate form of the deamidation reaction was not observed, probably because the succinimide intermediate can be hydrolyzed during trypsin digestion at pH 8.[2]

2.6.5 Hydrogen exchange mass spectrometry

HX-MS experiments on wild-type and mutant MBP were performed using the LC-MS system described in the preceding section. Deuterium labeling, quenching, and sample injection were performed by an H/DX PAL robot (LEAP Technologies, Morrisville, NC). Labeling with deuterium was initiated by diluting the protein samples (in 20 mM sodium phosphate, 100 mM sodium chloride, pH 7.0) 20-fold with D_2O buffer (20 mM sodium phosphate, 100 mM sodium chloride, pD 7.0 corrected for the isotope effect[5]). Labeling at 25°C was performed for between 30 s and 12 hr, each labeling time in triplicate. The exchange reaction was quenched by diluting 1:1 with quench buffer (200 mM sodium phosphate, pH 2.5) at 1°C. The quenched protein samples were injected into a refrigerated column compartment maintained at 0°C containing an immobilized pepsin column, trap, and a reversed-phase column. The quenched protein samples were digested by passing them over an immobilized pepsin column (50 mm × 2.1 mm, pepsin was immobilized and packed in-house[6]). The resulting peptides were captured and desalted using the reversed-phase trap and separated on the reversed-phase column using a 15 min gradient of 1-85% acetonitrile with 0.1% formic acid. To prepare deuteration control samples, peptic peptides of the proteins were generated by passing the proteins in 0.1% formic acid over the pepsin column at a flow rate of 200 µL min^{-1} at 4°C. The resulting peptic peptides were collected, vacuum dried, and then reconstituted in 20 mM sodium phosphate, 100 mM sodium chloride, and pH 7.0 buffer. The peptide samples were diluted 20-fold with D_2O buffer for 2 hr using H/DX PAL robot. Quenching and peptide separation were performed as described above.

The peptic peptides of MBP were identified by LC-MS/MS analysis and MS/MS spectral analysis was performed using MassHunter Qualitative Analysis software (version B.07.00). The identified

peptic peptides with asparagine sites of interest were used for HX-MS data processing. The HX-MS data were processed with HDExaminer software (Sierra Analytics, Modesto, CA).

Table 2.3.A-1: Differential scanning calorimetry (DSC) of MBP. DSC thermograms were obtained in triplicate for MBP WT and MBP W169G on a Microcal VP-Capillary DSC. Scans were performed from 10°C to 90°C at a scan rate of 1°C per minute. Both proteins were at concentrations of 0.5 mg/mL in protein storage buffer (20 mM phosphate, 100 mM NaCl, pH 7.0). Data was analyzed using Origin 7.0 software package with the MicroCal LLC DSC plug-in. Thermal melting (T_m) values were obtained by iteratively fitting the thermograms to a single peak non-2-state model for unfolding. Average T_m values, standard deviations, and ΔT_m from triplicate experiments are provided in the table.

Protein	Avg. T_m (°C)	SD	ΔT_m (°C)
Wild-type	**54.0**	**0.1**	**-**
W169G	**49.2**	**0.1**	**-4.8**

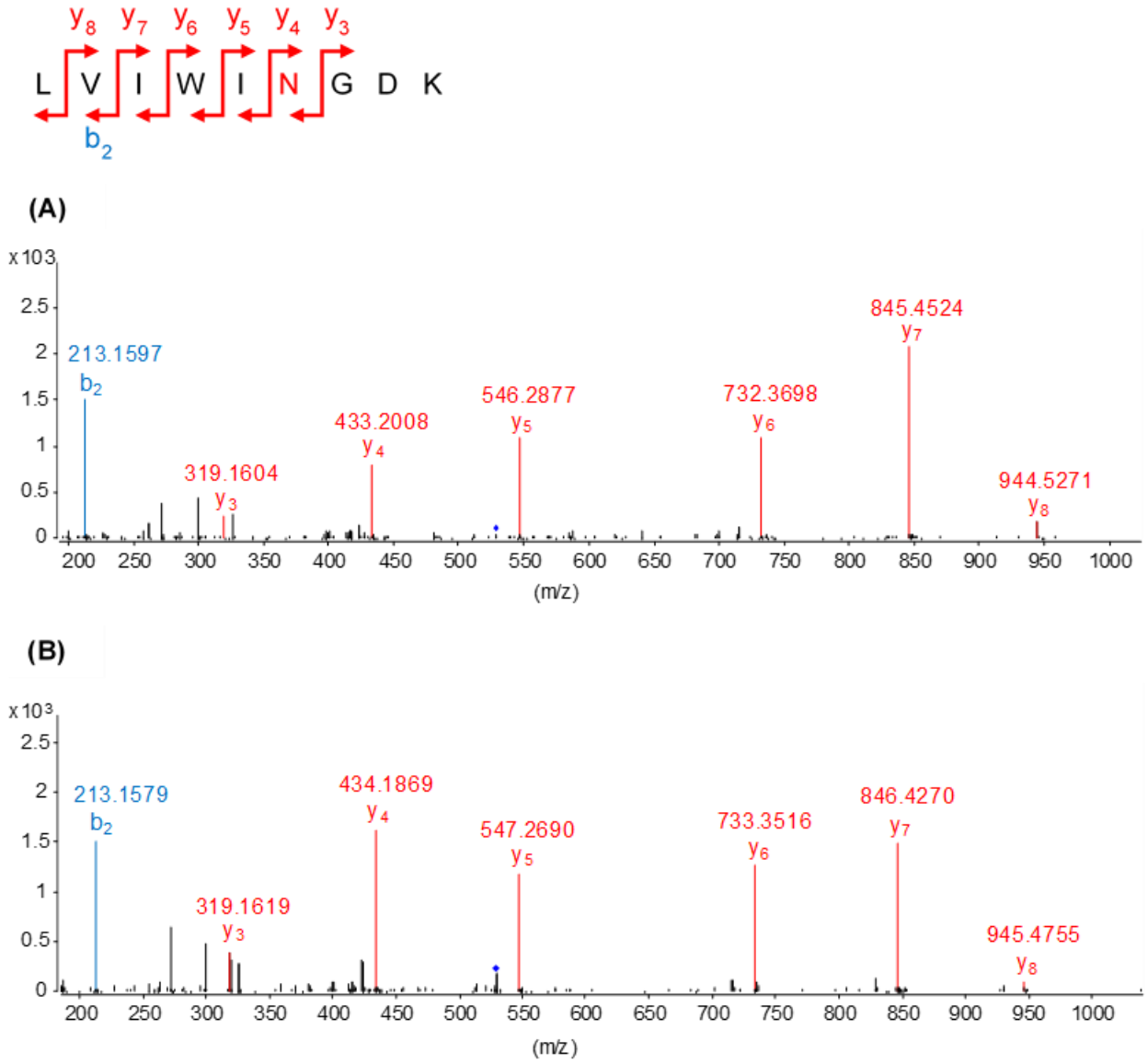


Figure 2.6.A-1: MS/MS spectra of the $(M+2H)^{2+}$ precursor ion of the (A) unmodified and (B) deamidated peptide, 18-LVIWINGDK-26. The deamidation site, N23, (shown in red) was identified by +0.98 Da mass difference of the y_4, y_5, y_6 and y_7 ions in the deamidated peptide relative to its unmodified peptide. The location of the deamidation site was further confirmed by y_3 and b_2 ions which did not show a mass difference in deamidated peptide relative to its unmodified peptide.

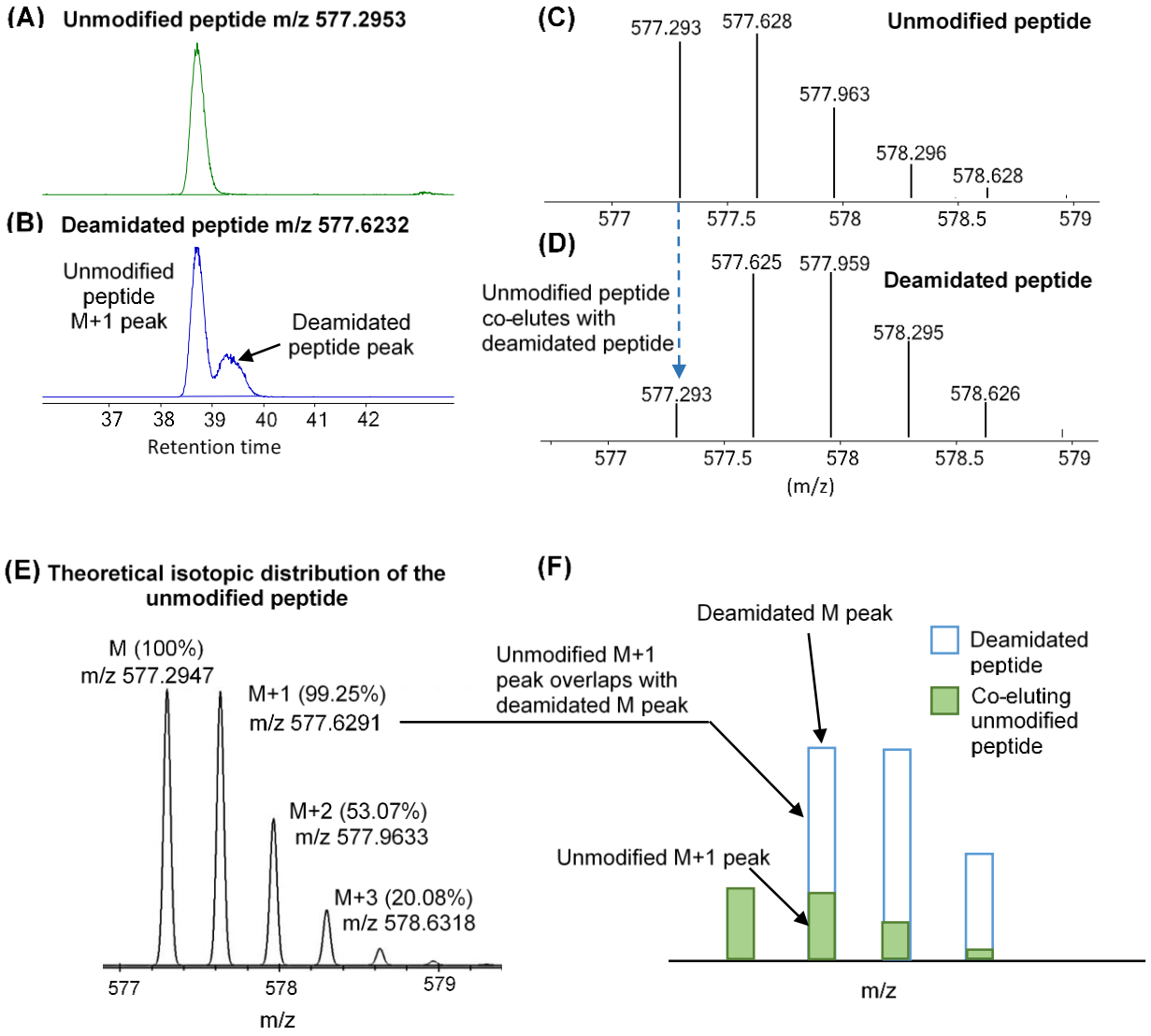


Figure 2.7.A-2: Correction for the coelution of unmodified and deamidated peptides of $(M+3H)^{3+}$ ion of the tryptic peptide 100-LYPFTWDAVRYNGK-113, adapted from reference 7,[7] (A) extracted ion chromatogram (EIC) of the unmodified peptide, (B) EIC of the deamidated peptides showing the co-elution of deamidated form with the unmodified form, (C) isotopic distribution of the unmodified peptide, (D) isotopic distribution of the deamidated peptide, (E) theoretical isotopic distribution of the unmodified peptide obtained using MassHunter Isotope Distribution Calculator software, and (F) cartoon representation of how M+1 peak of the unmodified peptide overlaps with the monoisotopic peak (M) of the deamidated peptide.

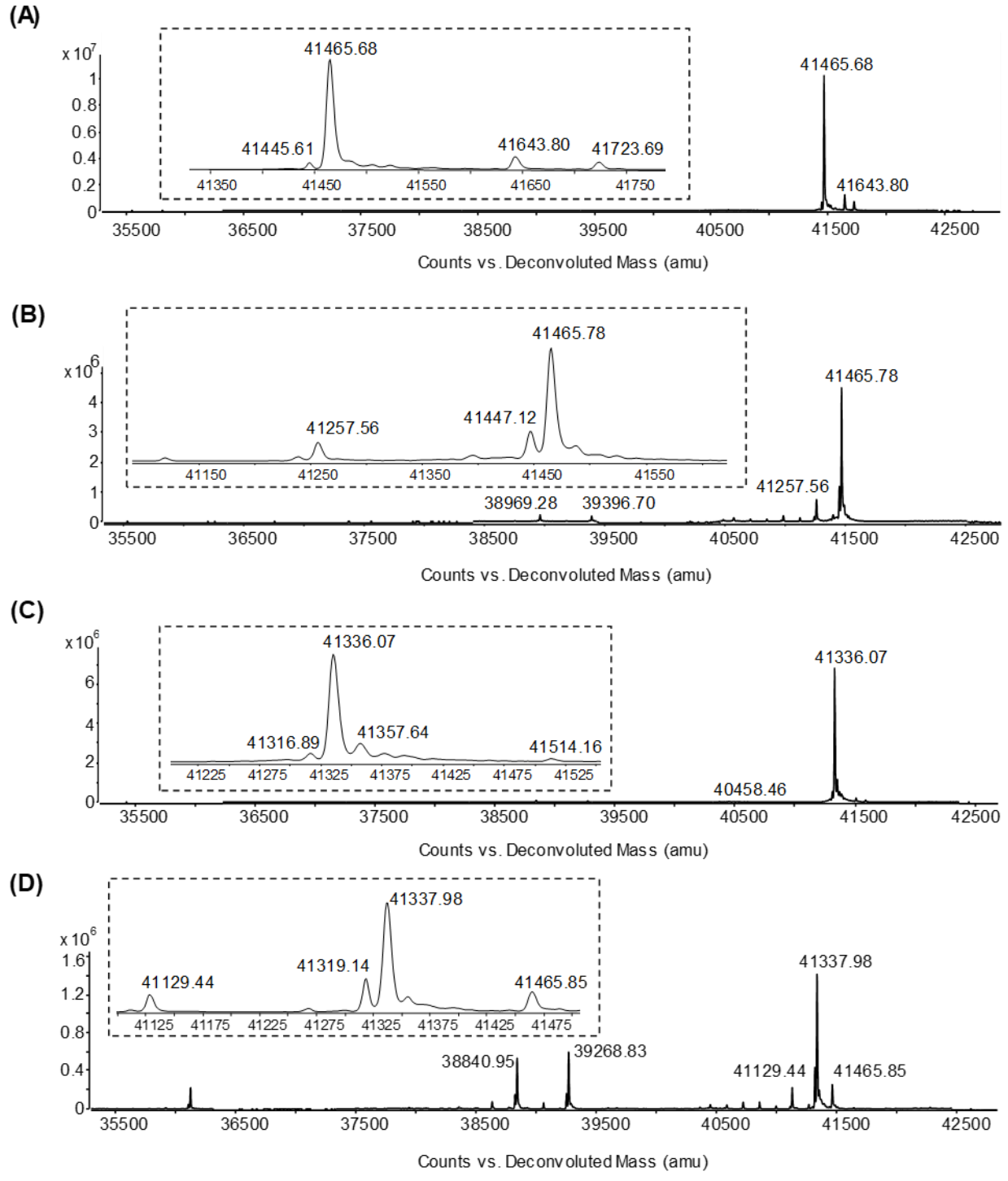


Figure 2.8.A-3: Deconvoluted mass spectra of proteins in the long-term stability study. (A) wild-type MBP at zero months, (B) wild-type MBP at 12 months, (C) mutant (W169G) MBP at zero months, and (D) mutant (W169G) MBP at 12 months. The insets show the expanded view of the major peak of each spectrum.

2.6.6 References

1. Li X, Cournoyer JJ, Lin C, O'Connor PB. Use of 18O labels to monitor deamidation during protein and peptide sample processing. J Am Soc Mass Spectrom 2008;19(6):855-864.

2. Chelius D, Rehder DS, Bondarenko PV. Identification and characterization of deamidation sites in the conserved regions of human immunoglobulin gamma antibodies. Anal Chem 2005;77(18):6004-6011.

3. Wang S, Kaltashov IA. An 18O-labeling assisted LC/MS method for assignment of aspartyl/isoaspartyl products from Asn deamidation and Asp isomerization in proteins. Anal Chem 2013;85(13):6446-6452.

4. DeGraan-Weber N, Zhang J, Reilly JP. Distinguishing Aspartic and Isoaspartic Acids in Peptides by Several Mass Spectrometric Fragmentation Methods. J Am Soc Mass Spectrom 2016;27(12):2041-2053.

5. Glasoe PK, Long FA. Use of glass electrodes to measure acidities in deuterium oxide1,2. J Phys Chem 1960;64(1):188-190.

6. Busby SA, Chalmers MJ, Griffin PR. Improving digestion efficiency under H/D exchange conditions with activated pepsinogen coupled columns. Int J Mass spectrom 2007;259(1):130-139.

7. Liu H, Wang F, Xu W, May K, Richardson D. Quantitation of asparagine deamidation by isotope labeling and liquid chromatography coupled with mass spectrometry analysis. Anal Biochem 2013;432(1):16-22.

3 CHAPTER 3: Identification of agitation-induced unfolding events causing aggregation of monoclonal antibodies using hydrogen exchange-mass spectrometry (HX-MS)

3.1 Introduction

Monoclonal antibodies (mAbs) are a major class of protein therapeutics that are increasingly used to treat various diseases including cancer and autoimmune disorders.[1-7] One of the key challenges in developing therapeutic mAbs is their propensity to form aggregates.[8-10] The presence of aggregates in therapeutic mAb formulations can enhance immunogenicity and might give rise to serious concerns in drug quality, safety, and efficacy.[11-13]

Interfacial stress is one of the major factors that induces aggregation of therapeutic mAbs during their manufacturing, shipping, long-term storage, and clinical administration.[14-16] Proteins, including mAbs, tend to adsorb to hydrophobic interfaces such as the air-water interface because of their amphiphilic and surface-active nature.[16,17] Agitation, another common stress that therapeutic mAbs encounter during manufacturing and shipping processes, increases interface exposure due to continuous regeneration of the air-water interface.[18-24] Surfactants such as polysorbates are commonly used in biopharmaceutical formulations to minimize interface-induced protein aggregation. Surfactants competitively adsorb to the air-liquid interface displacing proteins and effectively reducing the interface stress burden on a therapeutic protein. Decreased interface stress in turn reduces visible and sub-visible particle formation, a major safety concern for any parenterally dosed therapy.[25,26] However, polysorbates can degrade upon long-term storage affecting the stability of mAb formulations producing degradation products that are less soluble and lead to particle formation, ironically accelerating particle formation.[27-29] Minimizing the concentration of surfactants can reduce the rate of surfactant-based particle formation when it occurs, but finding the minimal concentration needed for protecting a therapeutic protein requires lengthy studies that are not conducive to rapid development. In either case, surfactant-based or

proteinaceous, the formation of sub-visible and visible particles can limit shelf-life. Some antibodies do not form particles under stress and others do; but the reasons for variable sensitivity to the air-water interface are unclear. An ability to design molecules impervious to interface stress requires a mechanistic understanding of these differences. Alternatively, minimizing the concentration of protective compounds like surfactants requires a way to quickly assess the particle forming propensity of any particular formulation.

Previous studies have reported interfacial properties of mAbs such as surface excess, surface pressure, hydrophobicity, and rheological behavior at the air-water interface.[14,30,31] It has been hypothesized that proteins adsorbed to the air-water interface may undergo conformational changes that expose hydrophobic regions and promote aggregate formation.[32,33] The studies by Leiske et al. have shown that hydrophobicity of mAbs increases with time at the air-water interface and that antibodies who form particles faster under agitation stress partition to the interface proportionately faster.[30] We sought to further understand the molecular events that occur at the air-water interface. Were the increases in hydrophobicity observed at the interface by Leiske et al. suggestive of global unfolding, or perhaps preferential orientation at the air-water interface to expose hydrophobic patches of structure. If local unfolding were observed, it would suggest that the molecule preferentially exposes a region to the interface. The addition of protective polysorbate was used as a control because these amphiphilic molecules coat the air-water interface eliminating its effect on particle formation. When present, one expects agitation to have no significant effect in these studies.

Hydrogen exchange-mass spectrometry (HX-MS) is a technique which provides information on structural dynamics of proteins at a sub-global resolution. HX-MS measures the isotopic

exchange of protein backbone amide hydrogens.[34,35] Protected from exchange by hydrogen bonds, the very hallmark of protein structure, by studying the variable rates by which these protons exchange, we are measuring protein dynamics. In HX-MS experiments, the protein of interest is diluted in a deuterated buffer allowing the exchange of hydrogen to deuterium to take place for predetermined time periods. Subsequently, the exchange reaction is quenched and the protein is subjected to acid proteolysis followed by LC-MS separation and mass analysis to determine the deuterium incorporation at a peptide level over time. Comparing the kinetics of deuterium exchange of a mAb with increased air-water interface exposure using mechanical agitation to a non-agitated control sample should identify any unfolding events promoted by interface exposure that might be responsible for downstream aggregation.

3.2 Experimental

Three IgG1 full-length mAbs called mAb-α, mAb-β and mAb-γ from Genentech were used for the study. The stock mAb-α sample was obtained in 10 mM histidine, 150 mM NaCl at pH 6.8 and the stock mAb-β and mAb-γ samples were obtained in 20 mM histidine-acetate buffer, 200 mM sucrose at pH 5.5. The stock mAb solutions were diluted in their corresponding buffers to get a concentration of 5 mg/mL for each mAb for the experiments.

3.2.1 Hydrogen exchange of the mAbs under agitation stress

The mAb samples (5 mg/mL) were diluted 12-fold in D_2O buffer for labelling in 2-cc glass vials. mAb-α was diluted in 10 mM histidine, 150 mM NaCl at pD 6.8 corrected for the isotope effect.[36] mAb-β and mAb-γ samples were diluted in 20 mM histidine-acetate buffer, 200 mM sucrose at pD 5.5 corrected for the isotope effect. Immediately after diluting with the deuterated buffer, the glass vials were sealed with caps and agitated using a mechanical arm shaker (built in-

house) at 100 cycles per minute at room temperature in the dark. For each mAb, agitation in deuterated buffer was performed for different time periods, ranging from 5 min to 24 hr (5, 15, 47, 167, 400, and 1440 min). Two replicates were performed for each time point and a fresh sample was used for each time duration for each mAb. At the end of each time point, agitation was stopped, and the samples were centrifuged at 10000 rpm for 3 min at 4 ˚C to remove any insoluble aggregates formed during the agitation stress. Then the labelling reaction was quenched by adding the supernatant into ice cold quench buffer (4 mM glycine, 4 M guanidine hydrochloride, and 0.5 TCEP at pH 2.3) at a 1:1 ratio. The quenched samples were stored at -70˚C until MS analysis. Along with the agitated samples, non-agitated control samples were prepared for each mAb following the same labelling procedure but without the agitation. Another set of experiments was performed in the presence of the surfactant polysorbate 20 (PS20) at a concentration of 0.05% v/v in the deuterated buffer to understand the effects of the surfactant on the mAbs under agitation stress.

3.2.2 LC-MS analysis

LC-MS analysis of the hydrogen exchanged samples was performed on a LC system (Agilent 1260 Infinity binary pump for washing steps, Waters nanoAcquity for gradient) coupled with an Orbitrap mass spectrometer (Thermo Orbitrap XL). For LC-MS analysis, the samples were thawed on ice and injected using an ice-cold syringe into a column compartment (as described previously[37]) that was maintained at 0 ˚C. The injected sample was first passed through an immobilized pepsin column (Porozyme) and the resulting peptides were trapped and desalted on a reversed-phase C18 trap (Acquity BEH VanGuard 2.1x5 mm) before separation using a reversed-phase C18 column (Acquity BEH 1 x 50mm, 1.7 uM) and then ionized into the mass spectrometer

by electrospray for mass measurement. Mobile phase A contained 0.1% formic acid and 0.04% trifluoroacetic acid in water and mobile phase B contained 0.1% formic acid and 0.04% trifluoroacetic acid (TFA) in acetonitrile. The peptides were separated and eluted using a 13 min gradient with an increasing mobile phase B from 5% to 35% at a flow rate of 50 µL/min.

Peptic peptides of each mAb were identified by LC-MS/MS analysis as described previously[37] using Mascot v2.1. Extracted ion chromatograms for each peptide at each timepoint were prepared using ExMS software v1[38] and in-house python scripts were used to define the level of deuterium carried and fraction of molecules in each observed population using an n-population binomial analysis described previously.[39] The contact map was generated using CMView software tool v1.1 and contacts were defined as any two residues whose backbone atoms are within 12 angstroms.[40] The Kabat numbering scheme was used for numbering antibody residues and we use 'H' and 'L' where appropriate to distinguish heavy or light chains, respectively.[41] All-atom homology models were constructed exactly as described recently.[42]

3.3 Results and discussion

Aggregation of therapeutic mAbs might affect the quality, safety, or efficacy of a drug product; therefore, it is essential to understand how various factors promote aggregation during manufacturing, shipping, storage, and clinical administration. Insight into the particular regions of individual mAbs that are most likely to denature when exposed to interfaces might enable engineering of molecules with fewer liabilities. Further, early identification of high-risk molecules allows later stages of the drug development process to prepare and potentially avoid timely delays in producing medicine for patients. Here, our goal was to explore whether the magnitudes of local

changes in structural dynamics of mAbs assessed by HX-MS under agitation stress might be related to their particle formation propensity in this environment.

Two antibodies were selected with notably different sensitivities to the air-water interface; mAb-β and mAb-γ. mAb-γ [30,43] (also known as mAb-1 in [30] and [43]) had a particularly high agitation-induced aggregation risk when a protective surfactant was not present, and was identified as being among the highest risk antibodies observed in previous studies. There, the tendency of antibodies to aggregate upon agitation was correlated with increased hydrophobic exposure at the air-liquid interface, and the authors hypothesized that this exposure was a result of a change in the folded state of the molecule. mAb-β (also known as mAb-7 in [14]) formed visible particles in the absence of a surfactant but was far less sensitive to aggregate formation than mAb-γ (also known as mAb-3 in [14]).[14] There were no measurements available on the aggregation propensity of mAb-α under agitation stress prior to this work.

If the air-water interface destabilized and partially unfolded antibody structure increasing their aggregation propensity exponentially when under agitation due to increased interface turnover with respect to time, we reasoned that increased relative sensitivity between molecules would be the result of either more extensive structural destabilization, or destabilization of a greater magnitude (perhaps local denaturation) upon interface exposure. Our hydrogen exchange experiments were designed to explore whether agitated samples, having an increased air-water surface area per unit time, had exchanged more deuterium when compared to non-agitated controls as would be consistent with destabilization or local unfolding occurring preferentially at the air-water interface. Therefore, both hydrogen exchange and agitation of the mAbs were performed simultaneously for different time periods ranging from 5 min to 24 hr. The mAb samples were

subjected to agitation stress immediately after they were diluted in an excess of deuterated buffer. Agitation stress was provided by shaking the samples using a mechanical arm shaker at room temperature in the dark. The samples were agitated both in the presence and absence of surfactant, polysorbate 20, to determine whether the surfactant could protect mAb molecules against agitation-induced destabilization. The control samples were prepared without agitation. As the focus of the study was to investigate the precursors to aggregation; therefore, the soluble protein fraction was separated from the aggregates by centrifuging the samples at the end of each agitation time period. Subsequently, only the soluble protein fraction was quenched to stop the H/D exchange reaction and analyzed by LC-MS to measure the differences in deuterium exchange between the agitated and the non-agitated samples.

Under agitation stress, all three mAbs formed visible aggregates in the absence of polysorbate 20. Visible aggregates were observed after 15 mins of agitation stress in mAb-α and mAb-γ samples and not until 167 minutes of agitation for mAb-β. None of the three mAbs formed visible aggregates when they were agitated in the presence of polysorbate 20 over the time duration of the experiment.

As shown in **Figure 3.1A**, mAb-α underwent faster exchange in some regions of both the Fab and Fc domains under agitation stress, consistent with increased rates of local unfolding (or deprotection). The faster exchanging region in the Fab domain was located in the light chain CDR domain and covered by the peptide L87-L115. Deuterium uptake curves for this peptide are shown in **Figure 3.1B**. Peptide L87-L115 contained 25 exchange sites (shorter peptides were not available covering this region) and had a substantial increase in deuterium uptake at early time points, with a difference of 0.7 Da, in the agitated molecule compared to its non-agitated control. In the Fc

domain, faster exchange was observed in two regions which were covered by the peptides H231-H248 and H376-H394 (their deuterium uptake curves are shown in **Figure 3.1C** and **3.1D**, respectively). Peptide H231-H248, containing 12 exchange sites, and peptide H376-H394, containing 14 exchange sites, had 0.7 Da and 0.5 Da differences in deuterium uptake in the agitated state at early time points compared to their non-agitated controls. Although deuterium uptake differences in Fab peptide L87-L115 and Fc peptide H231-H248 were similar, the 0.7 Da difference of the Fab peptide L87-L115 arose from a longer stretch of residues and thus the magnitude of change in exchange rate is notably smaller. The exchange rate differences in the Fab and Fc domains indicate that extent of local destabilization/unfolding occurs under agitation stress could be different from domain to domain within the same molecule. The deuterium uptake plots in **Figure 3.1** also compare the deuterium exchange of the peptides when agitated in the presence of polysorbate 20 and its non-agitated control. In the presence of polysorbate 20, the peptides did not have any substantial differences in deuterium exchange compared to the non-agitated control peptides with and without polysorbate 20. This demonstrates that polysorbate 20 was capable of protecting mAb structure against local unfolding accelerated by agitation stress.

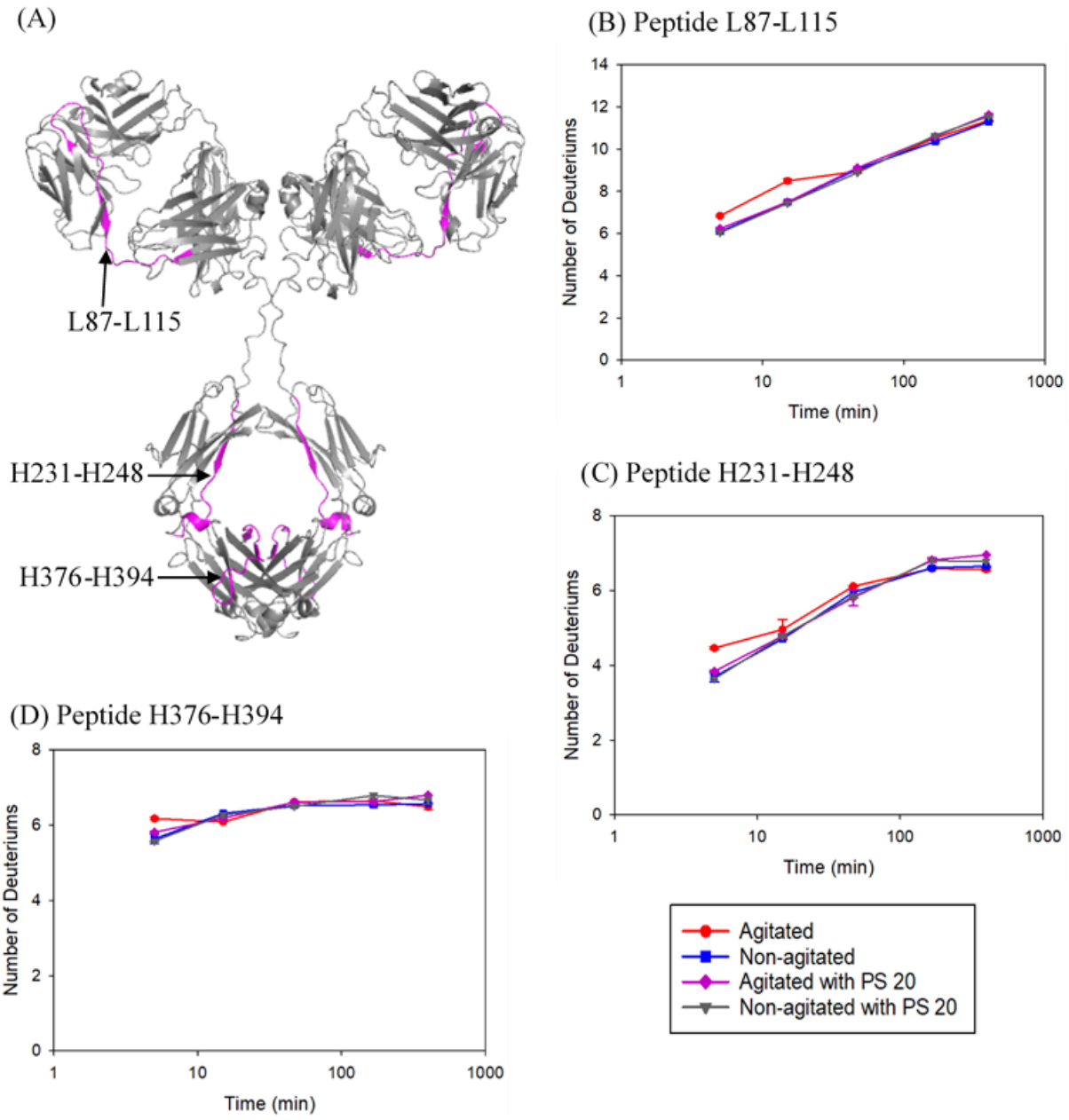


Figure 3.1: Regions of mAb-α that increase in observed deuterium incorporation when agitated. (A) Homology model showing the location of peptides (B) L87-L115, (C) H231-H248, (D) and H376-H394.

In mAb-β, faster exchange was observed only in the Fab domain upon agitation, these regions are shown in **Figure 3.2A**. Deuterium uptake curves of the representative peptides covering the faster exchanging regions are shown in **Figure 3.2B, 3.2C** and **3.2D**. All of these regions are partly located in the CDR regions of mAb-β. **Figure 3.2B** shows the peptide H33-H52 representing the region that exchanged faster in the heavy chain which contained 17 exchange sites. This peptide gave a deuterium uptake difference of 1.4 Da upon agitation compared to its non-agitated control. **Figure 3.2C** and **3.2D** show the peptide L53-L70 and peptide L47-L53, respectively, which covers the regions that exchanged faster in the light chain. The two light chain peptides; L53-L70 with 15 exchange sites and L47-L53 with 5 exchange sites, had deuterium uptake differences of 0.9 Da and 2.0 Da, respectively, in the agitated sample compared to its non-agitated control. The light chain peptide L47-L53, which is very short, gave the largest deuterium uptake difference among the three peptides, indicating that this region of the molecule underwent a relatively larger magnitude of local unfolding. mAb-β also did not have any substantial differences in deuterium uptake upon agitation in the presence of polysorbate 20 relative to its non-agitated controls, as shown in the deuterium uptake curves.

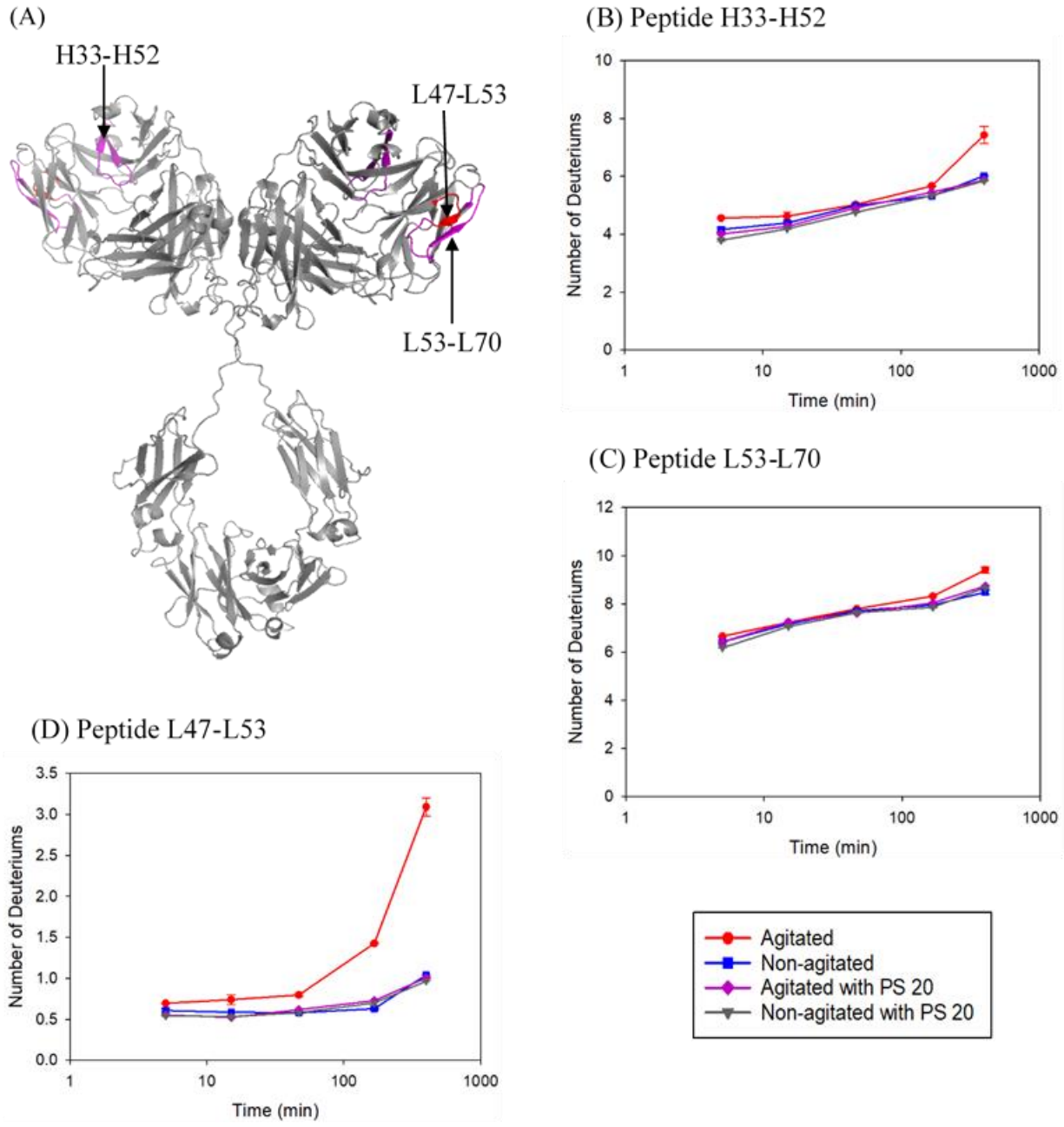


Figure 3.2: Regions of mAb-β that increase in observed deuterium incorporation when agitated. (A) Homology model showing the locations of affected regions and representative peptides (B) H33-H52 and (C) L53-L70, both shown in magenta to reflect similar magnitudes to all regions affected by agitation in mAb-α. (D) A much larger effect shown in red in panel A, and representative peptide L47-L53.

The effect of agitation stress on mAb-γ was different from mAb-α and mAb-β. Most regions of the molecule exchanged faster upon agitation compared to the non-agitated control. These faster exchanging regions are shown in **Figure 3.3A** in both red and yellow. Similar to the previous two molecules, inclusion of polysorbate 20 fully abrogated the effects of agitation. However, in contrast, the pre-aggregated soluble state of mAb-γ displayed both EX1 and EX2 exchange kinetics when agitated without polysorbate 20.

Under the EX2 kinetic regime, which is most common for native HX, an amide must undergo many de-protection/protection events before exchange occurs because the protein is well folded (e.g., protection rate much faster than deprotection and chemical exchange at physiological conditions). EX1 exchange kinetics, characterized in HX-MS by multi-modal isotopic distributions (see Appendix B **Figure 3.4.B-1**), takes place when the chemical exchange rate far exceeds the rate of protection for any particular amide causing each deprotection event to result in full exchange.[44,45]

Tracking the relative intensities of each deuterium distribution for HX-MS peptides showing EX1 kinetics, as shown in **Figure 3.3B** as a measure of the fraction of exchanged molecules, uncovered the presence of two cooperatively unfolding groups: regions with more than 50% conversion before our 167 min timepoint, shown in cyan, are termed “fast”, and peptides requiring more than 400 min, shown in orange, are termed “slow”. All peptides with EX1 kinetics were localized in the Fab domain.

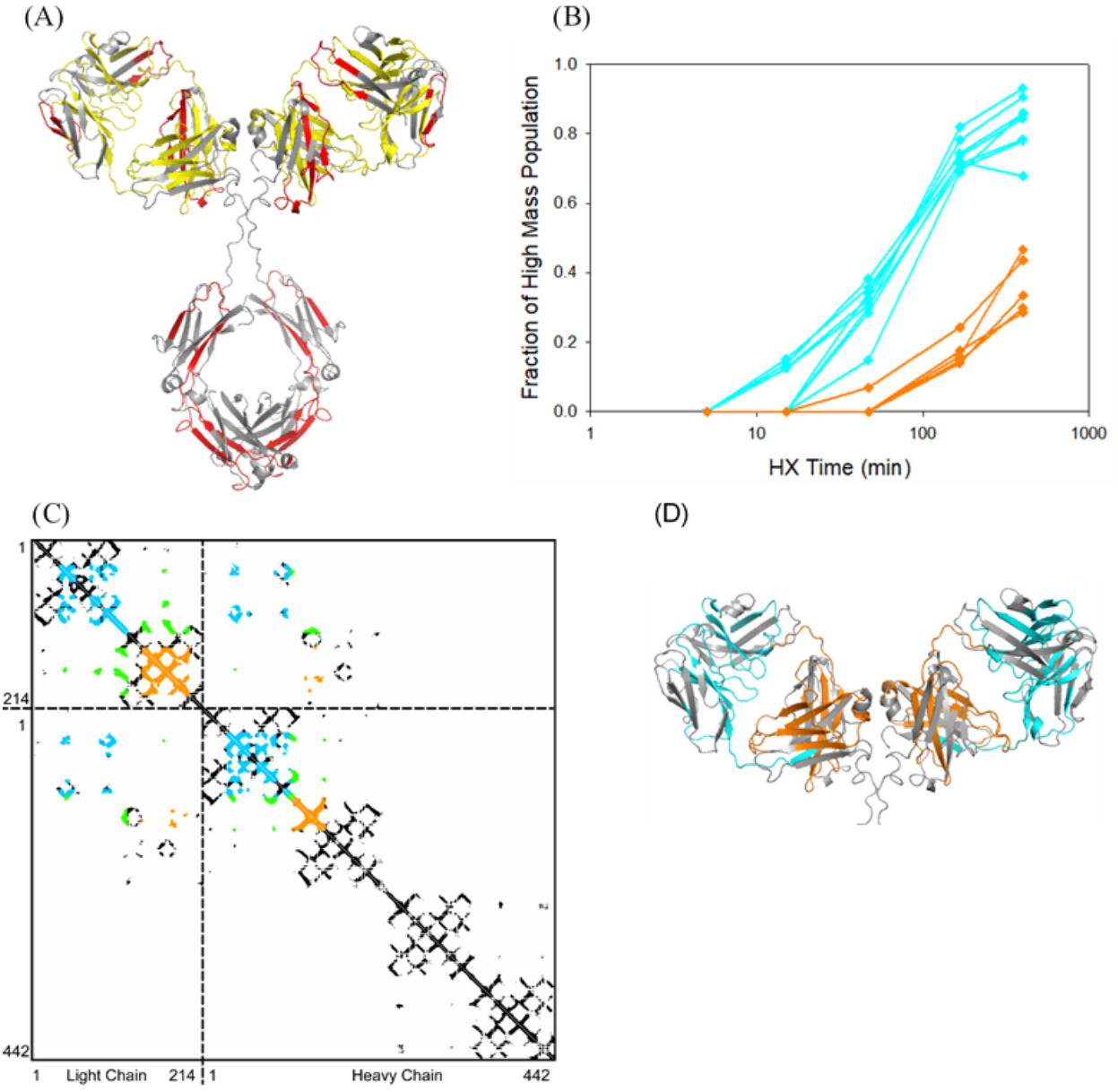


Figure 3.3: (A) Homology model of mAb-γ showing EX2 kinetic regions in red and EX1 kinetic regions in yellow. (B) Plot of fraction of the high mass population against HX time for EX1 kinetic peptides. Cyan and orange represent EX1 peptides having fast (reached 50% high mass population before 167 timepoint) and slow (required more than 400 min to reach 50% high mass population) agitation-induced unfolding, respectively. (C) Contact map showing the contacts (green off-diagonal) between the residues of the fast (cyan) group and slow (orange) group EX1 peptides. (D) Fab domain of mAb-γ showing the EX1 regions with fast agitation-induced unfolding in cyan and slow agitation-induced unfolding in orange.

To further explore the unfolding behavior in the EX1 kinetic regions of mAb-γ, a contact map (**Figure 3.3C**) was generated, using the structure of Raptiva as representative of the IgG1 fold, that shows all residues having a backbone atom from each separated by less than or equal to 12 Angstroms in the folded structure off-diagonal. Residues from the fast (cyan) or slow (orange) groups are highlighted. Contacts forming the interface of cyan and orange groups are colored green, i.e., a contact between a fast and slow residue. Clearly, seeing EX1 peptides separated by hundreds of residues, or even heavy/light chain that both cluster in exchange time and in structure suggests cooperative unfolding. Further, that the faster kinetics occur in Fab tips interfacing directly with residues showing slower EX1 kinetics in the constant regions (see **Figure 3.3D**) suggests that preferential exposure of the variable domain to the air-water interface leads to unfolding in the Fab tips first, and then in the heavy and light chains of the CH1 domain. These observations strongly support a preferential interaction of the variable Fab domain with the air-water interface whereby the Fab tips favor the air-water interface.

The results of our study clearly show that all three molecules undergo increases in exchange, consistent with increased local unfolding upon agitation. The magnitude and extent of said behavior was different among the three molecules upon exposure to the air-water interface and appear to be related to their aggregation propensities while under agitation. mAb-γ became the most destabilized upon agitation, we observed cooperative, sub-global unfolding in some regions of the Fab domain, confirming that exposure to air-water interfaces can destabilize and unfold the antibody. mAb-γ also formed visible aggregates after only 15 min of agitation stress, much faster than the other two molecules and consistent with cooperative unfolding at the air-water interface. mAb-α also formed visible aggregates after 15 min of agitation stress, but the solution was still relatively clear (results not shown). We postulate that this is because the destabilized, aggregation

prone regions of the molecule became unfolded in a much smaller fraction of molecules per unit time, consistent with smaller agitation-induced differences in exchange rates. mAb-β was the least destabilized upon agitation among the three molecules, forming visible aggregates only after 167 min of agitation stress in this study; however, size exclusion chromatography analysis of mAb-α and mAb-β (see Appendix B **Figure 3.5.B-2**, **Table 3.1.B-1** and **Figure 3.6.B-3**, **Table 3.2.B-2**, respectively) showed that the presence of high molecular species is relatively higher in agitated mAb-β samples relative to agitated mAb-α samples, indicating a higher tendency of mAb-β to form soluble aggregates.

Taken together, these results demonstrate that the extent of destabilization and magnitude of unfolding in antibody structure under agitation stress were different from molecule to molecule and are likely related to their propensities for agitation induced aggregation. In all cases, the fast regions showing EX1 kinetics that are likely the most destabilized regions of structure in these molecules also contained or flanked CDRs which are the only regions of sequence difference between all molecules in the study. If this issue had been discovered earlier in development, potentially an alternative molecule with similar affinity but different CDRs could have been selected. Further, because HX-MS pinpoints exactly where the unfolding events begin, it offers an opportunity to selectively re-engineer the molecule avoiding the problem altogether.

In all cases, polysorbate mitigated the effects of agitation in this small-scale study. However, shelf-life limiting particles will form over time from the precipitation of free fatty acid byproducts of polysorbate degradation creating a real driver to minimize or eliminating polysorbate.[27] The technique used here could also facilitate rapid concentration optimization and

testing of additives needed for mechanical protection, such as polysorbate or alternatives, without requiring lengthy, larger scale stability exercises that consume precious development time.

3.4 References

1. Castelli MS, McGonigle P, Hornby PJ. The pharmacology and therapeutic applications of monoclonal antibodies. Pharmacol Res Perspect 2019;7(6):e00535.

2. Buss NA, Henderson SJ, McFarlane M, Shenton JM, de Haan L. Monoclonal antibody therapeutics: history and future. Curr Opin Pharmacol 2012;12(5):615-622.

3. Shepard HM, Phillips GL, C DT, Feldmann M. Developments in therapy with monoclonal antibodies and related proteins. Clin Med (Lond) 2017;17(3):220-232.

4. Nelson AL, Dhimolea E, Reichert JM. Development trends for human monoclonal antibody therapeutics. Nat Rev Drug Discov 2010;9(10):767-774.

5. Bayer V. An Overview of Monoclonal Antibodies. Semin Oncol Nurs 2019;35(5):150927.

6. Singh S, Kumar NK, Dwiwedi P, Charan J, Kaur R, Sidhu P, Chugh VK. Monoclonal Antibodies: A Review. Curr Clin Pharmacol 2018;13(2):85-99.

7. Elgundi Z, Reslan M, Cruz E, Sifniotis V, Kayser V. The state-of-play and future of antibody therapeutics. Adv Drug Deliv Rev 2017;122:2-19.

8. Vázquez-Rey M, Lang DA. Aggregates in monoclonal antibody manufacturing processes. Biotechnology and Bioengineering 2011;108(7):1494-1508.

9. Shah M. Commentary: New perspectives on protein aggregation during Biopharmaceutical development. Int J Pharm 2018;552(1-2):1-6.

10. Cromwell MEM, Hilario E, Jacobson F. Protein aggregation and bioprocessing. The AAPS Journal 2006;8(3):E572-E579.

11. Wang W, Singh SK, Li N, Toler MR, King KR, Nema S. Immunogenicity of protein aggregates--concerns and realities. Int J Pharm 2012;431(1-2):1-11.

12. Moussa EM, Panchal JP, Moorthy BS, Blum JS, Joubert MK, Narhi LO, Topp EM. Immunogenicity of Therapeutic Protein Aggregates. J Pharm Sci 2016;105(2):417-430.

13. Rosenberg AS, Verthelyi D, Cherney BW. Managing uncertainty: a perspective on risk pertaining to product quality attributes as they bear on immunogenicity of therapeutic proteins. J Pharm Sci 2012;101(10):3560-3567.

14. Shieh IC, Patel AR. Predicting the Agitation-Induced Aggregation of Monoclonal Antibodies Using Surface Tensiometry. Molecular Pharmaceutics 2015;12(9):3184-3193.

15. Thomas CR, Geer D. Effects of shear on proteins in solution. Biotechnol Lett 2011;33(3):443-456.

16. Li J, Krause ME, Chen X, Cheng Y, Dai W, Hill JJ, Huang M, Jordan S, LaCasse D, Narhi L, Shalaev E, Shieh IC, Thomas JC, Tu R, Zheng S, Zhu L. Interfacial Stress in the Development of Biologics: Fundamental Understanding, Current Practice, and Future Perspective. The AAPS Journal 2019;21(3):44.

17. Malmsten M. Formation of Adsorbed Protein Layers. Journal of Colloid and Interface Science 1998;207(2):186-199.

18. Wang W. Protein aggregation and its inhibition in biopharmaceutics. Int J Pharm 2005;289(1-2):1-30.

19. Mahler HC, Friess W, Grauschopf U, Kiese S. Protein aggregation: pathways, induction factors and analysis. J Pharm Sci 2009;98(9):2909-2934.

20. Jayaraman M, Buck PM, Ignatius AA, King KR, Wang W. Agitation-induced aggregation and subvisible particulate formation in model proteins. Eur J Pharm Biopharm 2014;87(2):299-309.

21. Mahler HC, Müller R, Friess W, Delille A, Matheus S. Induction and analysis of aggregates in a liquid IgG1-antibody formulation. Eur J Pharm Biopharm 2005;59(3):407-417.

22. Kiese S, Papppenberger A, Friess W, Mahler HC. Shaken, not stirred: mechanical stress testing of an IgG1 antibody. J Pharm Sci 2008;97(10):4347-4366.

23. Fleischman ML, Chung J, Paul EP, Lewus RA. Shipping-Induced Aggregation in Therapeutic Antibodies: Utilization of a Scale-Down Model to Assess Degradation in Monoclonal Antibodies. J Pharm Sci 2017;106(4):994-1000.

24. Rudiuk S, Cohen-Tannoudji L, Huille S, Tribet C. Importance of the dynamics of adsorption and of a transient interfacial stress on the formation of aggregates of IgG antibodies. Soft Matter 2012;8(9):2651-2661.

25. Lee HJ, McAuley A, Schilke KF, McGuire J. Molecular origins of surfactant-mediated stabilization of protein drugs. Advanced Drug Delivery Reviews 2011;63(13):1160-1171.

26. Khan TA, Mahler H-C, Kishore RSK. Key interactions of surfactants in therapeutic protein formulations: A review. European Journal of Pharmaceutics and Biopharmaceutics 2015;97:60-67.

27. Tomlinson A, Demeule B, Lin B, Yadav S. Polysorbate 20 Degradation in Biopharmaceutical Formulations: Quantification of Free Fatty Acids, Characterization of Particulates, and Insights into the Degradation Mechanism. Molecular Pharmaceutics 2015;12(11):3805-3815.

28. Doshi N, Demeule B, Yadav S. Understanding Particle Formation: Solubility of Free Fatty Acids as Polysorbate 20 Degradation Byproducts in Therapeutic Monoclonal Antibody Formulations. Molecular Pharmaceutics 2015;12(11):3792-3804.

29. Kerwin BA. Polysorbates 20 and 80 Used in the Formulation of Protein Biotherapeutics: Structure and Degradation Pathways. Journal of Pharmaceutical Sciences 2008;97(8):2924-2935.

30. Leiske DL, Shieh IC, Tse ML. A Method To Measure Protein Unfolding at an Air–Liquid Interface. Langmuir 2016;32(39):9930-9937.

31. Koepf E, Eisele S, Schroeder R, Brezesinski G, Friess W. Notorious but not understood: How liquid-air interfacial stress triggers protein aggregation. Int J Pharm 2018;537(1-2):202-212.

32. Yano YF. Kinetics of protein unfolding at interfaces. Journal of Physics: Condensed Matter 2012;24(50):503101.

33. Yano YF, Uruga T, Tanida H, Toyokawa H, Terada Y, Takagaki M, Yamada H. Driving Force Behind Adsorption-Induced Protein Unfolding: A Time-Resolved X-ray Reflectivity Study on Lysozyme Adsorbed at an Air/Water Interface. Langmuir 2009;25(1):32-35.

34. Englander SW, Kallenbach NR. Hydrogen exchange and structural dynamics of proteins and nucleic acids. Quarterly Reviews of Biophysics 1983;16(4):521-655.

35. Wales TE, Engen JR. Hydrogen exchange mass spectrometry for the analysis of protein dynamics. Mass Spectrometry Reviews 2006;25(1):158-170.

36. Glasoe PK, Long FA. USE OF GLASS ELECTRODES TO MEASURE ACIDITIES IN DEUTERIUM OXIDE1,2. The Journal of Physical Chemistry 1960;64(1):188-190.

37. Mayne L, Kan ZY, Chetty PS, Ricciuti A, Walters BT, Englander SW. Many overlapping peptides for protein hydrogen exchange experiments by the fragment separation-mass spectrometry method. J Am Soc Mass Spectrom 2011;22(11):1898-1905.

38. Kan ZY, Mayne L, Chetty PS, Englander SW. ExMS: data analysis for HX-MS experiments. J Am Soc Mass Spectrom 2011;22(11):1906-1915.

39. Walters BT. 2013. Large protein folding and dynamics studied by advanced hydrogen exchange methods. Biochemistry and Molecular Biophysics, ed.: University of Pennsylvania: Philadelphia, PA. p p. 142.

40. Vehlow C, Stehr H, Winkelmann M, Duarte JM, Petzold L, Dinse J, Lappe M. CMView: interactive contact map visualization and analysis. Bioinformatics 2011;27(11):1573-1574.

41. Kabat EA, Wu, Tai Te, Bilofsky, Howard. 1979. Sequences of immunoglobulin chains: tabulation and analysis of amino acid sequences of precursors, V-regions, C-regions, J-chain and 2-microglobulins. ed.: National Institute of Health.

42. Izadi S, Patapoff TW, Walters BT. Multiscale Coarse-Grained Approach to Investigate Self-Association of Antibodies. Biophys J 2020;118(11):2741-2754.

43. Kannan A, Shieh IC, Leiske DL, Fuller GG. Monoclonal Antibody Interfaces: Dilatation Mechanics and Bubble Coalescence. Langmuir 2018;34(2):630-638.

44. Engen JR, Wales TE, Chen S, Marzluff EM, Hassell KM, Weis DD, Smithgall TE. Partial cooperative unfolding in proteins as observed by hydrogen exchange mass spectrometry. Int Rev Phys Chem 2013;32(1):96-127.

45. Ferraro DM, Lazo ND, Robertson AD. EX1 Hydrogen Exchange and Protein Folding. Biochemistry 2004;43(3):587-594.

3.5 Appendix B

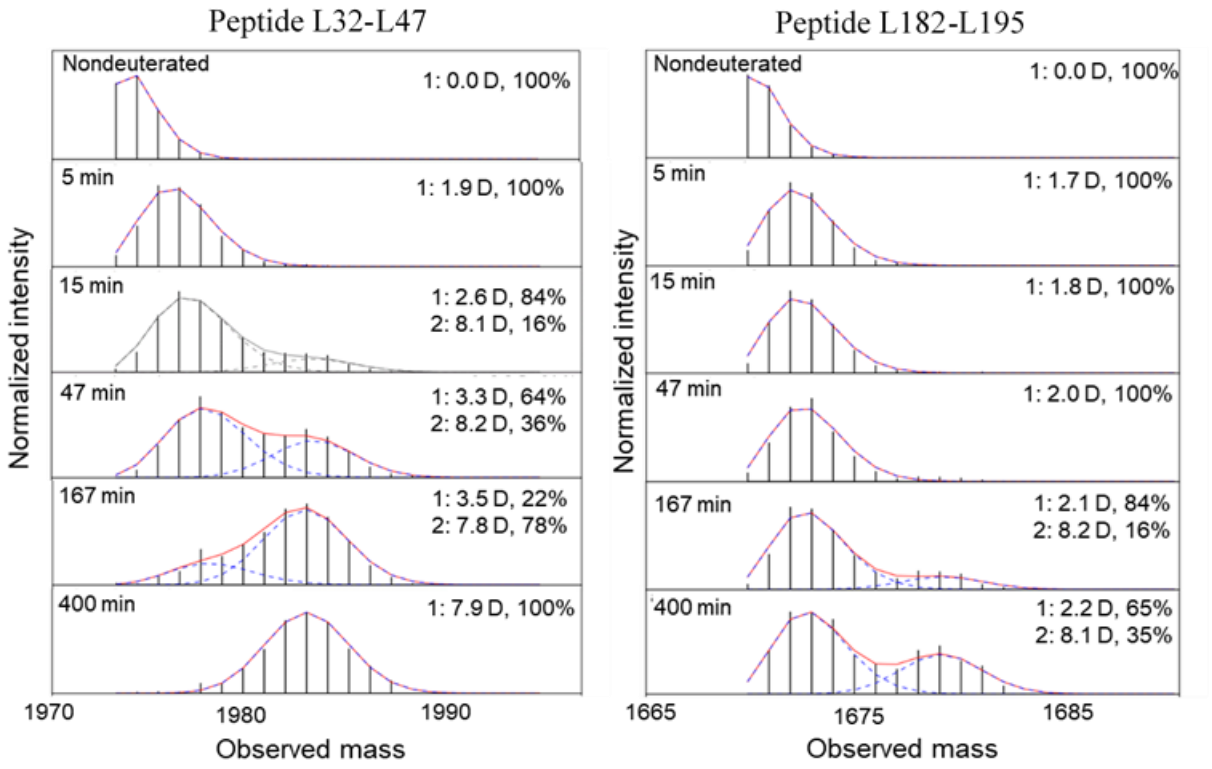


Figure 3.4.B-1: The mass spectra of two representative peptides of agitated mAb-γ displaying EX1 kinetics characterized by the presence of bimodal mass distributions.

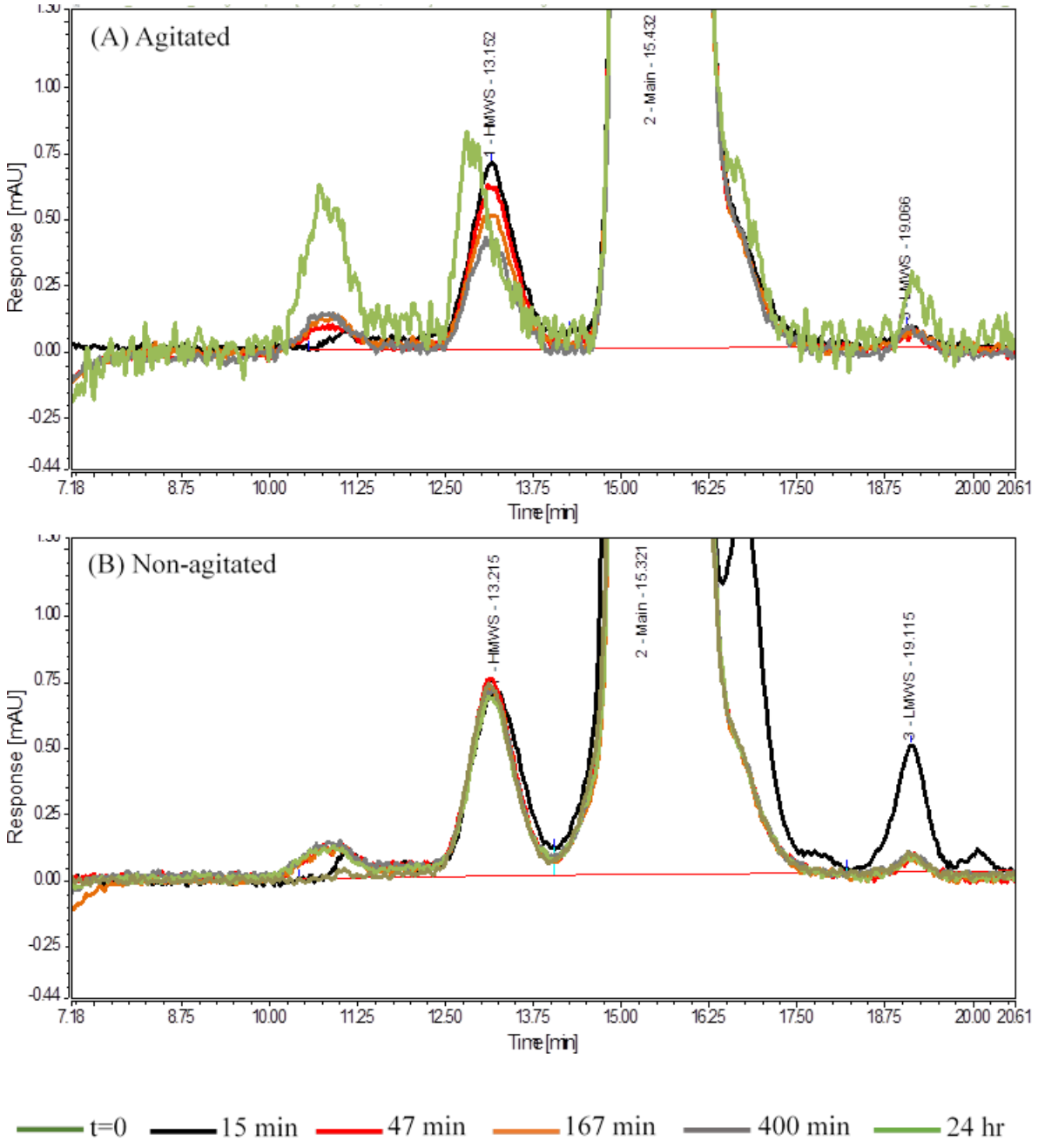


Figure 3.5.B-2: Size exclusion chromatograms of (A) agitated and (B) non-agitated mAb-α in the absence of polysorbate 20.

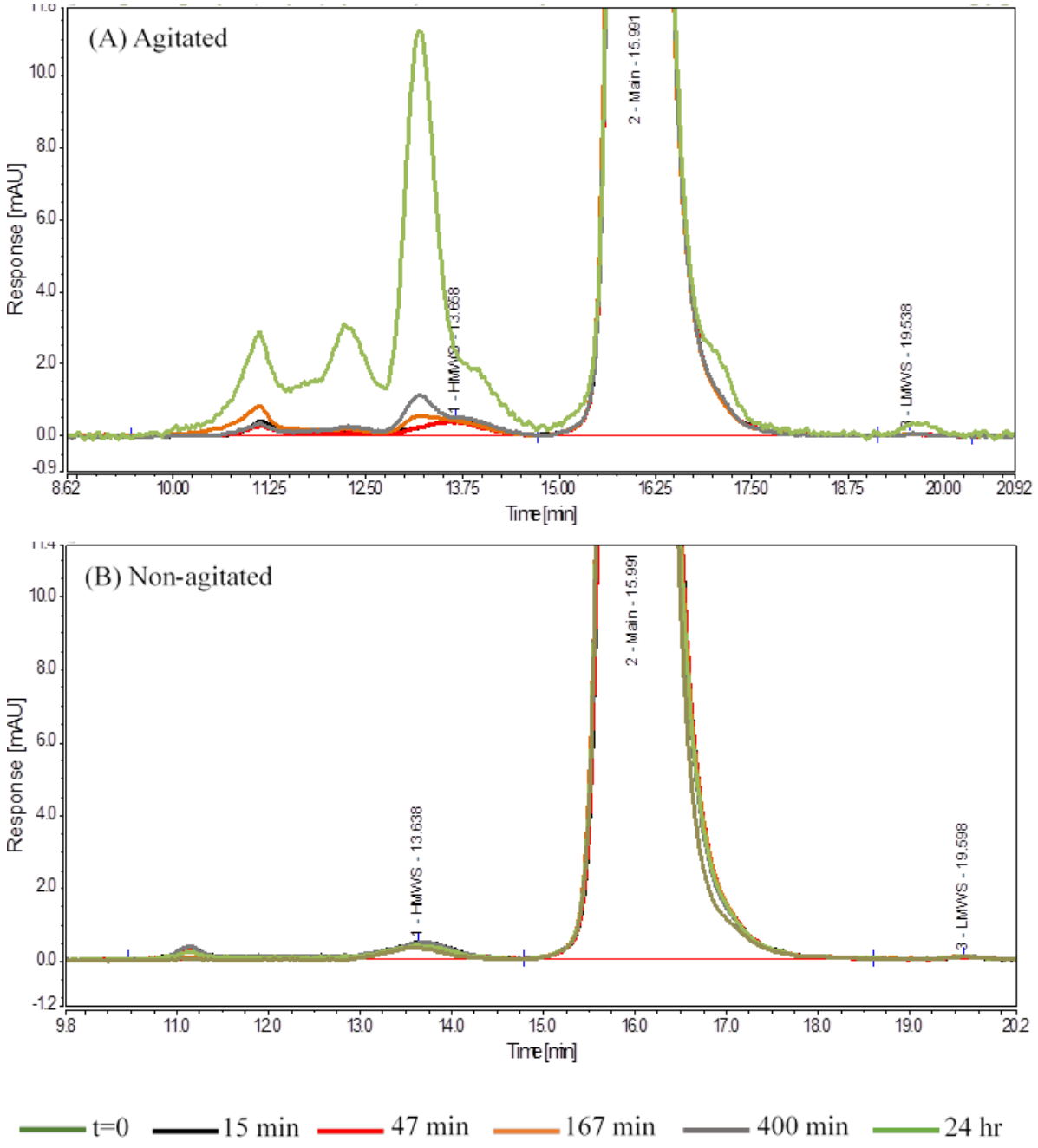


Figure 3.6.B-3: Size exclusion chromatograms of (A) agitated and (B) non-agitated mAb-β in the absence of polysorbate 20.

Table 3.1.B-1: The percentages of high molecular weight species present in each agitated and non-agitated sample of mAb-α at different time points in the absence of polysorbate 20.

Sample Name	Rel.Area% HMWS	Rel.Area% Main	Rel.Area% LMWS
t=0	0.78	99.16	0.06
Non-agitated 15 min	0.77	98.9	0.33
Non-agitated 47 min	0.67	99.28	0.05
Non-agitated 167 min	0.7	99.26	0.04
Non-agitated 400 min	0.77	99.17	0.06
Non-agitated 24 hr	0.67	99.28	0.06
Agitated 15 min	0.73	99.22	0.05
Agitated 47 min	0.57	99.38	0.05
Agitated 167 min	0.52	99.42	0.07
Agitated 400 min	0.44	99.55	0
Agitated 24 hr	1.25	98.5	0.25

Table 3.2.B-2: The percentages of high molecular weight species present in each agitated and non-agitated sample of mAb-β at different time points in the absence of polysorbate 20.

Sample Name	Rel.Area% HMWS	Rel.Area% Main	Rel.Area% LMWS
t=0	0.67	99.27	0.07
Non-agitated 15 min	1.17	98.78	0.05
Non-agitated 47 min	0.98	98.96	0.05
Non-agitated 167 min	0.6	99.35	0.05
Non-agitated 400 min	1.14	98.8	0.06
Non-agitated 24 hr	0.84	99.1	0.06
Agitated 15 min	1.26	98.68	0.05
Agitated 47 min	1.04	98.91	0.06
Agitated 167 min	1.94	98	0.06
Agitated 400 min	2.01	97.95	0.04
Agitated 24 hr	14.34	85.42	0.24

4 CHAPTER 4: Determination of relative biding affinities of V_HH-F5 mutants against ricin toxin subunit A to assess the suitability of using F5 as a model protein to define detection limits of HX-MS measurements

4.1 Introduction

HX-MS is increasingly being utilized to characterize higher-order structure of proteins in pharmaceutical discovery and development. Applications of HX-MS in higher-order protein structure analysis include studying protein conformational changes caused by various stresses, modifications, and different formulations, assessing comparability between production batches, studying protein-protein/ligand interactions. HX-MS has proven to be a reliable tool for protein higher-order structure assessment because of its potential to provide finger-print like information on higher-order protein structure.[1-5] However, the ability of HX-MS to detect subtle changes in protein higher-order structure or the detection limits for these measurements have not been rigorously tested.[3] Therefore, it is important to investigate what the smallest change that can be detected by HX-MS and if the detected changes are meaningful and could be related to functional alterations of a protein. Investigating these aspects will demonstrate the capabilities of HX-MS as an analytical tool in batch-to-batch comparability studies and comparing innovator or reference product with biosimilars for higher-order structural characterization of therapeutic proteins. The goal of this work is to generate a series of mutants of the protein V_HH-F5 with subtle structural and functional changes and test the suitability of these mutants as a panel of model proteins to define the limit of detection of HX-MS measurements in protein higher-order structural analysis.

4.1.1 Anti-ricin V_HH-F5 as a model protein

V_HHs are the variable domains of heavy chain-only antibodies. Heavy chain-only antibodies are a class of antibodies which are produced by camelids that contain only two heavy chains with molecular weight of about 95 kDa. The V_HH antibody fragment is about 15 kDa.[6, 7] V_HHs have an immunoglobin fold similar to variable heavy domain of a conventional antibody.

However, V_HHs contain modifications at the interface where variable heavy domain interacts with variable light domain of conventional antibodies. In V_HHs, hydrophobic residues at the interface are replaced with more hydrophilic amino acids.[7] V_HHs bind with their antigens with high specificity and affinity with dissociation constant values in nanomolar range.[8] Also, V_HHs have displayed higher stability relative to conventional antibodies. Moreover, V_HH antibody fragments can bind to enzyme active sites due to their smaller size, that is not reachable to conventional antibodies. [9-11]

For the purpose of our study, V_HHs serve as a good model protein because they are representative of antibodies which have become the prominent class of therapeutic proteins in the recent past. In addition, V_HHs can be easily expressed in *E. coli* and form high quality crystals.[12] The ability to form crystals is important for future work of this project because crystal structures of the V_HH selected for this study and a panel of its mutants will be solved to obtain structural data on binding of the selected V_HH and its mutants to its antigen. Moreover, our lab has participated in collaborations that studied the toxin neutralizing activity of a set of V_HHs against ricin toxin and characterized the binding of V_HHs with ricin toxin.[13, 14] Our collaborators have already developed binding assays to determine binding of the V_HHs to ricin toxin. Because of all the above reasons, we selected the V_HH called as F5, which has been identified as a ricin toxin subunit A (RTA) specific neutralizing V_HH antibody fragment, for this work. Further, we used V_HH-F5 and RTA as a model system to perform binding assays for our study.

RTA is one of the two subunits of ricin toxin which involves in inactivating eukaryotic ribosomes. There are ongoing efforts to develop ricin toxin neutralizing vaccines using V_HHs targeting RTA to understand mechanisms of neutralization.[15] RTA consists of three distinct

domains: folding domains I, II, and III. The folding domain I (residues 1-117) of RTA contains a six stranded β-sheet that terminates with a solvent exposed α-helix.[16] This α-helix (residues 97-108) is known as α-helix B and is the target binding site of F5.[12, 17] F5 binds with RTA with high affinity and has a dissociation constant (K_d) value of 2.2 nM.[13] The X-ray crystal structure of F5 in complex with RTA has been reported by Rudolph et.al[17] and is shown in **Figure 4.1**.

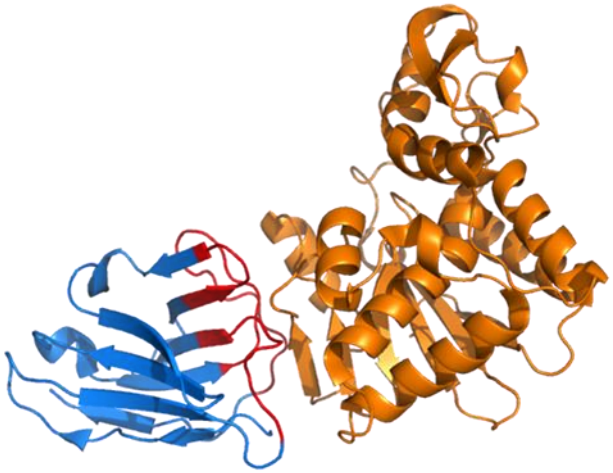


Figure 4.1: X-ray crystal structure of F5 (blue) in complex with RTA (orange). CDRs (complementarity-determining regions) of F5 are shown in red. PDB 4Z9K.

4.1.2 V_HH-F5 mutants

A series of point mutations were used with the aim of introducing subtle changes to the F5 structure and detune its binding to RTA and thereby generate a panel of proteins having small and progressive higher-order structural perturbations and detuned binding. In future work, this panel of proteins will be used to assess the ability of HX-MS to detect the introduced small changes in protein higher-order structure, thus limits of the detection. Further, this panel of proteins will be

used to investigate whether HX-MS would be capable of detecting all structural changes that are related to alterations in binding.

V_HHs poses a classical immunoglobin fold with nine β-strands arranged in two β-sheets and the CDRs (complementarity-determining regions) 1-3 are located on one face of the structure as shown in **Figure 4.1**. F5 contains two 3_{10} helices and two intramolecular disulfide bonds. One disulfide bond is located at C22 and C92 linking FR1(framework region 1) and FR2. The other disulfide bond is located at C50 and C97 linking CDR2 and CDR3. All three CDRs of F5 contribute to making binding interactions with RTA. F5 forms binding interactions with RTA at α-helix B (residues 98, 101-104, and 105), several residues immediately before α-helix B (residues 91, 92, and 94-97), β-strand h (residues 113-117), several residues in the loop (residues 111, 112, 118, and 119) before and after β-strand h, C-terminal region of α-helix D (residues 154), the loop immediately after α-helix D (residues 157 and 161) and R125 in α-helix C.[17]

The list of mutations is shown in **Table 4.1**. All the mutations were located either distal or proximal to the CDRs or the paratope of F5, so that the mutations are expected to cause only subtle changes in the structure of F5 and subtly detune the binding. No mutations were located in the CDRs because mutations in the CDRs would prevent or drastically change the binding of F5 to RTA. The targeted locations for the mutations were distal and proximal surface-exposed hydrophobic clefts, distal and proximal salt bridges, and distal reverse turns. As the goal of the mutations is to introduce subtle changes in the structure and binding, the mutations were designed to be either conservative or semi-conservative.

Table 4.1: Mutations in F5 and their locations that were used to introduce subtle changes to its structure and binding to RTA.

Location of mutations in F5	Mutation
Proximal hydrophobic cleft	I34L, I34V, V51A, V51I, I68L, I68V, V75A, V75I
Proximal salt bridge	E46D, K34R
Distal hydrophobic cleft	V12I, V12A, L18I, L18V, L82I, L82V, V88A, V88I
Distal salt bridge	D86E, R65K
Distal reverse turns	P14A, P41A, P84A
Disulfide bond	C50AC97A

Among the twenty-four mutations selected in the full project, five mutations were selected for this work as listed in **Table 4.2**. The mutations V51A and L18V are hypothesized to disturb the side-chain packing in the hydrophobic clefts. The salt bridge structure is predicted to be disrupted by mutating glutamate to aspartate (E46D) which shortens the side chain length by one carbon and arginine to lysine (R65K) which interrupts and weakens the electrostatic interactions. Proline to alanine mutation (P41A) at the reverse turn is expected to distort the reverse turns by changing the backbone dihedral angles. **Figure 4.2** shows each mutation site on the F5 structure.

Table 4.2: The set of mutations that were selected for this work and their locations on the F5 structure.

Location of the mutations in F5	Mutation
Proximal hydrophobic cleft	V51A
Proximal salt bridge	E46D
Distal hydrophobic cleft	L18V
Distal salt bridge	R65K
Distal reverse turns	P41A

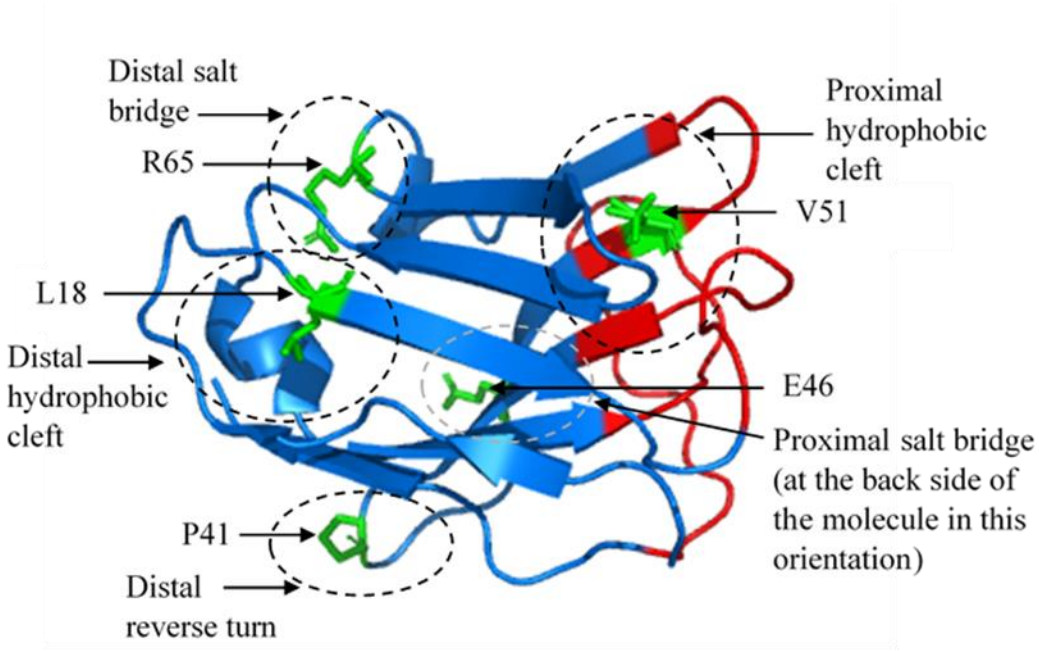


Figure 4.2: Each mutation site with the wild-type residue side chain is shown in green on the F5 structure (blue). The paratope of F5 is shown in red. The dotted circles show the targeted location of each mutation.

4.2 Experimental and results

4.2.1 SUMO-F5 plasmid

SUMO-F5 fusion protein DNA was prepared by GenScript (Piscataway, NJ) as an insert into the commercial pET-15b plasmid. The expression plasmid was mainly composed of the gene for F5, TEV (tobacco etch virus) protease recognition and cleavage site, SUMO (small ubiquitin-like modifier) tag, thrombin recognition and cleavage site, and 6xHis affinity tag. **Figure 4.3** shows a schematic diagram of main components of SUMO-F5 protein construct in panel (a) and the corresponding amino acid sequence in panel (b).

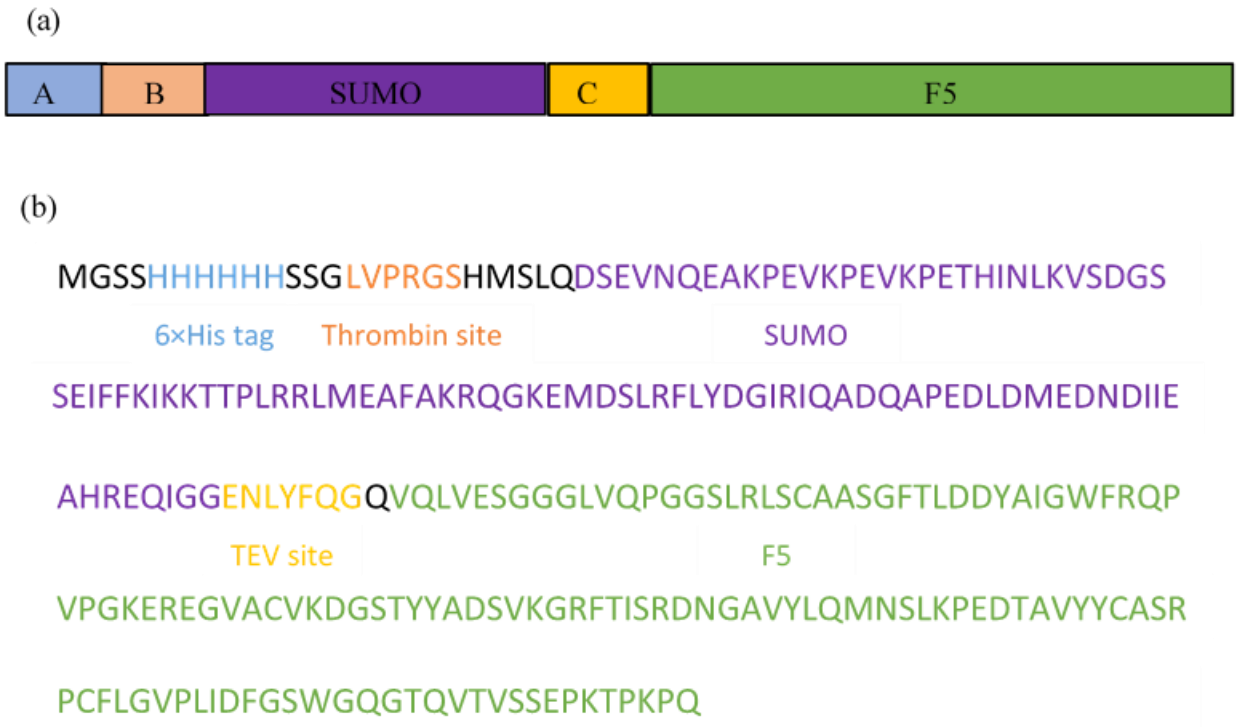


Figure 4.3: (a) Schematic diagram showing the main components of SUMO-F5 fusion protein construct. A, B, and C represents 6xHis affinity tag, thrombin recognition and cleavage site, and TEV protease recognition and cleavage site, respectively. (b) Amino acid sequence of SUMO-F5 protein construct.

4.2.2 Transformation of competent cells

SUMO-F5 plasmids were received in lyophilized form (4 µg) and stored at -20 °C. The plasmid DNA was prepared for transformation by reconstituting in sterile water. First, the plasmid DNA containing vial was thawed at room temperature and centrifuged at 6,000 × g for 1 min at 4 °C. Then, 20 µL of sterile water was added and vortexed for 1 min. The unused plasmid DNA was re-stored directly at -20 °C. Transformation of cells was performed using subcloning efficiency DH5α competent cells (Invitrogen) for preparing glycerol stocks for long-term plasmid storage and Origami 2 (DE3) singles competent cells (Novagen) for preparing glycerol stocks to use in the expression and purification of F5.

For the transformation of DH5α cells, which were stored at -70 °C, the cells were first thawed in an ice/water bath. Then the cells were gently mixed by swirling using a pipette tip and aliquoted 50 µL into a microcentrifuge tube kept on ice. The unused cells were refrozen in a dry ice/ethanol bath for 5 min and stored at -70 °C. F5 plasmid DNA (2.5 µL) was added into the cells and mixed gently. The resulting mixture was incubated in an ice/water bath for 30 min. After that, the cells were heat shocked for 20 sec in a 42 °C water bath and then again incubated in an ice/water bath for 2 min. Next, 950 µL of pre-warmed (at 37°C) LB (Luria-Bertani) broth was added to the mixture and incubated at 37°C for 1 hr shaking at 225 rpm. LB was prepared by dissolving 25 g of granulated LB (Miller granulated LB, Fisher Bioreagents) in 1 L of ultrapure water and adjusting the pH to 7.0. The transformed cells were selected by plating the transformation mixture on LB agar medium containing ampicillin because the pET-15b plasmid contains an ampicillin resistance gene. The final transformation reaction mixture (100 µL) was spread on a pre-warmed (at 37 °C) LB agar plate (Teknova), containing 100 µg/mL ampicillin. The plates were incubated overnight

at 37 °C. Glycerol stocks were prepared for storing the transformed cells. A single colony of the cells grown on the LB agar plate was inoculated into 5 mL of LB medium containing 100 µg/mL of ampicillin in a sterile tube. The LB medium was incubated overnight at 37 °C shaking at 175 rpm. A fraction (500 µL) of the grown culture was added to 500 µL of sterilized 20% glycerol in a cryovial and stored at -70 °C.

For the transformation of Origami 2 (DE3) singles competent cells, the cells were thawed in an ice/water bath and mixed gently by finger flicking the tubes a few times. An aliquot of the thawed cells (50 µL) was added into a chilled micro centrifuge tube and 5 µL of SUMO-F5 plasmid DNA was also added into the tube. The tube was incubated on ice for 5 min. Then, the tubes were heated in a 42 °C water bath for 30 sec and placed on ice again for 2 min. Room temperature SOC medium (250 µL) was added into the tube and incubated at 37 °C while shaking at 250 rpm for 1 hr. The total volume of the transformation mixture was spread on a pre-heated (at 37 °C) LB agar plate (Teknova) containing 100 µg/mL ampicillin and incubated overnight at 37 °C. A single colony from the plate was inoculated into 5 mL of pre-heated LB (at 37 °C) containing 100 µg/mL of ampicillin and incubated overnight at 37 °C shaking at 175 rpm. The glycerol stocks were prepared by adding 500 µL of the grown culture into 500 µL of sterilized 20% glycerol and stored at -70 °C.

4.2.3 Expression of SUMO-F5

The glycerol stocks containing transformed Origami 2 (DE3) host cells were used for SUMO-F5 expression. Origami 2 host strain is designed to enhance disulfide bond formation in the *E. coli* cytoplasm for the expression of proteins that require disulfide bond formation for proper folding. As SUMO-F5 contains two disulfide bonds, using Origami 2 cells for expression can

facilitate the proper disulfide bond formation. In addition, Origami 2 (DE3) strains are suitable for production of proteins from pET vectors cloned with target genes by IPTG (isopropyl β-D-1-thiogalactopyranoside) induction. The pET-15b plasmid contains the *lac* operon which can be induced by IPTG for the expression of targeted protein. Therefore, using Origami 2 cells with the pET-15b plasmid allowed the expression of SUMO-F5 induced by IPTG.

The glycerol stock was scraped using a sterile culture loop and streaked on a pre-warmed (at 37 °C) LB agar plate (Teknova) containing 100 µg/mL ampicillin. The plate was incubated overnight at 37 °C. The plate was sealed using parafilm and stored at 4 °C until use on the same day. A starter culture was prepared by inoculating a single colony from the streaked plate into 50 mL of LB containing 100 µg/mL ampicillin in a 250 mL Fernbach culture flask. The flask was incubated overnight at 37 °C shaking at 175 rpm. The entire starter culture was added to 1 L of LB containing 100 µg/mL ampicillin in a 4 L Erlenmeyer flask to scale up the culture size. The large culture was incubated at 37 °C shaking at 250 rpm until OD_{600} (optical density at 600 nm) reached between 0.6-0.8. At this point, IPTG was added as a solid and mixed well to have a concentration of 1 mM to induce the expression of SUMO-F5 gene. After adding IPTG, the culture was further incubated for 4 hr at 37 °C with shaking at 250 rpm. The cells were harvested from the culture by centrifuging at 5,000 × g for 10 min at 4 °C. The supernatant was removed and the cell pellet was washed by resuspending the pellet in a small volume of wash buffer (20 mM sodium phosphate, 300 mM sodium chloride, 10 mM imidazole, pH 7.4). The cell suspension was centrifuged at 5,000 × g for 10 min at 4 °C. The supernatant was discarded and the weight of the pellet was recorded. The weight of the cell pellet from 1 L of culture was varied between 1-2 g among the F5 mutants. The cell pellet was stored at -70 °C until use.

SUMO-F5 expression levels during the induction were monitored by SDS-PAGE. Fractions (0.5 mL) of the grown culture were removed into microcentrifuge tubes immediately after adding IPTG, after 2 hr and 4 hr after adding IPTG. The culture fractions were centrifuged at 5,000 × g for 10 min at 4 °C. The resulting cell pellet was drained and resuspended in 10 µL of water and 10 µL of 2X Laemmli sample buffer (Sigma). The cell suspension was heated at 92 °C for 10 min and centrifuged at 10,000 × g for 10 mins. The supernatant (5 µL) was loaded onto 4-12% bis-tris mini protein gel (Invitrogen) along with a protein ladder (Thermo Scientific). The gel was run in MES SDS running buffer (Invitrogen) at a 1X concentration for 60 min at 120 V. The gel was stained for imaging using a Coomassie based staining solution (InstantBlue, Expedeon) where the gel was shaken in the staining solution for 15 min. Imaging was performed after washing the gel in ultrapure water overnight. The gel image displaying the expression levels of SUMO-F5 (wild type) over the 4 hr induction period is shown in **Figure 4.4**. Expression of SUMO-F5 was indicated by the most intense bands at 30 kDa molecular weight level in the LB fractions after 2 hr (lane 3) and 4 hr (lane 4) of induction. The absence of the SUMO-F5 band in the fraction collected immediately after adding IPTG (lane 2) and the appearance of the SUMO-F5 band after induction for 2 to 4 hr indicate that IPTG had induced the protein expression with time.

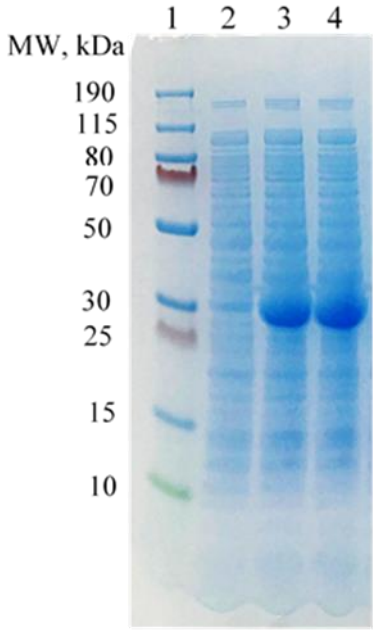


Figure 4.4: SDS-PAGE gel image showing the expression of SUMO-F5 over the 4 hr IPTG induction period. Lanes 1, 2, 3, and 4 contained protein ladder, LB grown culture immediately after adding IPTG, 2 hr after adding IPTG, and 4 hr after adding IPTG, respectively.

4.2.4 Purification of SUMO-F5

The cell pellet was thawed at room temperature and resuspended completely in 10 mL equilibration buffer (20 mM sodium phosphate, 300 mM sodium chloride, 10 mM imidazole, pH 7.4). Protease inhibitor cocktail (for use in purification of histidine-tagged proteins, Sigma-Aldrich, P8849) was added with a ratio of 1 mL for 20 g of cell pellet to inhibit the activity of proteases present in the cell extract. The cocktail contained five protease inhibitors; 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride, bestatin hydrochloride, N-(trans-epoxysuccinyl)-L-leucin 4- guanidinobutylamide, pepstatin A, and phosphoramidon disodium salt that have specificity for serine, cysteine, acid, thermolysin-like proteases and aminopeptidases. The cell suspension was sonicated to lyse the cells. The cell suspension containing tube was placed

on ice and sonicated using ultrasonic cell disruptor (Microson) in 10 sec bursts for 5-6 times at 50% power level. The cell suspension was allowed to cool down in between the sonication bursts for 1-2 min. The resulting mixture of cellular debris and cell lysate was centrifuged at 10,000 rpm for 20 min at 4 °C. The supernatant containing the cell lysate was transferred to a fresh tube for purification.

Purification of SUMO-F5 was performed using the HisPur Ni-NTA spin columns (Thermo Scientific) as the expressed protein contains a 6×His-tag. The spin column (3 mL resin-bed), placed in a centrifuge tube, was centrifuged at 700 × g for 2 min at 4 °C to remove the storage buffer and equilibrated with two resin-bed volumes of equilibration buffer (20 mM sodium phosphate, 300 mM sodium chloride, 10 mM imidazole, pH 7.4). The column was centrifuged again at 700 × g for 2 min at 4 °C to remove the buffer. The cell lysate was added to the spin column and mixed on an end-over-end mixer for 30 min at 4 °C. The column was centrifuged at 700 × g for 2 min at 4 °C and the flowthrough was collected in a fresh centrifuge tube. The resin was washed with two resin-bed volumes of wash buffer (20 mM sodium phosphate, 300 mM sodium chloride, 50 mM imidazole, pH 7.4) and centrifuged at 700 × g for 2 min at 4 °C. Washing was repeated for a total of five times and after each wash, the flow-through was collected in a separate centrifuge tube. Optimal removal of non-specifically bound proteins was achieved while minimizing the loss of specifically bound histidine-tagged protein by having an imidazole concentration of 50 mM in the wash buffer and performing five washing steps. To elute the His-tagged SUMO-F5 protein from the resin, one resin-bed volume of elution buffer (20 mM sodium phosphate, 300 mM sodium chloride, 500 mM imidazole, pH 7.4) was added and centrifuged at 700 × g for 2 min at 4 °C. Each eluted fraction was collected in a separate tube. The optimum imidazole concentration needed for maximum elution of His-tagged SUMO-F5 protein was

determined as 500 mM. The concentration of protein in each eluted fraction was determined by measuring the absorbance using a Nanodrop spectrophotometer (NanoDrop Lite, Thermo Scientific, Waltham, MA). NIST mAb reference material with a known concentration was used to perform calibration checks on the NanoDrop spectrophotometer. The elution step was repeated a total of eight times. The yield of protein after eight elution steps varied between 5-10 mg among the SUMO-F5 mutants. The progress of the purification process was also observed by running an SDS-PAGE gel. A fraction of 10 µL from the cell lysate, Ni spin column flowthrough, wash flowthroughs, and all the elutions were collected. Each fraction was mixed with 10 µL of Laemmli buffer and heated at 92 °C for 10 min. 10 µL volume of each of the denatured fractions were loaded onto a 4-12% bis-tris mini protein gel along with the protein ladder. The gel was run in MES SDS running buffer at a 1X concentration for 60 min at 120 V. The gel was stained for imaging using a Coomassie based staining solution (InstantBlue, Expedeon). The gel image displaying the progress of the purification of SUMO-F5 is shown in **Figure 4.5**. The gel band at 30 kDa indicates the presence of SUMO-F5. The intense band at 30 kDa in the cell lysate indicates the presence of extracted SUMO-F5 from the cells. The decrease in the intensities of other protein bands with the washing steps (lanes 4-6) demonstrate that proteins that were nonspecifically bound to the Ni-column were washed away gradually. Most of SUMO-F5 protein was eluted in the first few elutions (lanes 7-9) and decrease with the number of elutions (lanes 11-12).

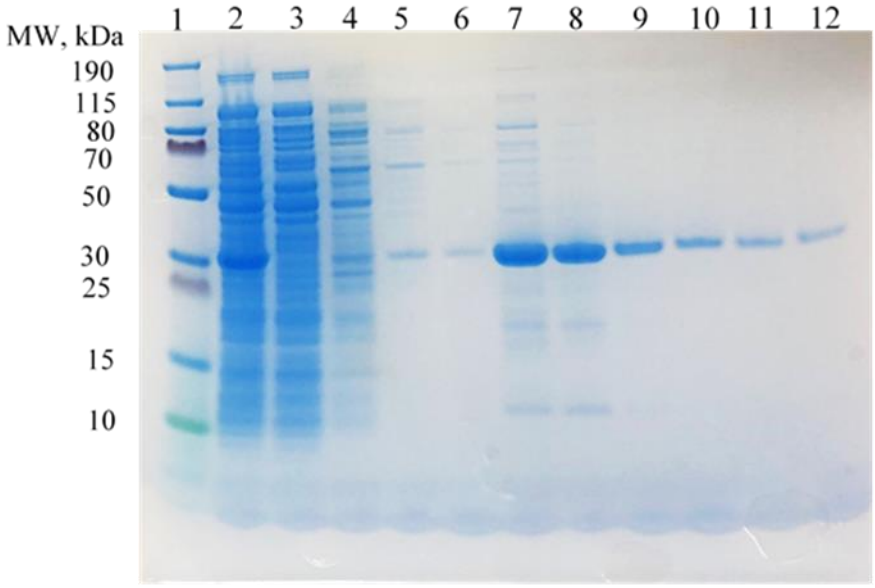
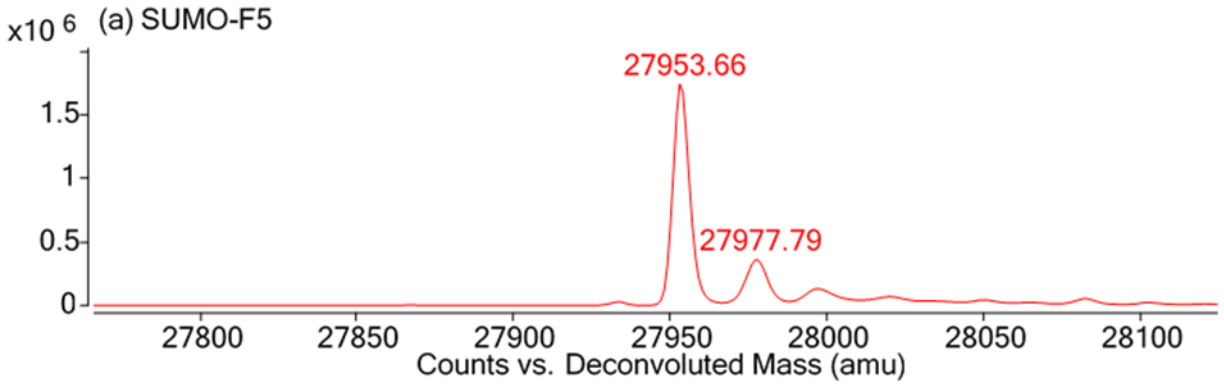


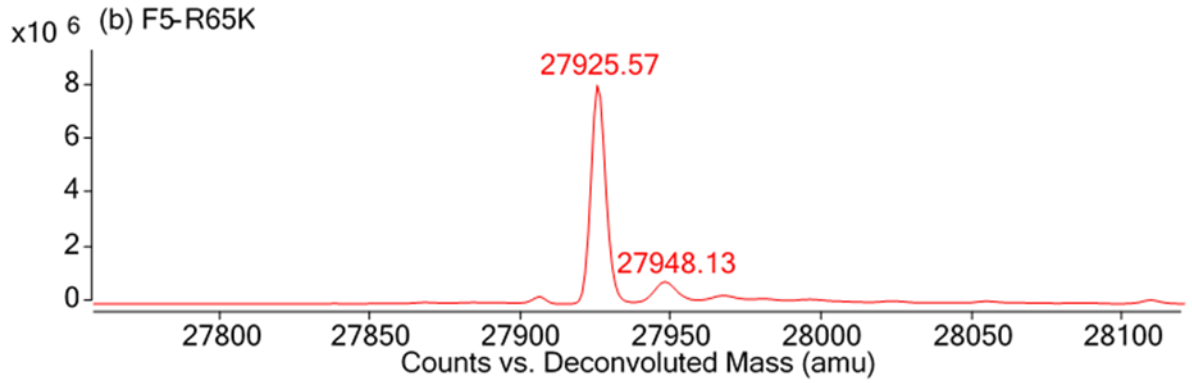
Figure 4.5: SDS-PAGE gel image showing the progress of SUMO-F5 purification using Ni-NTA spin columns. Lanes: 1-protein ladder, 2-cell lysate, 3-flowthrough, 4-wash 1, 5-wash 3, 6-wash 5, 7-elution 1, 8-elution 3, 9-elution 5, 10-elution 6, 11-elution 7, 12-elution 8.

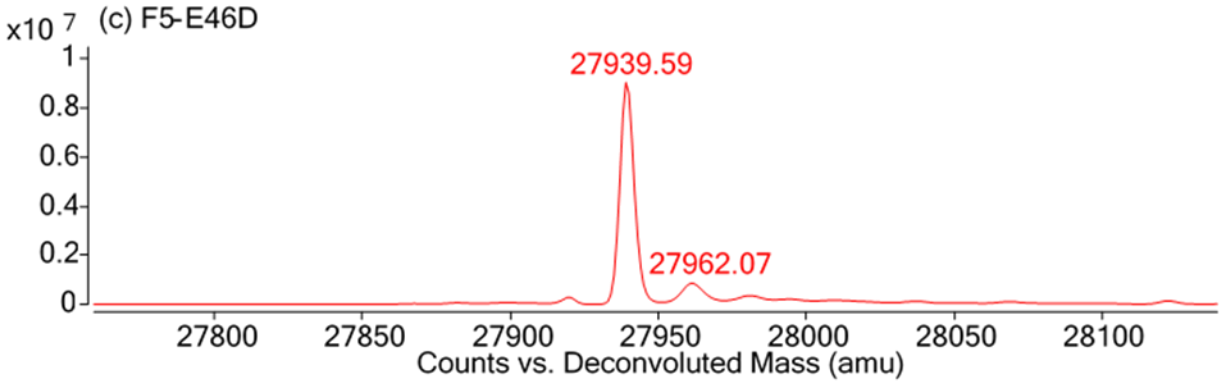
The eluted fractions containing the protein were combined and buffer exchanged to tris buffer (20 mM Tris-HCl, 100 mM NaCl, pH 7.5) using 10K, 15 mL centrifugal filter units (Millipore Sigma) by centrifuging at 5000 × g for 15 min at 4 °C. The buffer exchanged protein was flash frozen in liquid nitrogen and stored at -80 °C until further use. To regenerate the resin, residual imidazole and any nonspecifically adsorbed protein to the Ni-spin column was removed by washing the resin with ten resin-bed volumes of MES buffer (20 mM 2-(N-morpholine)-ethanesulfonic acid, 0.1 M sodium chloride, pH 5.0) and ten resin-bed volumes of ultrapure water. The resin was stored as a 50% slurry in 20% ethanol at 4 °C.

4.2.5 Intact protein analysis of SUMO-F5 by LC-MS

Intact protein MS analysis was performed to confirm the mass of the expressed and purified SUMO-F5. A liquid chromatography (LC) system (Agilent 1260, Santa Clara, CA) coupled with a quadrupole time-of-flight (QTOF) mass analyzer (Agilent 6530, Santa Clara, CA) was used for the intact protein analysis. Mobile phase A and B consisted of 0.1% formic acid in water and 0.1% formic acid in acetonitrile. For desalting, the purified protein sample was passed through a reversed phase C3 trap (2.1 x 12.5mm, 5-micron, ZORBAX 300SB) using the gradient; 1-20% B in 2 min, 20-60% B from 2-7 min, 60-95% from 7-10 min, 95% from 10-11 min and 95-1% from 11-13 min, at a flow rate of 200 µL/min. Intact protein mass spectra were analyzed using MassHunter Qualitative Analysis (version B.07.00) software. Deconvoluted intact protein mass spectra of SUMO-F5 proteins that were expressed and purified are shown in **Figure 4.6**. The deconvoluted masses were compared with the theoretical sequence masses of the SUMO-F5 proteins to confirm that the expression and purification procedure had yielded the expected mutants. The sequence mass also included the mass of the 6×His tag, thrombin recognition site, SUMO tag, and TEV protease recognition site. In addition, the theoretical sequence masses were calculated for the oxidized form of the proteins to incorporate the two disulfide bonds which were expected to present in SUMO-F5. **Table 4.3** lists the theoretical sequence masses and the deconvoluted masses for each SUMO-F5 protein and the mass differences calculated between the two mass values. The mass differences were calculated by subtracting the theoretical sequence mass from the deconvoluted mass for the oxidized form of the proteins containing four oxidized cysteines forming two disulfide bonds. The mass difference between theoretical sequence mass and deconvoluted mass for each SUMO-F5 protein was less than 1 Da indicating that the proteins were expressed and purified with the expected mutations and two intact disulfide bonds.


x10 6
(a) SUMO-F5
27953.66
27977.79
1.5
1
0.5
0
27800
27850
27900
27950
28000
28050
28100
Counts vs. Deconvoluted Mass (amu)


x10 6
(b) F5-R65K
27925.57
27948.13
8
6
4
2
0
27800
27850
27900
27950
28000
28050
28100
Counts vs. Deconvoluted Mass (amu)


x10 7
(c) F5-E46D
27939.59
27962.07
1
0.8
0.6
0.4
0.2
0
27800
27850
27900
27950
28000
28050
28100
Counts vs. Deconvoluted Mass (amu)

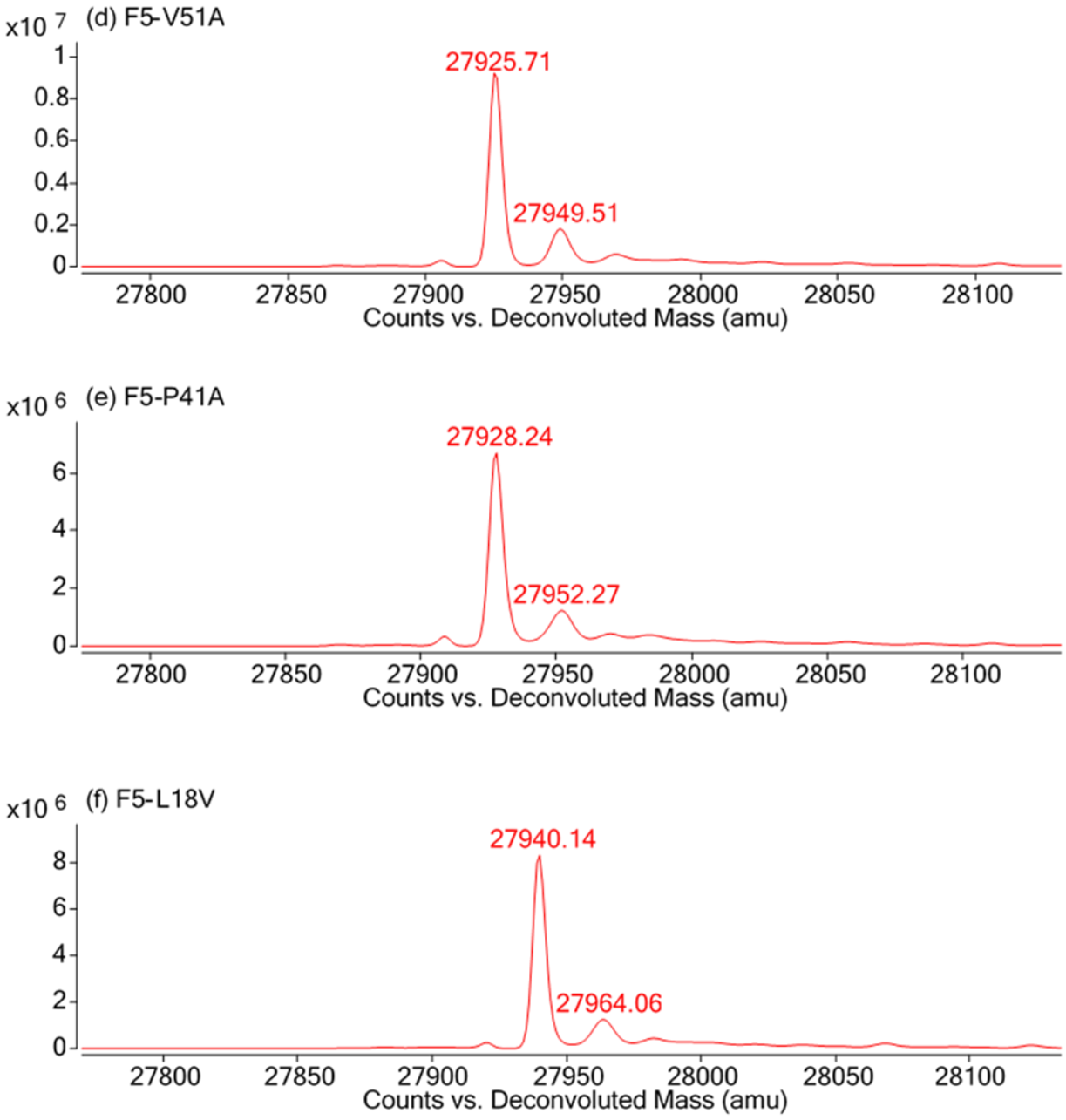


Figure 4.6: Deconvoluted intact protein mass spectra of the expressed and purified SUMO-F5 proteins: (a) SUMO-F5, (b) F5-R65K, (c) F5-E46D, (d) F5-V51A, (e) F5-P41A, and (f) F5-L18V. The satellite peak observed at +24 Da in each spectrum may be resulting from a sodium adduct.

Table 4.3: Theoretical sequence masses and deconvoluted masses of each SUMO-F5 protein. The mass difference was calculated by subtracting theoretical sequence mass from the deconvoluted mass.

F5 protein	Sequence mass (Da)	Deconvoluted mass (Da)	Mass difference (Da)
F5_WT	27953.50	27953.66	0.16
F5_R65K	27925.25	27925.57	0.32
F5_E46D	27939.24	27939.59	0.35
F5_V51A	27925.21	27925.71	0.50
F5_P41A	27927.46	27928.24	0.78
F5_L18V	27939.47	27940.14	0.67

4.2.6 Removal of His-SUMO tag by TEV cleavage reaction

The expressed and purified SUMO_F5 contained a 6×His-tag and a SUMO fusion protein tag. These two tags need to be cleaved from the F5 protein structure to perform HX-MS analysis of the F5 mutants. The F5 protein also contained a TEV protease recognition and cleavage site after His-SUMO tags which can be used remove the tags from F5 proteins (see **Figure 4.3**). Therefore, His-SUMO tag was removed using His-tagged TEV protease, which was produced by COBRE protein production group at the University of Kansas. Reaction with TEV protease will cleave SUMO-F5 at the TEV protease recognition site. Cleaved F5 and His-SUMO tags along with His-tagged TEV protease can be separated by Ni affinity purification. The His-SUMO fragment and His-tagged TEV protease will bind to the Ni column while F5 protein will elute with the flowthrough. The TEV protease storage buffer contained 25 mM tris-HCl, 0.5 mM EDTA, 1 mM DTT, 0.25 M NaCl and 50% glycerol at pH 8.0. The presence of DTT is not favorable because

F5 contains two disulfide bonds and DTT can reduce disulfide bonds. The presence of EDTA in the reaction mixture can adversely affect the Ni affinity column performance when separating the cleaved protein by Ni affinity separation. However, DTT is used in standard TEV cleavage reactions to provide the required reducing power for TEV protease to work. Therefore, the TEV protease was buffer exchanged to a reaction buffer containing 25 mM tris-HCl, 100 mM NaCl, 3 mM glutathione, and 0.3 mM oxidized glutathione at pH 8.0, which can maintain disulfide bonds while providing redox power for TEV reaction.[18] Buffer exchange was performed using 3K, 4 mL molecular weight cutoff filter centrifugal filter units (Millipore Sigma) by centrifuging at 7000 × g for 30 min at 4 °C. Buffer exchanged TEV protease and SUMO-F5 was reacted at a ratio of 1:2 for 3 hr at RT and then overnight at 4 °C. The reaction mixture was separated using a HisTrap HP Ni affinity column (5 mL, GE Healthcare).

First, the Ni affinity column was equilibrated with five column volumes of the equilibration buffer (20 mM sodium phosphate, 300 mM sodium chloride, 20 mM imidazole, pH 7.4). Next, the TEV protease/F5 reaction mixture was passed through the column. The flowthrough was collected. Then, 5 mL portions of equilibration buffer were passed through the Ni column two times and the flowthroughs were collected. Cleaved F5 eluted with these flowthrough fractions and the concentrations of protein in these fractions were measured using a Nanodrop spectrophotometer (NanoDrop Lite, Thermo Scientific, Waltham, MA). The His-SUMO tag and His-tagged TEV protease bound to the Ni column was eluted by passing the elution buffer (20 mM sodium phosphate, 300 mM sodium chloride, 500 mM imidazole, pH 7.4). The Ni-column was re-equilibrated with ultrapure water. The progress of the TEV cleavage reaction and separation of the cleaved F5 by Ni column was observed by running an SDS-PAGE gel. **Figure 4.7** shows the SDS-PAGE gel displaying the progress of the TEV cleavage reaction for SUMO-F5 and the separation

of cleaved F5 (lane 1- protein ladder, lane 3-SUMO-F5, lane 4-His-tagged TEV protease, lane 5-TEV/F5 reaction mixture after overnight incubation, lane 6-cleaved F5 separated using Ni affinity column). The bands on lane 5 of the gel shows that the TEV/F5 reaction mixture after overnight incubation contains cleaved F5, cleaved His-SUMO tag, His-tagged TEV protease and some fraction of un-cleaved SUMO-F5. The protein band on lane 6 confirms that the Ni column separation of the reaction mixture results in the separation of cleaved F5 in the flowthrough with a good purity as there were no other bands observed on lane 6. Moreover, lane 3 and 5 shows a band at around 12 kDa which may be an impurity that might have not been removed in the protein purification step. The flowthrough fraction containing the cleaved F5 was buffer exchanged to PBS buffer (20 mM sodium phosphate, 300 mM sodium chloride, pH 7.4). The buffer exchange was carried out using 3K, 15 mL molecular weight cutoff centrifugal filter units (Millipore Sigma) by centrifuging at 5000 × g for 30 min at 4 °C. The final yield of F5 obtained after cleaving 1 mg of SUMO-F5 followed by separation and buffer exchange steps was 0.4 mg, yielding 42% of the final product. Cleaved F5 in PBS buffer was flash frozen in liquid nitrogen and stored at -80 °C.

Intact protein analysis of cleaved F5 was performed to confirm the mass of the cleaved protein as described in section 4.2.5 and the deconvoluted mass spectrum is shown in **Figure 4.8**. Theoretical sequence mass of cleaved F5 containing four oxidized cysteines is 13671.46. The deconvoluted mass determined by intact protein analysis was 13671.62 which had only a 0.16 Da mass difference from the theoretical sequence mass value. This indicates that TEV cleavage of SUMO-F5 occurred at the expected cleavage site resulting the expected F5 protein product with the two disulfide bonds intact.

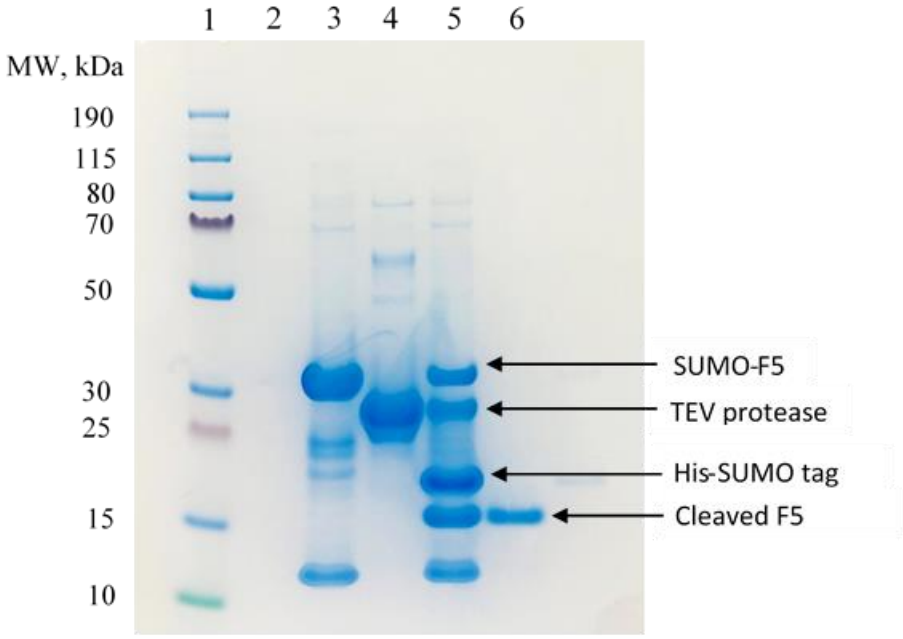


Figure 4.7: SDS-PAGE gel displaying the progress of the TEV cleavage reaction for SUMO-F5. Lanes: 1-protein ladder, 2-empty, 3-SUMO-F5, 4-His-tagged TEV protease, 5-TEV/F5 reaction mixture after overnight incubation, 6-cleaved F5 separated using Ni affinity column.

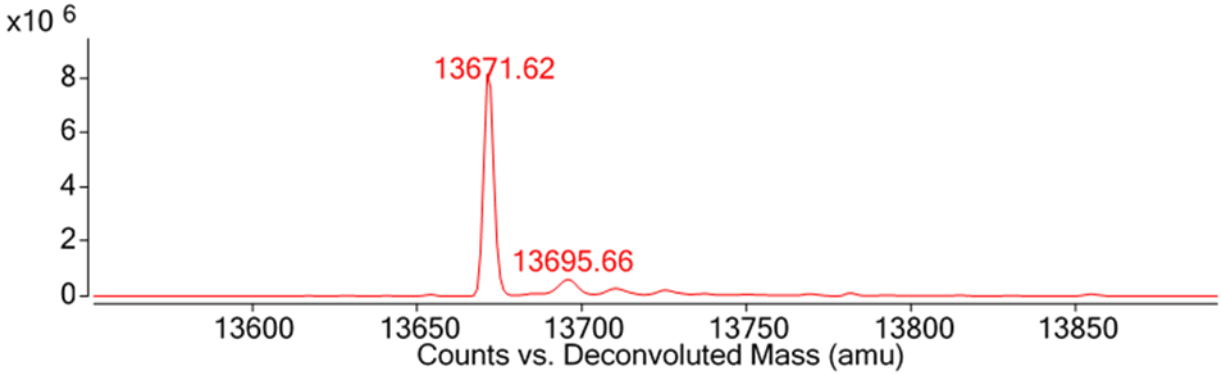


Figure 4.8: Deconvoluted intact mass spectrum of cleaved F5. Theoretical sequence mass for cleaved F5 in the oxidized form was 13671.46. The satellite peak observed at +24 Da may be resulting from a sodium adduct.

4.2.7 ELISA for determining relative binding affinities of V_HH-F5 mutants to RTA

Relative binding affinities of F5 wild type and its mutants to RTA were determined using ELISA (enzyme-linked immunosorbent assay). RTA used in this work was a mutant form of native RTA and was inactive in terms of enzymatic activity.[19] RTA was expressed in *E. coli* and purified in the lab as described previously.[20] F5 proteins used for the ELISA contained the N-terminal 6×His-tag and SUMO tag. The initial concentration of stock RTA and F5 protein samples used for the ELISA was determined using a NanoDrop UV spectrophotometer (NanoDrop Lite, Thermo Scientific, Waltham, MA). ELISA was performed as follows. A 96-well plate (Immulon 4 HBX, clear flat-bottom nonsterile, Thermo Fisher) was coated with 100 µL of 2 µg/mL RTA prepared in PBS buffer (phosphate buffered saline, pH 7.4, Thermo Fisher). The RTA coated plate was incubated overnight at 4 °C and the RTA coating solution was removed. The plate was washed three times by filling the wells with wash buffer (150 mM sodium chloride, 20 mM sodium phosphate, 0.1% Tween 20, pH 7.4). Blocking buffer (200 µL) containing 2% goat serum (Gibco) in the wash buffer was added into the wells and incubated at room temperature for 2 hr. The blocking buffer was removed. F5 and the mutants were prepared at a concentration of 1 µg/mL by diluting the protein stock solutions with blocking buffer. The F5 protein samples (200 µL) were applied only into the wells of the first row of the plate. The rest of the wells starting from second row to the eighth row were filled with 100 µL of blocking buffer. Dilutions of the protein samples at 1:2 were performed down the plate using serial dilution as follows. First, 100 µL of the protein samples from the first row was aspirated using a microchannel pipette, dispensed into the second row of wells, and mixed by pipetting up and down for three times. Then, 100 µL from the second row was aspirated and dispensed into the third row. Similarly, 1:2 dilutions were performed all the way down to the eighth row. Triplicates in columns of the plate were performed for F5 and each

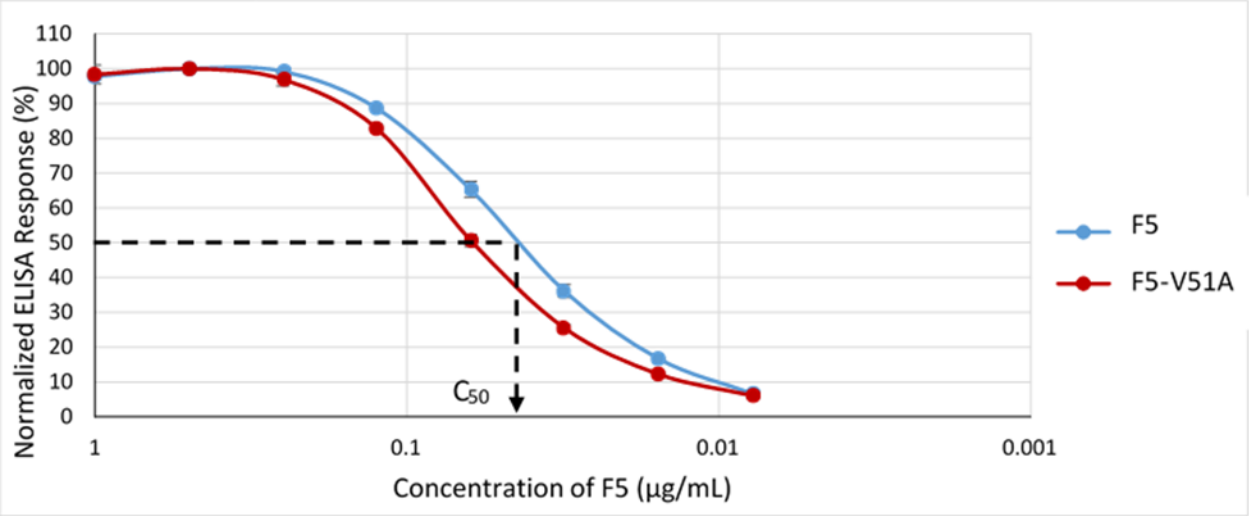
F5 mutant. Blanks were performed by adding blocking buffer instead of a protein sample. The plate was incubated for 1 hr at room temperature. Then, the F5 samples were removed and the plate was washed three times with the wash buffer. Anti-His-HRP, anti 6×His-tag monoclonal antibody-HRP (horseradish peroxidase) conjugate (Invitrogen, MA1-80218) was used as the detection antibody for the F5 proteins as they contained a His tag. Anti-His-HRP (100 µL) at 2 µg/mL concentration was added into the wells and incubated for 1 hr at room temperature. The anti-His-HRP solution was removed and the plate was washed three times with the wash buffer. SureBlue TMB (3,3',5,5'-tetramethylbenzinidine) peroxidase (SeraCare) was used as a substrate for HRP and the reaction between the two produces a blue color. TMB solution (100 µL) at room temperature was added into the wells and incubated for 6 min at room temperature. A blue color was observed in the solution. The reaction was quenched by adding 100 µL of 1 M phosphoric acid, which resulted in a yellow color. The absorbance of each well was measured at 450 nm using a microplate reader (Versa max, Molecular Devices, San Jose, CA).

The absorbance of each well for F5 and each of its mutants was plotted against their concentration in each well to generate binding affinity curves. The absorbance values of each F5 were blank-corrected by subtracting the absorbance of the blank from the absorbance of F5. The blank-corrected absorbance values were normalized by dividing each blank-corrected absorbance value from the highest blank-corrected absorbance value for each F5. The binding affinity curves were obtained by plotting the percent normalized absorbance values on the y-axis called as normalized ELISA response and the corresponding concentration values in a log scale on the reversed x-axis for each F5. **Figure 4.9** shows the binding affinity curves obtained for each F5 mutant in comparison with F5 wild type. Among the five F5 mutants, binding affinities of F5-V51A and F5-P41A were less compared to F5 as expected from the mutations, which were

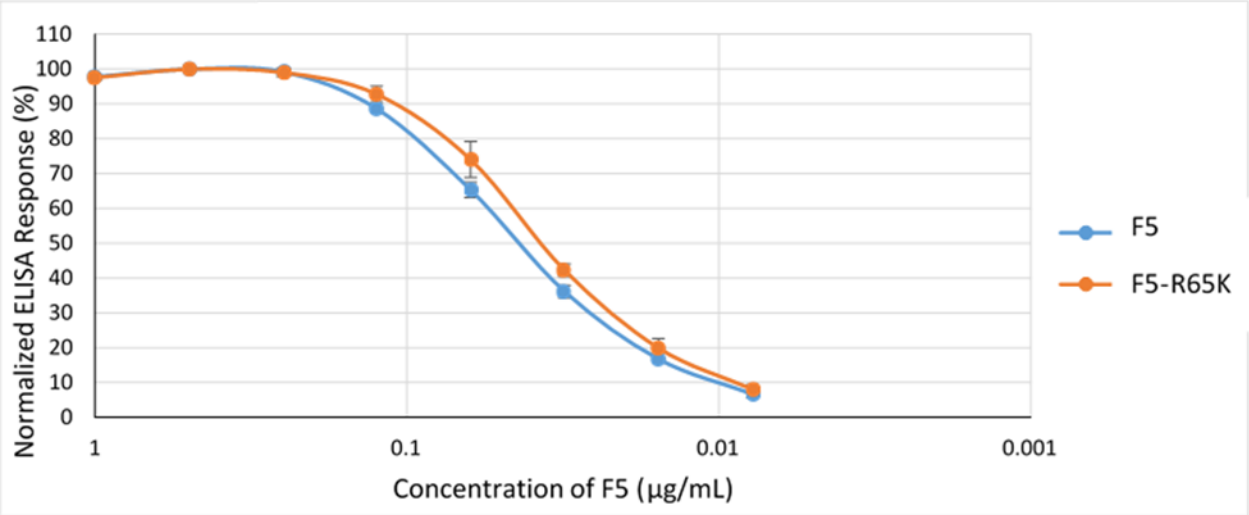
intended to detune the binding affinity of F5. However, binding affinities of F5-R65K and F5-L18V were relatively higher to the wild type, suggesting that these two mutations could have enhanced the affinity of F5 to RTA. Binding affinity curves of F5 and F5-E46D overlapped demonstrating that their binding affinities to RTA were similar.

To compare relative binding affinities of the F5 mutants, the F5 concentration at which ELISA response decreased to 50% , C_{50}, was determined using inverse regression.[21] The uncertainty associated with C_{50} values were calculated at 95% confidence interval. The C_{50} values calculated for each mutant are listed in **Table 4.4**. A C_{50} value which is higher than the C_{50} of the wild type indicates weaker binding to RTA while a C_{50} value that is less than the wild type C_{50} indicates stronger binding to RTA.

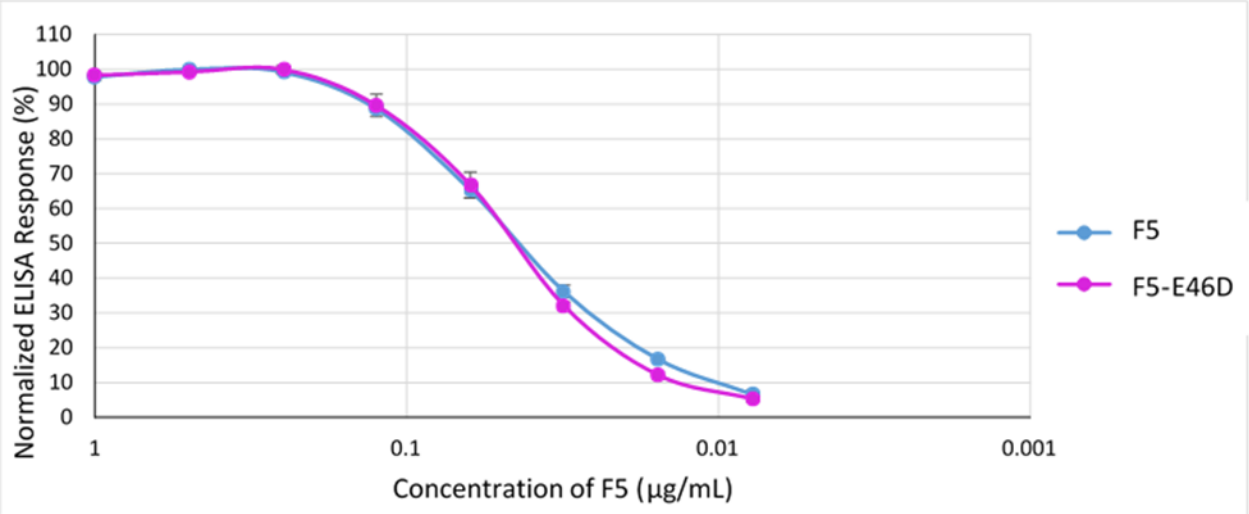
(a) F5-V51A



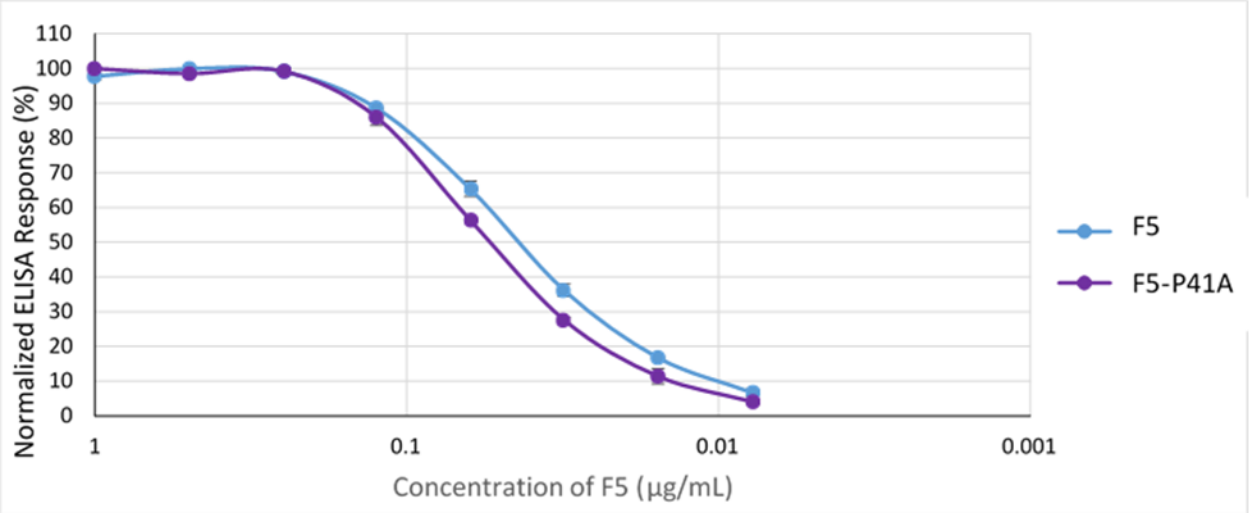
(b) F5-R65K



(C) F5-E46D



(d) F5-P41A



(e) F5-L18V

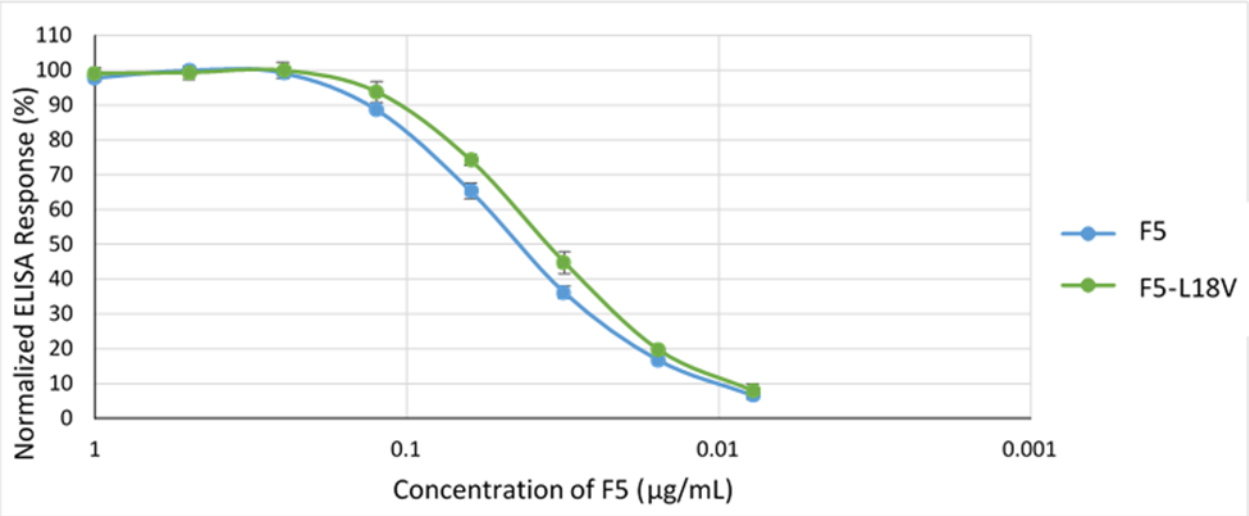


Figure 4.9: Binding affinity curves obtained for each of the five F5 mutants in comparison with binding affinity curve for F5 wild type. The arrow in panel (a) points to the C50 value of F5 wild type.

Table 4.4: C_{50} values (the F5 concentration at which ELISA response decreased to 50%) of the F5 mutants calculated using inverse regression.

Mutant	C_{50} (μg/mL) ± 95% confidence interval
F5	0.046 ± 0.002
F5-V51A	0.062 ± 0.002
F5-R65K	0.038 ± 0.004
F5-E46D	0.047 ± 0.003
F5-P41A	0.056 ± 0.001
F5-L18V	0.037 ± 0.003

Figure 4.10 shows a bar graph representing the C_{50} values and their 95% confidence intervals as error bars for each F5 mutant. The F5 mutants marked with stars; F5-V51A, F5-R65K, F5-P41A, and F5-L18V did not have overlapping confidence intervals in their C_{50} values with F5, indicating that their relative binding affinities were significantly different from the wild type. F5-V51A and F5-P41A had higher C_{50} values relative to the F5 demonstrating that these had relatively weaker binding affinities to RTA compared to the wild type. F5-R65K and F5-L18V had C_{50} values which were lower than F5 suggesting that they were able to bind with RTA with a relatively higher affinity compared to the wild type. F5-E46D had an overlapping confidence interval with thC_{50} value of F5 indicating that the mutant did not have any significant difference in its binding affinity for RTA relative to the wild-type affinity.

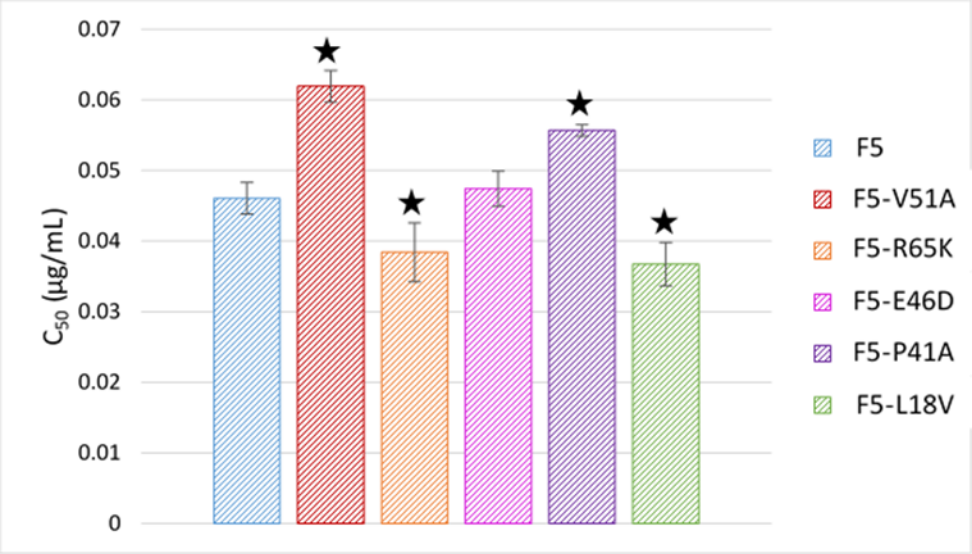


Figure 4.10: Bar graph showing the C_{50} values and their 95% confidence intervals for each F5 mutant. The stars represent the mutants that had significantly different relative binding affinities to the F5 wild type.

4.3 Discussion

The goal of this work was to generate a series of mutants of V_HH-F5 with subtle structural changes that would subtly detune its binding to RTA. Then, relative binding affinities of the mutants were determined using ELISA to assess if the mutants contain the expected detuning in binding to RTA. V_HH-F5 was selected for this work because this protein is well characterized in its crystal structure and binding to its antigen RTA. The mutations in F5 were carefully designed to be either conservative or semi-conservative to introduce only subtle alterations to its structure and thereby binding to RTA. Further, all the mutation sites were either proximal or distal to the paratope of F5.

All the F5 proteins were expressed and purified in the lab. The intact protein MS analysis confirmed that the F5 proteins were obtained with the expected sequences and the intact disulfide bonds. ELISA was performed to determine relative binding affinities of F5 and five of the mutants: V51A, R65K, E46D, P41A, and L18V to RTA. Relative binding affinities of the mutants can be used to assess whether the mutations were successful in detuning the binding of F5 to RTA. The largest binding affinity difference with the highest C_{50} (concentration of F5 at which the ELISA response decreased to 50%) value relative to F5 wild type was identified for the mutant F5-V51A. This demonstrates that V51A mutation diminished the binding of F5 to RTA. This mutation was located in the CDR framework region in the hydrophobic cleft proximal to the F5 paratope. Among all mutations, this mutation was the closest to the paratope (see **Figure 4.2**) and it had the largest effect on detuning F5 binding to RTA. The mutant F5_P41A also had a significant binding affinity difference with a higher C_{50} value relative to F5 but was less than F5-V51A. Therefore, this mutation also indicated weaker binding to RTA. The P41A mutation site was located in a reverse turn structure distal to the paratope of F5 as shown in **Figure 4.2**. It was interesting to see that the two mutants, F5-R65K in a distal salt bridge and F5-L18V at a distal hydrophobic cleft, had C_{50} values significantly smaller than F5. As these two mutations are considered to be conservative, these mutations could possibly have improved the binding of F5 to RTA. The mutant F5-E46D which was located at a proximal salt bridge did not have any significant difference in its binding affinity compared to F5. Therefore, the relative binding affinities of the mutants indicate that the mutations that were successful in causing subtle alterations in F5 binding to RTA were F5-V51A, F5-R65K, and F5-L18V.

As a continuation to this work, the remaining set of the mutants will also be expressed and purified in the lab and tested by ELISA for determining their relative binding affinities. Thereafter,

the mutants that are identified as having subtle differences in their binding will be examined by HX-MS. The goal of performing HX-MS will be to identify structural changes caused by the mutations. This will enable exploring if the mutants identified as having significant differences in binding will also be identified as having significant structural changes by HX measurements. Further, this study would enable assessing if all structural changes identified as significant by HX will be functionally meaningful as well. In addition, HX measurements obtained from the mutants will enable determining how small of a structural change that has a functional effect could be confidently detected by HX-MS and thereby establishing detection limits of HX-MS measurements.

4.4 References

1. Huang, R. Y.; Chen, G., Higher order structure characterization of protein therapeutics by hydrogen/deuterium exchange mass spectrometry. *Anal Bioanal Chem* **2014,** *406* (26), 6541-58.

2. Federici, M.; Lubiniecki, A.; Manikwar, P.; Volkin, D. B., Analytical lessons learned from selected therapeutic protein drug comparability studies. *Biologicals* **2013,** *41* (3), 131-47.

3. Houde, D.; Berkowitz, S. A.; Engen, J. R., The utility of hydrogen/deuterium exchange mass spectrometry in biopharmaceutical comparability studies. *J Pharm Sci* **2011,** *100* (6), 2071-86.

4. Wei, H.; Mo, J.; Tao, L.; Russell, R. J.; Tymiak, A. A.; Chen, G.; Iacob, R. E.; Engen, J. R., Hydrogen/deuterium exchange mass spectrometry for probing higher order structure of protein therapeutics: methodology and applications. *Drug Discovery Today* **2014,** *19* (1), 95-102.

5. Benhaim, M.; Lee, K. K.; Guttman, M., Tracking Higher Order Protein Structure by Hydrogen-Deuterium Exchange Mass Spectrometry. *Protein Pept Lett* **2019,** *26* (1), 16-26.

6. Muyldermans, S., Nanobodies: natural single-domain antibodies. *Annu Rev Biochem* **2013,** *82*, 775-97.

7. Dumoulin, M.; Conrath, K.; Van Meirhaeghe, A.; Meersman, F.; Heremans, K.; Frenken, L. G.; Muyldermans, S.; Wyns, L.; Matagne, A., Single-domain antibody fragments with high conformational stability. *Protein Sci* **2002,** *11* (3), 500-15.

8. Lauwereys, M.; Arbabi Ghahroudi, M.; Desmyter, A.; Kinne, J.; Hölzer, W.; De Genst, E.; Wyns, L.; Muyldermans, S., Potent enzyme inhibitors derived from dromedary heavy-chain antibodies. *EMBO J* **1998,** *17* (13), 3512-3520.

9. Muyldermans, S.; Cambillau, C.; Wyns, L., Recognition of antigens by single-domain antibody fragments: the superfluous luxury of paired domains. *Trends Biochem Sci* **2001,** *26* (4), 230-5.

10. Riechmann, L.; Muyldermans, S., Single domain antibodies: comparison of camel VH and camelised human VH domains. *J Immunol Methods* **1999,** *231* (1-2), 25-38.

11. Transue, T. R.; De Genst, E.; Ghahroudi, M. A.; Wyns, L.; Muyldermans, S., Camel single-domain antibody inhibits enzyme by mimicking carbohydrate substrate. *Proteins* **1998,** *32* (4), 515-22.

12. Rudolph, M. J.; Vance, D. J.; Cheung, J.; Franklin, M. C.; Burshteyn, F.; Cassidy, M. S.; Gary, E. N.; Herrera, C.; Shoemaker, C. B.; Mantis, N. J., Crystal structures of ricin toxin's enzymatic subunit (RTA) in complex with neutralizing and non-neutralizing single-chain antibodies. *J Mol Biol* **2014,** *426* (17), 3057-3068.

13. Vance, D. J.; Tremblay, J. M.; Mantis, N. J.; Shoemaker, C. B., Stepwise engineering of heterodimeric single domain camelid VHH antibodies that passively protect mice from ricin toxin. *J Biol Chem* **2013,** *288* (51), 36538-47.

14. Vance, D. J.; Tremblay, J. M.; Rong, Y.; Angalakurthi, S. K.; Volkin, D. B.; Middaugh, C. R.; Weis, D. D.; Shoemaker, C. B.; Mantis, N. J., High-Resolution Epitope Positioning of a Large Collection of Neutralizing and Nonneutralizing Single-Domain Antibodies on the Enzymatic and Binding Subunits of Ricin Toxin. *Clin Vaccine Immunol* **2017,** *24* (12).

15. Vance, D. J.; Mantis, N. J., Progress and challenges associated with the development of ricin toxin subunit vaccines. *Expert Rev Vaccines* **2016,** *15* (9), 1213-22.

16. Montfort, W.; Villafranca, J. E.; Monzingo, A. F.; Ernst, S. R.; Katzin, B.; Rutenber, E.; Xuong, N. H.; Hamlin, R.; Robertus, J. D., The three-dimensional structure of ricin at 2.8 A. *J Biol Chem* **1987,** *262* (11), 5398-403.

17. Rudolph, M. J.; Vance, D. J.; Cassidy, M. S.; Rong, Y.; Shoemaker, C. B.; Mantis, N. J., Structural analysis of nested neutralizing and non-neutralizing B cell epitopes on ricin toxin's enzymatic subunit. *Proteins* **2016,** *84* (8), 1162-72.

18. Waugh, D. S. TEV Protease FAQ. https://structbio.vanderbilt.edu/wetlab/private/vectors/Tev/tevnotes.Waugh.pdf (accessed October 1, 2021).

19. Angalakurthi, S. K.; Vance, D. J.; Rong, Y.; Nguyen, C. M. T.; Rudolph, M. J.; Volkin, D.; Middaugh, C. R.; Weis, D. D.; Mantis, N. J., A Collection of Single-Domain Antibodies that Crowd Ricin Toxin's Active Site. *Antibodies* **2018,** *7* (4), 45.

20. Thomas, J. C.; O'Hara, J. M.; Hu, L.; Gao, F. P.; Joshi, S. B.; Volkin, D. B.; Brey, R. N.; Fang, J.; Karanicolas, J.; Mantis, N. J.; Middaugh, C. R., Effect of single-point mutations on the stability and immunogenicity of a recombinant ricin A chain subunit vaccine antigen. *Human Vaccines & Immunotherapeutics* **2013,** *9* (4), 744-752.

21. Norman R. Draper, H. S., *Applied Regression Analysis* Third Edition ed.; John Wiley & Sons, Inc.: New York, 1998.

5 CHAPTER 5: Summary and future directions

5.1 Summary

In chapter 2, we assessed the potential of using HX-MS to predict deamidation propensities of proteins by investigating the correlation between hydrogen exchange and deamidation kinetics of proteins. Using MBP and its mutant as model proteins, we identified a power law correlation between hydrogen exchange half-life and deamidation half-life at NG site of the two proteins. This correlation demonstrated that HX-MS can be used to rank the deamidation sites in their order of deamidation propensities within a protein or among the protein states at different conditions. Our findings suggest that HX-MS can be used as a rapid method for predicting deamidation propensities of proteins compared to the conventional accelerated and long-term stability studies used for deamidation predictions.

In chapter 3, we identified unfolding events in three mAbs with variable sensitivities to aggregation when exposed to the air-water interface under agitation stress that leads their aggregation using HX-MS. The magnitudes of unfolding and destabilization observed in three mAbs were proportional to their aggregate formation propensities. Understanding structural dynamics of mAbs at the air-water interface causing them to aggregate is important to minimize particle formation in drug formulations because particles in drug formulations poses a significant risk in the safety of a drug.

In chapter 4, we expressed and purified a set of V_HH-F5 mutants to generate a panel of proteins with subtle higher-order structural changes and detuned binding to its antigen. Relative binding affinities of the V_HH-F5 mutants were determined using an ELISA assay to identify the mutant with detuned binding. The V_HH-F5 mutants with detuned binding will be used to assess the ability

of HX-MS to detect the introduced small changes in protein higher-order structure that altered the binding, thus limits of the detection in future work.

5.2 Concluding overview and future directions

In chapter 2, we identified a power law correlation between hydrogen exchange and deamidation kinetics at NG sites in the model protein MBP. This suggests that hydrogen exchange kinetics can be used to rank deamidation sites of a protein in order of their deamidation rates and compare deamidation rates at a deamidation site of a protein at different states or conditions. Therefore, this study demonstrates that HX-MS can offer a rapid way of predicting deamidation rates of proteins compared to time consuming accelerated and long-term stability studies.

This study can further be extended to explore the correlation between hydrogen exchange and deamidation kinetics in a variety of proteins such as monoclonal antibodies. This will be important to investigate if the correlation observed in this study could be common to all proteins or protein-specific. Moreover, to improve the resolution of the data from peptide averaged to residue level, conventional HX-MS approach can be coupled with ETD to obtain hydrogen exchange kinetics at the residue level. Residue level hydrogen exchange kinetics will provide more accurate prediction of the deamidation propensity at a deamidation site.

Aggregation that occurs in drug formulations in manufacturing and transportation under agitation stress is a concern because the presence of aggregates can pose potential risks to the safety, quality, and efficacy of a drug. Therefore, in chapter 2, we studied structural dynamics of three mAbs (mAb-α, mAb-β, and mAb-γ) leading to their aggregation when exposed to the air-water interface under agitation stress using HX-MS. These three mAbs had different propensities

to aggregate under agitation stress. The mAb which which had the lowest propensity to aggregate (mAb-β) partially unfolded in some regions of the Fab domain partly including the CDR domains. The mAb with intermediate propensity (mAb-α) to form aggregates partially unfolded in the LC CDR domain in the Fab and in some regions of the Fc domain. The mAb with the highest risk to form aggregates under agitation destabilized and unfolded extensively showing cooperative unfolding in the Fab domains. The presence of polysorbate 20 in the agitated samples was able to prevent the destabilization and unfolding occurred in all the three molecules under agitation stress. These observations indicate that mAbs get destabilized at the air-water interface and the extent of destabilization is different among the mAbs and likely to related to their aggregation propensities upon agitation.

This study demonstrates that HX-MS can be used to identify the regions of a molecule that undergoes destabilization under agitation stress and reveals information useful for reengineering molecules with potential liabilities. This study can be further extended to explore agitation-induced destabilization of variety of antibodies at the air-water interface leading them to form particles in drug formulations. Moreover, the HX-MS approach used in this work can be used to study the effect of different surfactants on the stability of drug formulations under agitation stress.

The goal of chapter 4 was to generate a panel of proteins with subtle differences in higher-order structure and binding affinities using the model protein V_HH-F5, to be used to determine the detection limits of HX-MS as a higher-order structural characterization technique. In this work, we were able to optimize the protein expression and purification workflow to successfully express and purify V_HH-F5 in the lab and optimize the ELISA assay to determine relative binding affinities of the expressed and purified V_HH-F5 and its mutants. One of the aims of generating the V_HH-F5

mutants was to identify mutants that detune the binding of V_HH-F5. The relative binding affinities determined by the ELISA assay indicated that two of the V_HH-F5 mutants among the five mutants studied had detuned binding to its antigen RTA compared to V_HH-F5 wild-type. Therefore, these two mutants will be used in the future work as two mutants of V_HH-F5 that are successful in achieving detuned binding.

As a continuation to the work done in chapter 4, the rest of the V_HH-F5 mutants designed for the study will be expressed and purified and their relative binding affinities will be determined using the ELISA assay. The V_HH-F5 mutants with detuned binding will be selected and analyzed by HX-MS to determine if HX-MS can detect changes in the higher-order structure of these mutants. Therefore, this panel of mutants will be used to explore if HX-MS can detect the structural changes arising from all the mutations that detuned binding to a significant extent. In addition, this panel of mutants will also be subjected to x-ray crystallographic analysis to determine the structural changes caused by the mutations as a collaborative work at NYSBC. Finally, the HX-MS measurements obtained from this panel of mutants will be used to explore how much of a small difference in higher-order structure that led to significant difference in binding affinity can be detected by HX-MS and thereby establish the detection limits of HX-MS.