University of Massachusetts Amherst

ScholarWorks@UMass Amherst

July 2020

STRUCTURAL ANALYSIS OF PROTEIN THERAPEUTICS USING COVALENT LABELING – MASS SPECTROMETRY

Patanachai Limpikirati

Follow this and additional works at: https://scholarworks.umass.edu/dissertations_2

⚙ Part of the Amino Acids, Peptides, and Proteins Commons, Analytical Chemistry Commons, Other Biochemistry, Biophysics, and Structural Biology Commons, and the Other Pharmacy and Pharmaceutical Sciences Commons

STRUCTURAL ANALYSIS OF PROTEIN THERAPEUTICS USING COVALENT LABELING – MASS SPECTROMETRY

A Dissertation Presented

by

PATANACHAI LIMPIKIRATI

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2020

Department of Chemistry

STRUCTURAL ANALYSIS OF PROTEIN THERAPEUTICS USING COVALENT LABELING – MASS SPECTROMETRY

A Dissertation Presented

by

PATANACHAI LIMPIKIRATI

Approved as to style and content by:

_____
Richard W. Vachet, Chair

_____
Igor A. Kaltashov, Member

_____
Min Chen, Member

_____
Stephen J. Eyles, Outside Member

                                  _____
                                   Ricardo B. Metz, Department Head
                                   Chemistry Department

DEDICATION

To my family, friends, and other people who always love and support me.

ACKNOWLEDGMENTS

I would like to thank my dissertation committee members, Prof. Igor Kaltashov, Prof. Min Chen, and Prof. Stephen Eyles. I appreciate their advice and discussions, as well as challenging questions, which helped refine my prospectus, ORP and dissertation.

I am grateful to the Institute for Applied Life Sciences, UMass where many state-of-the-art instruments belong to. I acknowledge Dr. Stephen Eyles, Dr. Adam Graichen and Dr. Cedric Bobst for their efforts in operating the mass spectrometry center. Dr. Lizz Bartlett is thanked for her role in the biophysical characterization core facility. I sincerely thank Steve and Liz for their generosity and patience in helping me learn how to use instruments and troubleshoot issues, and they always suggest me useful ideas for my experiments. I also thank Adam for helping our group fix instruments many times, at certain point even as late as 1.30 am!

I also thank my collaborators, Eric Graban, Dr. Robert Vaughan, and Dr. John Hale (QuarryBio Inc.). I am very glad that I had a good chance to work with QuarryBio team. I thank all of them for providing experimental materials, helping with data analysis pipeline, and giving feedbacks to my research. We had a good teamwork and we could finally publish our nice paper. They are gratefully thanked for allowing our group to use their custom software pipeline for data analysis.

I would like to thank the past and present members in Vachet group – Nick, Gokhan, Alyssa, Zhe, MAC, Tyler, Bo, Meizhe, Tianying, Kristen, Blaise, Laura, Xiao, Catherine, Stacey, Zack, Dheeraj, Jack, Plabo, Orn and Ping. Thanks for interacting with and helping me in the lab, and thanks all of you for making an atmosphere of this lab lively, interactive and secure. Thank Nick, MAC, Tyler, Tianying, and Bo for training me how to use instruments. Thanks Kristen, Blaise, Tyler, and Catherine for helping me edit some of

my writings. Special thanks to MAC, Bo, Meizhe, Tianying, Laura, Xiao, Stacey, Orn and Ping whom I talked a lot and shared many feelings with. Ellen is also thanked for being our lab's administrative assistant. She is always nice, kind and very helpful. I also thank a former group member Tan who gave me a lot of advice about Ph.D. study at UMass and life in Amherst.

Over the course of my doctoral study, I thank my past and present friends in Chemistry and MCB programs, especially Jake, Rob, Joe, Mahdieh, Lin Hui, Jiale, Kirandeep, Vikash, Jiaming, Kingshuk, Fanjun, etc for helping me and supporting me. Special thanks to Lin Hui and Jiale whom I talked a lot with. I thank faculty and staff members of the UMass Chemistry department who keep this Ph.D. program running well. With special thanks to former and current graduate program coordinators, jms and Chandra, and thanks Bob who always help our group fix instruments.

In addition, I am very grateful to a doctoral fellowship from the Faculty of Pharmaceutical Sciences, Chulalongkorn University (Bangkok, Thailand). I would like to thank faculty and staff members there who helped me regarding this scholarship, and thanks those who stay in touch with me. I also acknowledge the International Programs Office, UMass and the Office of Educational Affairs, Royal Thai Embassy, Washington DC for their roles in coordinating this fellowship.

I would like to thank my past and present Thai friends at UMass - Chanti, Amorn, Chompoo, Kade, Champi & Dave, Ying, Ye & Justin, Tick, Aom, Palm, Bird, Kratae, Gam, Nok, AP, Tae & Victor, Ping, Piyachai, New, Feem, Ed, Nutt, Kik, Boss, Orn, Firm, Nui, Kaew, Mind, etc. It is truly relaxing to have activities with friends after hard work. Thanks for having fun time together. We have had many parties and trips, and I have

developed my skills to be a (professional) trip planner, taxi driver, and bathroom cleaner! I thank many of them whom I can share my feelings with. I really thank Feem, Piyachai, and AP for being wonderful roommates to me. Special thanks to Nok whom I met in my first year, she is like a big sister who keeps in touch with me and always makes me feel positive. I also thank my aunt and uncle in California and all other friends in the U.S. and back in Thailand. Thanks for spending time with me and always giving me a moral support during my study. All their supports indeed encourage me to finish this career.

I would like to thank my church family at the First Baptist Church where I worship and serve my Lord. Spiritually, I have felt His presence in my life, and I am blessed with the ability and confidence to get through many difficult times during the study. I have special thanks to the Chinese Ministry and Pastor Lily Soh. Lily is my spiritual mentor who helps me grow in faith and counsels me many times regarding my personal issues. I believe that through His grace I have been given the opportunity to overcome my problems and pursue the purpose of life.

Last but not least, I thank my parents "Poh" and "Mae" and my sister "Nong Pa" who unconditionally love and care for me. Studying abroad is a big journey for me, and I have to live very far away from them for many years. Despite the distance, we are still connected by our love. They constantly trust and support me and always offer me what they can do for me. I love and thank them for everything!

ABSTRACT

STRUCTURAL ANALYSIS OF PROTEIN THERAPEUTICS USING COVALENT LABELING – MASS SPECTROMETRY

MAY 2020

PATANACHAI LIMPIKIRATI,

B.Sc. (Pharm), CHULALONGKORN UNIVERSITY

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Richard W. Vachet

Using mass spectrometry (MS) to obtain information about a higher order structure of protein requires that a protein's structural properties are encoded into the mass of that protein. Covalent labeling (CL) with reagents that can irreversibly modify solvent accessible amino acid side chains is an effective way to encode structural information into the mass of a protein, as this information can be read-out in a straightforward manner using standard MS-based proteomics techniques. The differential reactivity of proteins under two or more conditions can be used to distinguish protein topologies, conformations, and/or binding sites. CL-MS methods have been effectively used for the structural analysis of proteins and protein therapeutics.

This dissertation focuses on the use of a diethylpyrocarbonate (DEPC-based CL-MS method to characterize the higher-order structure of protein therapeutics. DEPC is a simple to use, commercially-available covalent labeling reagent that can readily react with a range of nucleophilic residues in proteins. We find that in intact proteins weakly nucleophilic side chains (Ser, Thr, and Tyr) can be modified by DEPC in addition to other residues such as His, Lys, and Cys, providing very good structural resolution. We

hypothesize that the microenvironment around these side chains, as formed by a protein's higher order structure, tunes their reactivity such that they can be labeled. To test this hypothesis, we compare DEPC labeling reactivity of Ser, Thr, and Tyr residues in intact proteins with peptide fragments from the same proteins. Results indicate that these residues almost never react with DEPC in free peptides, supporting the hypothesis that a protein's local microenvironment tunes the reactivity of these residues. From a close examination of the structural features near the reactive residues, we find that nearby hydrophobic residues are essential, suggesting that the enhanced reactivity of certain Ser, Thr, and Tyr residues occurs due to higher local concentrations of DEPC.

Monoclonal antibodies (mAbs) are among the fastest growing therapeutics in the pharmaceutical industry. Detecting higher-order structure changes of antibodies upon storage or mishandling, however, is a challenging problem. In this dissertation, we describe the use of DEPC-based CL-MS to detect conformational changes caused by heat stress, using rituximab as a model system. The structural resolution obtained from DEPC CL-MS is high enough to probe subtle conformation changes that are not detectable by common biophysical techniques. Results demonstrate that DEPC CL-MS can detect and identify sites of conformational changes at the temperatures below the antibody melting temperature (e.g., 55 ^{O}C). The observed labeling changes at lower temperatures are validated by activity assays that indicate changes in the F_{ab} region. At higher temperatures (e.g., 65 ^{O}C), conformational changes and aggregation sites are identified from changes in CL levels, and these results are confirmed by complementary biophysical and activity measurements. Given the sensitivity and simplicity of DEPC CL-MS, this method should be amenable to the structural investigations of other antibody therapeutics.

Reliable information about antibody higher-order structure can be obtained, though, only when the protein's structural integrity is preserved during labeling. In this dissertation, we have evaluated the applicability of DEPC reaction kinetics for ensuring the structural integrity of mAbs during labeling. By monitoring the modification extent of selected proteolytic fragments as a function of DEPC concentration, we find that a common DEPC concentration can be used for different monoclonal antibodies in formulated samples without perturbing their higher-order structure. Under these labeling conditions, we find that the antibodies can accommodate up to four DEPC modifications without being structurally perturbed, indicating that multi-domain proteins can withstand more than one label, which contrasts to previously studied single-domain proteins. This more extensive labeling provides a more sensitive measure of structure, making DEPC-based CL-MS suitable for the higher-order structural analyses of mAbs.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

COVALENT LABELING – MASS SPECTROMETRY FOR STRUCTURAL

ANALYSIS OF PROTEINS AND PROTEIN THERAPEUTICS

1.1 Introduction to Covalent Labeling – Mass Spectrometry

A protein's higher order structure (HOS) determines its function, and so methods that provide insight into protein structure are important for understanding protein reactivity. Traditionally, NMR spectroscopy or X-ray crystallography have been the methods of choice because of the atomic-level resolution afforded by these techniques. Recently, cryo-electron microscopy (cryo-EM) has also emerged as a very powerful tool for studying protein structure. In some cases, however, protein HOS cannot be properly studied by these techniques because of limited sample amounts, a given protein's tendency to aggregate, or sample incompatibility. Moreover, in certain applications (e.g. protein therapeutic design) more rapid structural analysis tools are needed. Because of the inherent sensitivity, specificity, and speed of mass spectrometry (MS), methods based on this technique have emerged for the analysis of a protein's solution HOS. These MS-based methods add to the toolbox of protein biochemists by offering approaches that give much higher structural resolution than techniques such as circular dichroism (CD) spectroscopy or fluorescence spectroscopy while at the same time being more routinely applicable and devoid of the limitations associated with NMR, X-ray crystallography, or cryo-EM.

Using MS to obtain information about a protein's structure in solution requires that the protein's structural properties are encoded into the mass of the protein. Three primary methods of encoding this structural information have been utilized, including hydrogen-deuterium exchange (HDX),[1-7] cross-linking,[8-10] and covalent labeling.[11-13] When coupled with MS, HDX relies on the replacement of hydrogens by deuteriums on backbone amides, leading to mass increases in regions of the protein that are the least protected from this exchange. In doing so, HDX/MS provides insight into structured/unstructured and rigid/dynamic regions of a protein that can then be related to HOS. Cross-linking uses multi-functional reagents that can form new intra- or inter-molecular bonds between amino acid side chains in a protein or protein-protein complex. The choice of the cross-linking agent leads to defined distance constraints for the linked side chains that enables one to deduce HOS information. Covalent labeling (CL) methods are analogous to cross-linking methods in that they modify protein side chains, but they report on a protein's surface structure, which can then be used to infer information about its HOS.

All three MS-based approaches typically use proteolytic digestion, liquid chromatographic separation, and mass spectrometric analysis to identify modification sites as a way to provide localized or amino acid-level structural information; however, HDX and cross-linking methods have analysis challenges that are not present in CL techniques. In HDX, the inherent reversibility and lability of the modification can lead to back-exchange and scrambling that must be minimized via the use of specialized sample handling techniques. In cross-linking experiments, branched polypeptide chains are necessarily produced, requiring custom software and often specially-designed reagents to facilitate identification of the cross-linked sites. CL methods do not suffer from these

limitations. Side chain modifications are irreversible, enabling the use of well-established proteomics workflows and sample handling techniques to facilitate the identification of protein modification sites. In comparison to cross-linking methods, identifying modified peptides and pinpointing their labeled residues is more straightforward.

CL methods also offer some valuable attributes for the structural analysis of proteins. CL reactions occur on protein side chains, providing complementary information to HDX/MS, which probes the protein backbone. Because most protein complexes (e.g. protein-protein, protein-ligand) are primarily mediated by side chain interactions, it can be argued that amongst MS-based methods CL techniques are perhaps best suited for identifying interaction sites in protein complexes. A unique attribute of some CL reagents is the ability to study protein folding reactions that occur on the μsec timescale, which is a timeframe inaccessible by most techniques. The very fast reaction kinetics of reagents such as hydroxyl radicals and carbenes makes this possible. Given the advantages associated with analyzing covalently labeled proteins and peptides, as well as the distinctive features this technique offers, there has been an increasing interest in developing CL methods for studying protein HOS.

The goal of this introductory chapter is to provide an overview of the most commonly used CL approaches. While there have been a large number of studies using amino acid specific labeling reagents together with MS to study protein structure[12], this chapter will focus on labeling reagents that are non-specific. Non-specific reagents are ones capable of modifying a wide variety of amino acid side chains simultaneously, thereby enabling broader structural coverage in a single experiment. The basic CL experiment is illustrated in **Figure 1.1**. A protein or protein complex of interest is exposed to a particular

labeling reagent (see below for details), allowing solvent exposed amino acids to be modified. After halting the reaction, the modification sites on the protein are then identified using MS. Typically, this is done via proteolytic digestion and LC/MS/MS analysis to pinpoint the specific amino acid sites that have been modified, so structural information can be deduced. Structural information about a protein or protein complex is usually gathered by comparing a protein's differential reactivity under two conditions (e.g. monomer vs. dimer to determine a binding interface as illustrated in **Figure 1.1**). A key assumption in all CL experiments is that amino acid residue will react to an extent that depends on their solvent accessibility, although solvent accessibility is not the only factor that influences the reactivity of a given amino acid residue.
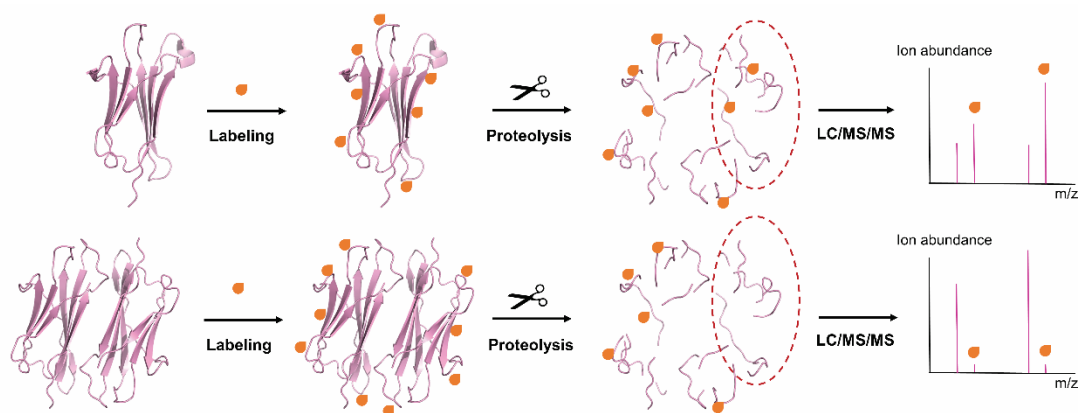
Figure 1.1: Scheme showing CL with MS detection.
The label modifies solvent accessible amino acids. The modified protein or protein complex is subjected to proteolytic digestion and the modified peptides are analyzed using LC-MS/MS. Sites of protein conformational changes and/or protein-protein interactions can be revealed by changes in the extent of labeling at specific residues.

1.2 Factors Affecting Covalent Labeling Reactivity

1.2.1 Reagent factors

Several reagents have been used to covalently modify amino acid side chains through oxidation or bioconjugate chemical reactions.[12, 14, 15] A portion of these reagents

are used in covalent labeling – mass spectrometry (CL-MS) techniques for structural analysis of proteins and other macromolecules. In this section, we will describe how CL reactions are initiated and quenched, and the types of products that are generated.

1.2.1.1 Hydroxyl Radicals

Hydroxyl radicals ($^\bullet$OH) have been the most commonly used CL reagents when combined with MS detection. They are typically generated through radiolysis or photolysis of water or hydrogen peroxide (H_2O_2).[11] The resulting $^\bullet$OH radicals then modify amino acid side chains on a protein's surface via a cascade of reactions that typically begins with hydrogen abstraction and ends with formation of a new covalent bond on the side chain. High-flux X-ray or γ-ray radiation sources (e.g from a synchrotron) can be used to directly generate $^\bullet$OH radicals via the radiolysis of water.[16] Several excellent reviews from the Chance group have described synchrotron footprinting and its applications.[17, 18] One challenge associated with this CL approach is that a synchrotron source is necessary to directly produce $^\bullet$OH radicals, which limits its wide availability and applicability. In this synchrotron approach, $^\bullet$OH radicals are generated on a millisecond timescale, meaning protein unfolding could conceivably occur and compete with the labeling,[19, 20] although adding radical scavengers to the solution can shorten the lifetime of the hydroxyl radicals and therefore shorten the reaction time frame. It is important to stress that the applicability of oxidative labeling reactions is determined not only by the radical lifetime but also the timescale of radical formation.

UV-light induced photolysis of H_2O_2 in aqueous solution is another common way to generate $^\bullet$OH radicals;[21] however, long UV irradiation times and concentrations of H_2O_2 up to 15% by volume are needed unless a UV laser is used.[22, 23] Laser photolysis by a

Nd:YAG laser (266 nm)[22] or KrF excimer (248 nm) laser[23] was developed by Sze and co-workers and Gross and co-workers, and this approach allows H_2O_2 concentrations of less than 1% to be used, thereby limiting direct oxidation of sulfur groups by H_2O_2. Short laser pulses can shorten the time scale of radical formation to the nsec and µsec timescale.[22, 23] When used with the appropriate solution additives, this laser irradiation approach can oxidize proteins faster than they unfold, ensuring that there is no structural perturbation during labeling.[24] Gross and co-workers later coined the term "fast photochemical oxidation of protein" or FPOP for this approach.

Other reactions can also be used to generate $^\bullet$OH radicals for macromolecular labeling, including electron pulse radiolysis,[25] metal-catalyzed oxidation (MCO) reactions,[26-28] Fenton chemistry,[29-31] disproportionation of peroxynitrous acid,[32, 33] high voltage electrical discharge,[34-36] electrochemical oxidation,[37] and more recently, plasma generation.[38] Each of these $^\bullet$OH radical generation methods has unique aspects to their chemistry, but because they are not as widely used in CL of proteins, they will not be discussed further here. Interested readers are referred to a comprehensive review of hydroxyl radical CL by Xu and Chance.[11]

The general oxidation reactions that occur during radiolysis or UV photolysis are briefly summarized in **Figure 1.2a**. These reactions typically follow pseudo first-order kinetics[16, 24, 39] because of the excess concentration of $^\bullet$OH radicals. At least 14 of the 20 common amino acid side chains can be modified by $^\bullet$OH radicals, and over 50 different modification types on proteins can be generated.[11, 13] The most common products of aliphatic amino acids are addition of a hydroxyl group or a carbonyl group on the side chain, resulting in mass shifts of +16 Da and +14 Da, respectively. For aromatic residues,

the addition of one or multiple hydroxyl groups to the aromatic ring results in mass shifts of +16. Asp and Glu residues generally undergo a -30 Da mass change due to oxidative decarboxylation, while basic side chains give rise to a series of unique oxidized products that include mass additions and side chain cleavages.[11, 13] More information about the residue-specific oxidation products can be found elsewhere.[11]

While $^{\bullet}$OH radicals can undergo self-quenching reactions in aqueous solution, supplemental reagents are typically added to solution to control the lifetime of the radicals or prevent unwanted side reactions. Secondary reactions from the presence of H_2O_2 and other oxidative species produced upon radiolysis are found to over-oxidize Met and Cys residues, thereby affecting analytical reproducibility and data interpretation regarding solvent accessibility.[40] Often, amino acids or other molecules possessing good reactivities towards $^{\bullet}$OH are selected as reaction quenchers. Addition of catalase or methionine have been found to minimize this secondary oxidation and improve quantitative protein labeling [40]. In FPOP, Gln or Phe is usually added prior to irradiation to scavenge $^{\bullet}$OH radicals and shorten their lifetimes in solution to the μsec time scale, so that oxidation can occur before any significant protein structural changes.[23, 24] Even with these solution additives to control the highly reactive $^{\bullet}$OH radical chemistry, evidence has been provided that the timescale of protein oxidation can extend to tens of msec because of the formation of longer-lived secondary or higher order radicals.[41]
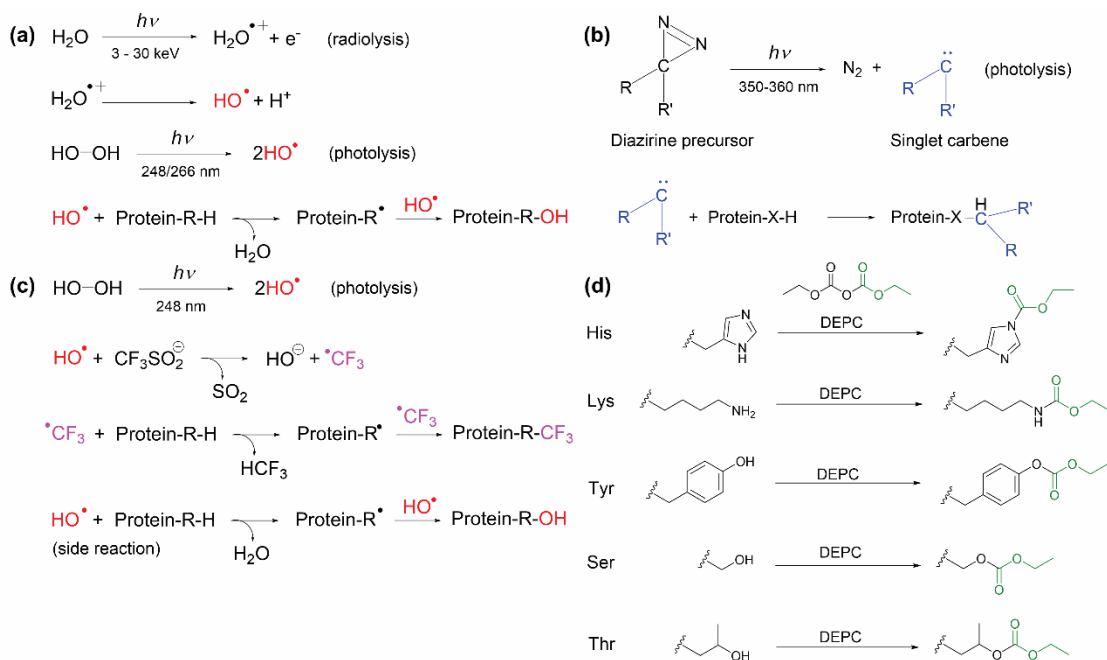
Figure 1.2: Modification reactions of amino acid residues used in non-specific covalent labeling with (a) hydroxyl radicals, (b) carbenes, (c) trifluoromethylation, and (d) diethylpyrocarbonate.

1.2.1.2 Carbenes

Highly reactive singlet carbenes can be generated from the photolysis of diazirine derivatives **(Figure 1.2b)** by using near-UV wavelengths ($\approx$350 nm) and like hydroxyl radicals, they can also be used for CL.[42-44] Carbenes can rapidly insert into any X−H bond (where X can be C, O, N, or S), and thus can potentially label any amino acid residues on a protein surface.[44, 45] Diazirine gas (CH_2N_2) was first used as a carbene precursor,[46-49] but its poor solubility limits the extent to which proteins can be labeled with this reagent. A more useful carbene precursor is L-2-amino-4,4-azipentanoic acid (or photoleucine), which has good stability and solubility in water and has been successfully used to label proteins in several studies.[44, 50, 51] Recently, Manzi *et al*. reported a new aromatic diazirine

precursor, 3-trifluoromethyl-3-phenyldiazirine, that has higher reaction efficiency than photoleucine.[52]

The common reactions that occur during carbene production and labeling are briefly summarized in **Figure 1.2b**. The products of carbene labeling will have a substituent group of the diazirine precursor inserted into the amino acid side chains, whose mass shift is readily detected using MS. Carbene labeling is rapid, irreversible, and independent of protein concentration, implying the zero-order kinetics.[44] However, protein-dependent reactivity has been observed in some protein systems.[52] Addition of quenching agents is not necessary as carbenes are readily quenched by water. The lifetime of carbenes in aqueous solution is on the nsec time scale due to its rapid reaction with water, which allows carbene labeling to be faster than protein unfolding.[44, 50] A range of intermediates generated from aliphatic diazirines upon photolysis, such as triplet-state carbenes and diazo-mediated carbocations (diazo isomers), can result in side reactions.[53] These side reactions have been reported to be reduced somewhat by tuning functional groups on the diazirine precursor.[52, 54] Oxidation side products have also been observed after irradiation but typically at low levels.[53]

1.2.1.3 Trifluoromethylation (CF$_3$)

Very recently, $^\bullet$CF$_3$ radicals have been introduced as labeling reagents that can be generated by pulsed laser photolysis of triflinate (Langlois' reagent).[55] Laser irradiation is performed in the presence of H_2O_2 ($^\bullet$OH source) and the water-soluble salt $NaSO_2CF_3$. The resulting $^\bullet$OH rapidly reacts with excess $[SO_2CF_3]^-$ to generate $^\bullet$CF$_3$ (**Figure 1.2c**). The highly reactive $^\bullet$CF$_3$ can insert into X-H bonds (where X is C, O, N, or S) of amino acid side chains, resulting in a mass shift of +67.987 Da. $^\bullet$CF$_3$ reacts with 18 of the 20 common

amino acids, including those that are relatively unreactive with hydroxyl radicals (Gly, Ala, Ser, Thr, Asp, and Glu). Because hydroxyl radicals are necessary to generate $^\bullet$CF$_3$, side reactions with $^\bullet$OH or dissolved oxygen species during photolysis yield oxygen-incorporated products, although CF$_3$-substitued products dominate.[55] Radical trifluoromethylation likely follows pseudo first-order kinetics because of the excess concentration of [SO$_2$CF$_3$]$^-$ in the reaction mixture, but this has not been verified yet. There is no need for addition of quenchers as the reactive $^\bullet$CF$_3$ is readily quenched in water, having a lifetime of $\sim$ 30 msec.[56] Using the FPOP platform along with the appropriate solution quenchers (e.g. catalase and methionine), $^\bullet$CF$_3$-based CL should be faster than protein unfolding, ensuring structural integrity upon labeling.[24]

1.2.1.4 Diethylpyrocarbonate (DEPC)

Diethylpyrocarbonate (DEPC) is a commercially-available reagent that can react with a range of nucleophilic residues. Unlike the radical reagents where specialized equipment is needed to generate the radicals, DEPC directly labels proteins when added to solution. A stock solution of the reagent is typically prepared in anhydrous acetonitrile due to its limited solubility and propensity to be hydrolyzed in water. DEPC concentrations ranging from 0.01 mM to 40 mM are readily soluble in aqueous solutions and are also useful concentrations for labeling proteins.[57] While DEPC will eventually be hydrolyzed in water, reactions between this reagent and proteins are typically quenched after a short time (10 sec – 60 sec) by adding relatively high concentrations (~20 mM) of a nucleophilic compound such as imidazole.[12, 58]

DEPC is a reactive electrophile that can modify nucleophilic side chains (Cys, His, Lys, Thr, Tyr, Ser) and N-termini via nucleophilic substitution reactions (**Figure 1.2d**).[12,

[57-59] Under the most relevant protein labeling conditions, the reaction follows second-order kinetics, meaning the reaction rate depends on both protein and DEPC concentrations.[58] Because DEPC reacts more slowly than radical reagents, its concentration and reaction time must be carefully controlled to prevent over-modification and preserve the structural integrity of the protein during the labeling reaction. Carbethoxylated products with a mass shift of +72.021 Da are obtained for Cys, His, Lys, Thr, Tyr, Ser, and the N-terminus. One key advantage of this labeling reagent is that it generates a single type of product, which simplifies identification of labeled sites and improves sensitivity as the signal is not distributed among numerous products. The addition of a second carbethoxyl group to His residues has been reported, but this modification is typically avoided by labeling at the low DEPC concentrations necessary to ensure the structural integrity of proteins.[12] Some modified residues, especially Ser and Thr, are subject to hydrolysis and thus label loss if they remain in solution too long, so proteolytic digestion and LC/MS/MS analysis must be performed soon after the labeling reaction is completed.[59] Moreover, DEPC label scrambling can happen in solution if free Cys residues are available. This label scrambling can be easily avoided, though, by alkylating any free thiols after disulfide reduction,[60] as is typically done in most proteomics experiments.

1.2.2 Protein Factors

Structural analysis of proteins using CL-MS relies on the fact that in different protein conformers a certain set of residues possesses differential reactivity with the labeling reagent, and the resulting difference in modification extents of those amino acids can be used to distinguish protein topologies, conformations, and/or binding sites. The protein factors that affect the reactivity of amino acid side chains are described below.

11

1.2.2.1 Solvent Accessibility

All CL reactions used to study protein structures are performed in the aqueous phase, to preserve the HOS of proteins. The implicit assumption in CL-MS is that amino acids that are exposed to solvent and accessible to the CL reagent can be modified. Meanwhile, buried residues will be modified slowly or not at all (**Figure 1.3**). However, solvent accessibility is not the sole factor that governs reactivity. Different amino acid side chains can react differently with a given labeling reagent, as will be discussed in **Section 1.2.2.2.** Several groups have established that a qualitative relationship exists between solvent accessible surface area (SASA) and extent of labeling for certain types of residues.[11] For example, Chance and co-workers have found that hydroxyl radicals primarily react with surface accessible residues[11, 16, 61, 62] and that residues at protein-protein interfaces are generally protected from modification.[61] Moreover, in several proteins, oxidation rates are found to correlate fairly well with calculated SASAs of residues from known NMR and X-ray crystal structures.[21, 23, 62] A clear relationship between SASA of individual amino acid residues and carbene reactivity has not been observed, possibly due to the carbene precursor's affinity for certain residues, the high inherent reactivity of the carbene itself, and the different intrinsic reactivities of amino acid residues towards carbene labeling.[44, 50] Despite these issues, Delfino and co-workers have shown using diazirine gas as a carbene precursor that global levels of solvent accessibility for different α-lactalbumin and β-lactamase conformers are related to carbene labeling.[47, 49] This group also found a similar relationship for antibody-bound lysozyme.[48] More recently, an improved correlation between SASA and extent of carbene labels at the residue level has been demonstrated in experiments where the protein sample is flash-frozen during

irradiation[53] (see **Section 1.2.2.2** for further discussion of this observation). For trifluoromethylation, Gross *et al*. have reported that this reagent's reactivity patterns are consistent with SASA in a model membrane protein.[55] Our group has also demonstrated that DEPC labeling can be used to study protein topology, and DEPC modification extents for certain types of residues are found to be consistent with solvent accessibility.[58] Similarly, decreases in DEPC modification levels upon binding to transition metals are consistent with the general trend, and actually allow metal binding sites to be determined.[28, 58, 63]

Differences in SASA of reactive side chains, resulting from their involvement in a conformational change and/or protein-ligand binding, will also cause amino acids to be more or less accessible to a given CL reagent and therefore, react to greater or lesser extents. Most recently, researchers have attempted to further refine and quantify the relationship between SASA and labeling extent. Yang *et al.* and Sharp *et al.* have proposed the use a protection factor (PF) to quantitatively relate hydroxyl radical reaction rates to amino acid SASA using well-characterized protein models.[64, 65] These PF are obtained by normalizing the measured labeling rate constant with the side chain's intrinsic reactivity. More details about the applications of PF in mapping conformation changes and in modeling protein structures can be found elsewhere.[13, 66]
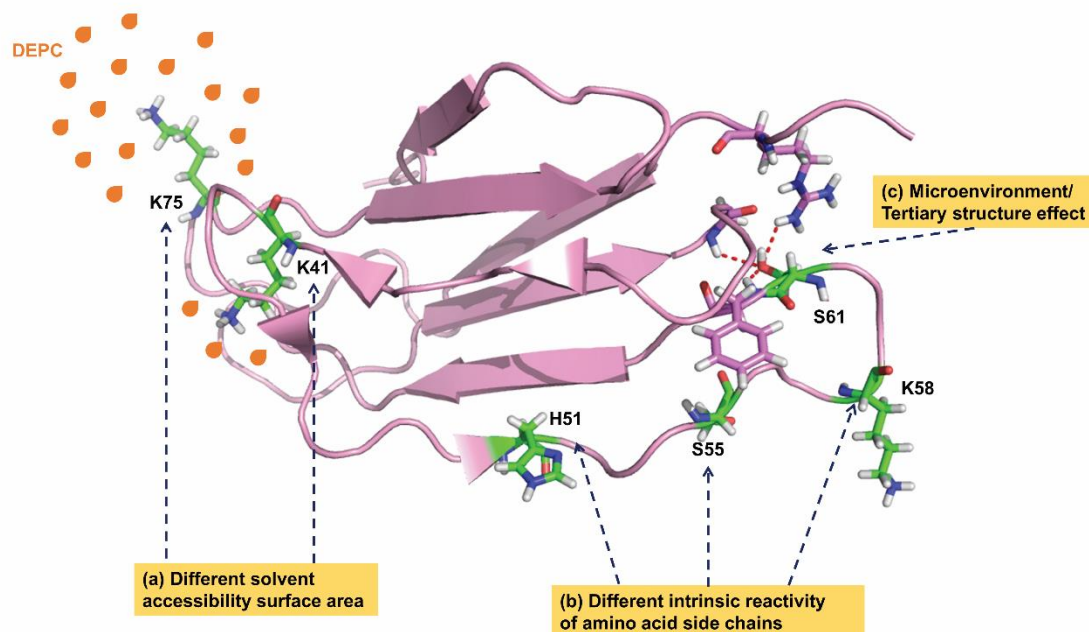
Figure 1.3: Protein factors that affect the reactivity of CL reagents.
Factors include (a) solvent accessible surface area (SASA), (b) the intrinsic reactivity of amino acid residues, and (c) primary and tertiary structure.

1.2.2.2 Intrinsic Reactivity of Amino Acid Side Chains

Attempts to use a PF to quantify the relationship between SASA and reactivity imply that the intrinsic reactivity of a given amino acid side chain influences its reactivity (**Figure 1.3**). Thus, this factor must be considered when interpreting CL data. While $^{\bullet}$OH is highly reactive and can modify a wide range of amino acid side chains, rate constants for reactions with the 20 common amino acid residues vary over three orders of magnitude with the following reactivity order: Cys > Met > Trp > Tyr > Phe > His > Leu, Ile > Arg, Lys, Val > Ser, Thr, Pro > Gln, Glu > Asp, Asn > Ala > Gly.[67] Although every amino acid can react with hydroxyl radicals, practically speaking Gly, Ala, Ser, Thr, Asp, and Glu are often not found labeled during protein CL experiments with hydroxyl radicals. Therefore, only 14 of the 20 side chains are useful in typical labeling experiments with hydroxyl

radicals, and these residues typically account for ∼65% of the sequence of the average protein.[11, 67] Sharp *et al.* have demonstrated that normalization of a residue's intrinsic reactivity via a PF is needed to obtain a qualitative correlation between SASA and labeling extent.[65]

While CL with carbenes was initially predicted to be affected less by differences in amino acid side chain chemistry than hydroxyl radicals, studies suggest that there is a decidedly non-uniform distribution in carbene labeling of proteins.[44] Schriemer and co-workers have found that using photoleucine as a diazirine precursor leads to Glu residues being preferentially modified, which might be explained by the reagent's affinity for negatively charged residues or the reactive carbenes favoring polar protic bond (O-H) insertion rather than C-H bond insertion.[50] Oldham *et al.* have also found that hydrophobic and basic residues are favored in the carbene labeling using an aryldiazirine precursor.[52] In the more recent study from Schriemer *et al*, surface bias is still found in the CL with carbenes generated from other different aliphatic diazirine precursors.[53] By performing carbene footprinting on flash-frozen samples to limit the diffusion of carbenes to sites of higher reactivity during irradiation, the preferred sites of labeling tend to reflect the sidechain interactions with diazirine precursor, which is influenced by the diazirine substituents rather than the intrinsic reactivity of residues.[50, 52, 53] The resulting surface biases can be avoided somewhat by tailoring the diazirine precursor or changing solution conditions to reduce the reagent's affinity toward specific residues.[50, 52]

Because •CF₃ radicals have only recently emerged as CL reagents, the full details of their reactivity have not been studied enough yet. However, aromatic residues (e.g. Trp and Phe) are found to be modified at relatively high levels, indicating their good intrinsic

reactivity toward $^\bullet CF_3$ labeling.[55] For DEPC labeling, the intrinsic reactivity of side chains

is governed by the nucleophilicity of residues.[57-59] Thus, the reactivity of Tyr, Ser, and Thr

residues is lower than the reactivity of Cys, His, and Lys residues due to their weakly

nucleophilic hydroxyl groups.[58]

1.2.2.3 Primary and Higher Order Structure Effects

In the structure of an intact protein, the sequence context (primary structure) may

contribute to the reactivity of amino acid due to electron donating and/or electron

withdrawing effects. This fact appears to be true even for hydroxyl radical labeling. Sharp

and co-workers have demonstrated that accurate measures of SASA for residues with poor

intrinsic reactivity (e.g. Arg, Lys, Val, Thr, Ser, Pro, Glu, Gln, Asn, Asp, Ala) during

hydroxyl radical CL experiments require the normalization of sequence effects by

comparing label profiles of native and denatured structures.[65] For highly reactive residues

(e.g. Trp, Tyr, Phe, His, Ile), the effects of primary structure on reactivity are relatively

minimal.[65] In addition, the microenvironment (**Figure 1.3**) caused by the tertiary structure

around an amino acid residue can influence the acid/base characteristics of side chains

because of charge-charge interactions, charge-dipole interactions, dipole-dipole

interactions, and/or hydrophobic effects.[68-71] For example, in dimethyl labeling proteins,

Wang and co-workers found that hydrogen bonding and electrostatic interactions around

lysine residues can influence the methylation reactivity of lysine residues.[72] The effects of

local structural contacts of amino acid side chains, especially those with lower intrinsic

reactivity, have been found to contribute to their reactivity with hydroxyl radicals.[64, 65] In

CL with carbenes, the surface bias caused by the reagent's affinity for specific protein

surface sites might be explained by primary and/or tertiary effects. Sequence context and

non-covalent molecular interactions (higher-order structures) may drive and steer the selectivity of diazirine precursors toward certain residues.[50, 52, 53] However, no definitive experiments have been performed to demonstrate this hypothesis yet. For reagents like DEPC that are inherently less reactive than hydroxyl radicals and carbenes, the effect of primary or tertiary structure might be expected to be more pronounced since changes in the pK_a values of a given residue lead to changes in its protonation state and thus nucleophilicity. The relatively poor correlation observed between SASA and the DEPC reactivity of Ser and Thr residues[58, 59, 73] might imply that the microenvironment around these amino acids, as formed by a protein's tertiary structure, tunes their reactivity.

1.3 Using Covalent Labeling with Mass Spectrometry Detection

1.3.1 Principle of Experiment Design

During CL experiments only solvent exposed residues are modified and thus provide direct structural information, while the buried residues are unmodified and thus indirectly provide structural information. As illustrated in **Figure 1.1**, any structural information obtained from a CL experiment comes from comparing the labeling ratio of the protein reacted under at least two conditions, with one of the conditions usually being the protein in its native state. The resulting differential reactivity is used to deduce structural information. Another important principle of CL experiments is that the labeling reagent must be used in such a way as to not perturb a protein's HOS. Clearly, if the probe itself changes the protein's structure, then the resulting data will not correctly report on the protein's structure.

There are at least three general ways that are used to assess that a protein's structure is not perturbed during CL. The first is to use a complementary measurement such as CD, fluorescence, or an activity assay to monitor if the protein's structure has changed during labeling.[52, 58, 73, 74] Our group has demonstrated in previous work, that while these methods are easy to implement, they tend to indicate global structural changes and lack the resolution to identify local protein structural changes.[58] A comparable approach is to measure the charge-state distributions of a protein via electrospray ionization (ESI)-MS[75] as a protein will undergo a shift in its charge-state distribution upon a structural change. However, like with CD and fluorescence spectroscopy, changes in ESI-MS charge-state distributions are typically not sensitive enough to monitor local changes in HOS, so it is debatable whether these approaches are sufficient for assessing if a protein's structure is perturbed or not from the labeling reaction.

A second commonly used approach for assessing a protein's structural integrity during CL is to measure labeling reaction kinetics (**Figure 1.4**). Chance and co-workers pioneered this approach with synchrotron-based hydroxyl radical labeling in which the unmodified fraction of a peptide or protein is monitored as a function of the reagent dose (e.g. radiolysis time). In this way, dose-response plots can be generated, and adherence to the proper reaction order kinetics indicates whether a protein's structure has been perturbed or not.[76, 77] These plots can be generated for all modified peptides that are produced from a proteolytic digestion and provide a measure of the structural integrity throughout the entire protein. Our group has successfully used an analogous approach with DEPC labeling[58, 59, 63, 78, 79], which follows second order kinetics, and we have found this approach

to offer much higher sensitivity than CD or fluorescence spectroscopy for ensuring

structural integrity through all regions of a protein.
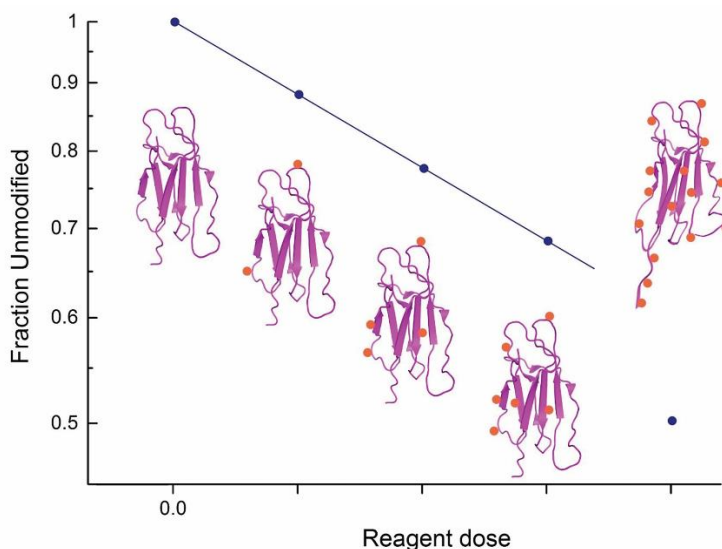
Figure 1.4: Hypothetical dose-response plot, showing first-order labeling reaction kinetics.
When the extent of labeling is low enough to prevent structural perturbation, reaction kinetics are monotonic (i.e. four lowest reagent doses in the plot). When over-labeling perturbs the structure of the protein, the reaction kinetics change, indicating variations in the microenvironment around one or more residue.

A third method of guaranteeing a protein's HOS is not been perturbed by CL is to

ensure labeling chemistry happens faster than structural changes can occur. As indicated

in previous sections (see **Section 1.2.1**), FPOP-based hydroxyl radical labeling and carbene

labeling have extremely fast labeling rates that are comparable to or faster than protein

folding rates.[23-25, 44, 50] The fastest protein structural changes occur in a few µsec,[20, 80] and

thus if hydroxyl radical or carbene labeling occurs faster, then protein modification should

take place before any significant protein structural changes occur. In the FPOP

experimental setup, hydroxyl radicals are generated during the nsec pulse of the laser,[23]

and radical scavengers that are present in solution limit the overall lifetime of the hydroxyl

radicals. Calculated lifetime profiles[23] and time-resolved UV spectroscopy[25] suggest that

reactive hydroxyl radical species are consumed in 0.1-1 µsec when radical scavengers are present. Such short lifetimes allow the radical-protein reactions to happen faster than protein structural changes. Further proof that proteins are modified before they can undergo structural changes comes from the Poisson distribution of modification extents, which suggest that only a single protein species is reacting.[24] Despite evidence that hydroxyl radical labeling by FPOP occurs on the µsec timescale, recent work by Konermann and coworkers suggests that the time window for FPOP-based labeling may be much longer than one millisecond as secondary and higher order radicals have longer lifetimes and are possibly not effectively quenched.[41] At this point, however, it is unclear the extent to which secondary radicals are responsible for the measured protein oxidation.

It should be noted that there are numerous examples in the literature in which nothing is done to ensure a protein's structural integrity during CL. This approach can be successful if "single-hit" conditions are achieved.[81] In other words, as long as there is one label per protein on average, the labeling chemistry should have successfully probed the original protein structure. Any additional label, by definition, would be probing a possibly perturbed structure because the protein would now have a covalent modification that it did not have originally. While this has commonly been acknowledged, recent work by Madsen *et al*[82] on monoclonal antibodies (mAbs) suggests that multiple labels can be accommodated while still obtaining the desired information. It is unclear at this point how generalizable this observation is.

1.3.2 Experiment Design and Workflow

The goal of a CL experiment is to encode structural information into the mass of the protein. The commonly used workflow for CL experiments is sample preparation,

labeling reaction, and, if necessary, quenching to consume excess reagent. After the CL experiment, the labeled protein sample can be directly analyzed by MS or LC-MS, or more typically, the labeled protein can be proteolytically digested prior to LC-MS (**Figure 1.5**). While standard proteomics techniques are typically used to analyze covalently labeled proteins to obtain the desired structural information, the following unique aspects of the CL approach must be taken into consideration.
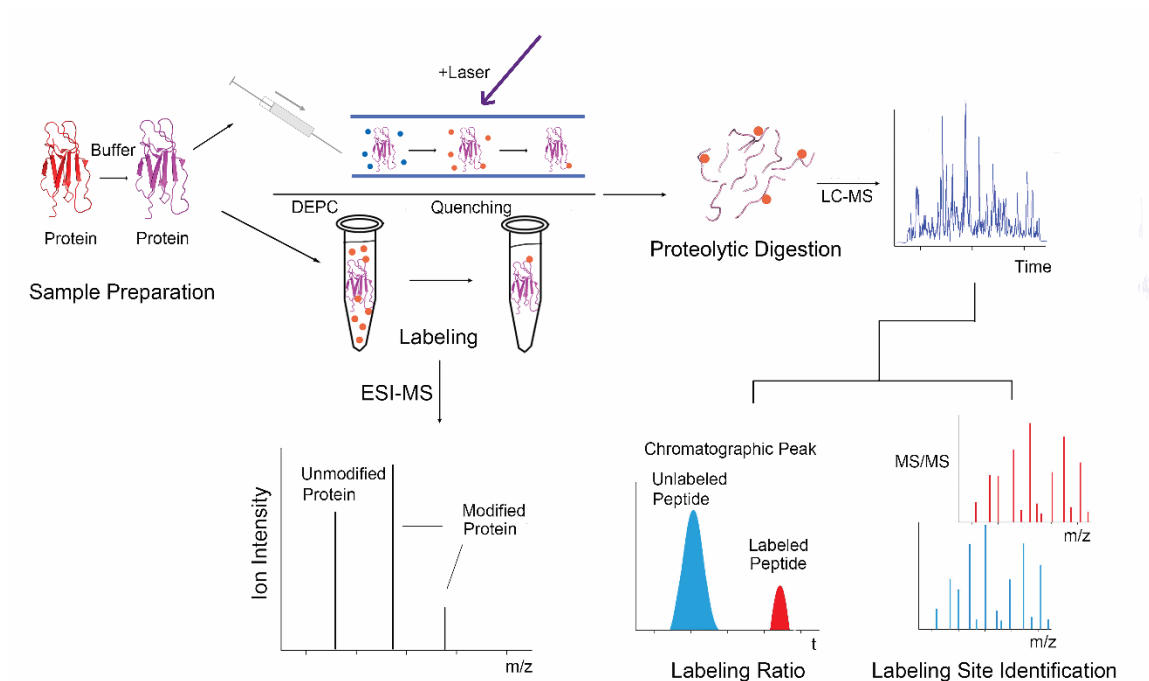
Figure 1.5: Example workflow for covalent labeling combined MS.
Before MS analysis, a protein sample is prepared and labeled, during which the structural information is encoded into the mass of the protein. The labeled intact protein sample can be analyzed by ESI-MS to monitor the overall extent of modification, or more typically the labeled protein is subjected to "bottom-up" analysis via proteolytic digestion and LC/MS/MS to determine labeling ratios at each modified residue.

Unlike in typical proteomics experiments, CL-MS experiments require complete or near complete protein sequence coverage to obtain the desired structural information. Moreover, a semi-quantitative measure of labeling extents at residues throughout the protein is needed to deduce structural information in a differential experiment (see **Figure**

21

1.1). Like in proteomics studies, the two primary strategies for determining labeling sites are "top-down" sequencing[83-85] and "bottom-up" sequencing.[83, 86, 87] In top-down sequencing, the entire, intact protein is analyzed with MS/MS, and therefore complete sequence coverage is inherently available. In top-down sequencing, it is possible to select and apply MS/MS to the protein with only one label. This is beneficial if there is a concern that multiple labeling events could have perturbed the structure ("single-hit" principle). Despite these advantages, very limited work has been published using top-down MS to analyze proteins covalently labeled by non-specific reagents. The limited utility of top-down sequencing for CL-MS is mostly due to the relatively poor ability of top-down sequencing to localize modification sites, especially for large protein systems. Often large fragment ions containing multiple possible modification sites are produced, making it difficult to assign labeling sites to specific residues with high confidence. This difficulty effectively results in low structural resolution. In addition, semi-quantitative information that is usually required in CL-MS experiments is challenging to obtain with top-down sequencing. One study by Gross *et al*. compared top-down and bottom-up sequencing for analyzing covalently labeled proteins and concluded that top-down sequencing had limited sensitivity, a low dynamic range, and sometimes difficulty pinpointing modified residues in regions with numerous modified sites.[88] It should be noted that the utility of top-down sequencing for CL-MS experiments might improve with the use of UV photodissociation, as this dissociation technique has shown utility for identifying modified sites during amino-acid specific labeling experiments.[89]

Bottom-up sequencing is a more commonly used strategy to analyze covalently labeled proteins. Because MS/MS is more effective on smaller peptides, this strategy

provides high spatial resolution, usually pinpointing labeled sites down to 1 or 2 amino acids. Moreover, modification extents can usually be determined as low as 0.1% or lower. Bottom-up sequencing of covalently labeled proteins still suffers from the usual challenges associated with bottom-up sequencing, such as biases associated with proteolytic cleavages and peptides going undetected during LC/MS experiments. For large proteins, where sequence coverage during bottom-up sequencing can be low, using multiple enzymes for proteolytic digestions can be used to improve coverage.[90]

An important consideration associated with bottom-up sequencing is the choice of dissociation techniques during the MS/MS stage of the experiment. Collision-induced dissociation (CID) remains the most commonly used dissociation technique and provides accurate information in most cases. There are examples, however, in which CID has difficulty with certain types of modifications and modified amino acid residues. For example, CID provides very poor sequence coverage for peptides containing oxidized cysteine or methionine residues.[91] The dissociation technique has also been known to lead to label misassignments for peptides in which histidines are oxidized, especially when more than one modification site exists in a peptide, as oxidation causes new dissociation pathways to emerge.[92] Sharp and co-workers reported a similar conclusion based on studies of several oxidatively modified peptides.[93] Significant neutral losses are also observed in CID of peptides modified by carbenes, making it difficult to pinpoint modification sites.[50] Another issue observed with a subset of DEPC-labeled peptides is the possibility of the label to "scramble" from one site on a peptide to another during CID, presumably due to an intramolecular nucleophilic attack during the CID process.[94] In each of the cases in which CID provided incomplete or misleading information about modification sites,

electron transfer dissociation (ETD) was successfully used to correct the problem. In fact, it has been argued that ETD generally provides improved identification and quantitation of labeled sites produced during oxidative labeling experiments.[50, 94, 95]

1.3.3 Data Analysis and Presentation

The desired results from CL-MS experiments are measurable labeling levels at residues throughout the protein so that structural information can be obtained. The more labeled sites that are measured, the higher the structural resolution will be from the CL-MS experiment. In addition, some semi-quantitative measure of the labeling extent at each residue is desired so that any corresponding structural change can be deduced. It is rare that a particular residue is labeled in 100% of the protein molecules. More commonly, only a fraction of the protein molecules in a sample are labeled at a given site, and this labeling extent is usually provided as a labeling percentage from LC-MS measurements. Decreases in labeling percentages at given residues typically indicate a drop in solvent accessibility that could, for example, indicate a protein-protein interaction site (**Figure 1.6**). In contrast, increases in labeling percentages at a given site would suggest unfolding in that region of the protein.

While residue labeling identification is achieved using tandem MS, the labeling extent is usually determined by the ion abundance or chromatographic peak area of the labeled peptide that contains the modified residue. In addition to the use of standard software packages for analyzing such data (e.g. Mascot[96, 97]), other programs such as ByOnic,[51] ProtMapMS,[98] a custom-designed version of Proteome Discoverer,[99] and other specialized software pipelines[73] have been created to facilitate CL-MS data analysis. A key difference between these specialized software and traditional proteomics software is that

they are much more efficient at finding labeling sites, especially at low levels, while also providing a measure of the labeling percentage at every identified site.

While it is important to obtain a measure of labeling extent, it should be noted that explicit quantitation is not attainable through most CL-MS experiments. There are many reasons for this fact. First off, modified and unmodified peptides are distinct chemical species that have different inherent ionization efficiencies. Moreover, a difference in chromatographic retention times of modified and unmodified peptides during reversed-phase LC gradient elutions leads to a different percentage of organic solvents from which each peptide is electrosprayed, resulting in further differences in ionization efficiencies. These ionization efficiency differences cause there to be a distinct relationship between solution concentration and ion abundance for modified and unmodified peptides. To remediate the ionization efficiency bias that occurs during LC-MS analyses, Sharp and co-workers have employed size-exclusion chromatography (SEC) instead of reversed-phase LC to separate labeled peptides. SEC ensures all peptides are electrosprayed under common solvent conditions, thereby eliminating one cause of ionization efficiency differences. A drawback of this approach is that isomeric labeled peptides co-elute, making it more difficult to quantify labeling extents for peptides with more than one modified residue. To some degree, this drawback can be overcome using ETD, as tandem mass spectra from ETD can often correctly report modification levels of each residue in an isomeric mixture.[93]
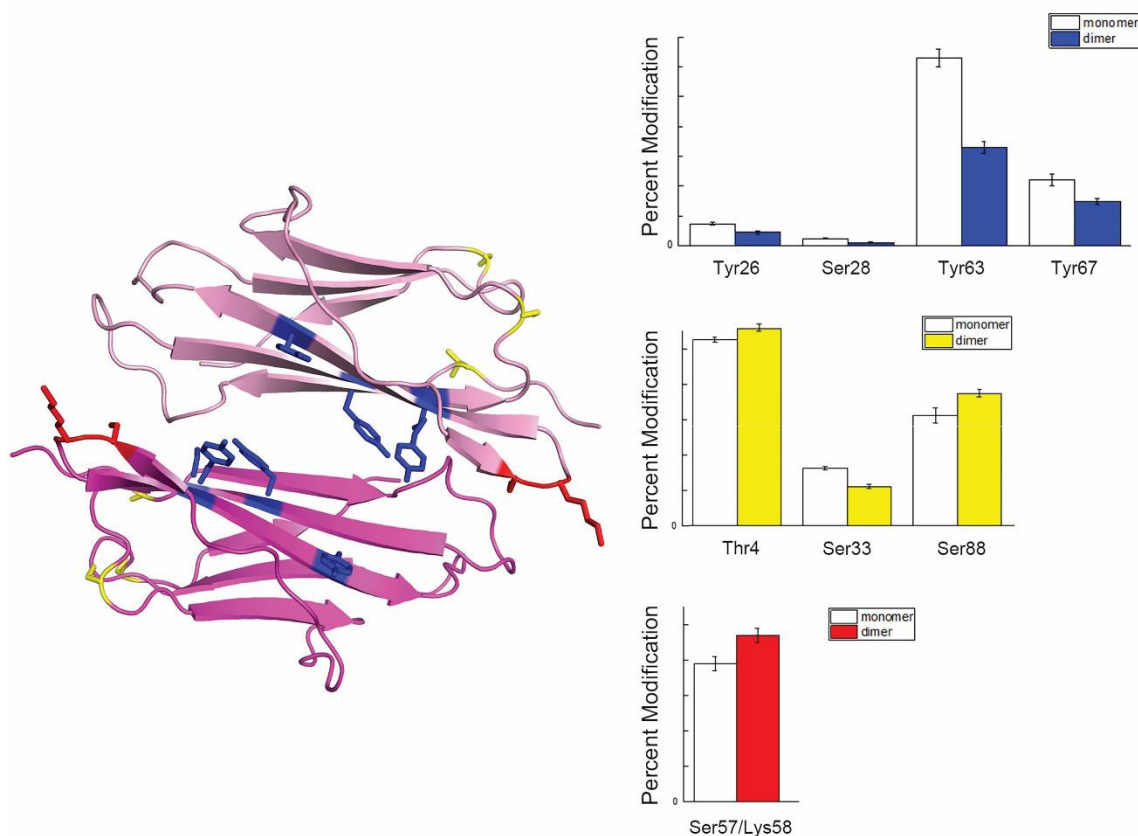
Figure 1.6: An example of how changes in labeling level can be applied to predict the protein dimer interface demonstrated with β2m dimer structure extracted from β2m H13F hexamer (PDB 3CIQ).
Blue colored residues indicate significant decreases in labeling level from monomer to dimer, which helps locate the dimer interface. Yellow colored residues show no significant change in labeling level, which indicates these residues are at a region away from dimer interface. The red colored residues are examples of residues with increases in labeling levels, indicating these residues becomes more solvent exposed upon dimer formation.

1.4 Applications of Covalent Labeling – Mass Spectrometry Methods

The use of CL-MS methods to obtain structural information about protein topology has gained popularity because of its moderate to high structural resolution, straightforward implementation, and/or ability to probe structural changes on a µsec time scale. In this section, we describe some select examples of CL-MS applications.

1.4.1 Protein Folding/Unfolding

In most proteins, specific biological functions can be exerted only after folding into their appropriately folded forms. Deciphering protein folding/unfolding is necessary to understand protein-energy landscapes, the biological processes of protein misfolding disorders,[100-102] and is essential for the development of therapeutics against these disorders.[101] While numerous techniques, including other MS-based techniques, have been developed to monitor conformation dynamics during protein folding/unfolding,[3, 103] protein surface mapping with CL-MS can provide information with high spatial and temporal resolution, in a way that many other techniques cannot. Laser-based CL methods, such as FPOP, have proven to be useful tools for protein folding studies because proteins are labeled on the µsec time scale, which is faster than protein unfolding, due to the pulsed nature of reagent production and the rapid reaction quenching.[23, 24] Using FPOP together with denaturants such as temperature and pH, sub-msec protein folding experiments can be performed. For example, Gross and co-workers used temperature jump experiments to study the region-specific folding of the protein barstar. Barstar in aqueous buffer undergoes reversible unfolding at low temperature ($\sim 0\ ^\circ C$)[104], and 'pump/probe' experiments can be conducted to study the protein's refolding (**Figure 1.7a**). Refolding of unfolded barstar was created by rapid IR laser-induced heating of a solution containing the protein ('pump' from $\sim 0\ ^\circ C$ to $\sim$ room temperature), and then FPOP with a 'probe' UV laser was used to oxidize the protein at different times after it began refolding. By limiting the lifetime of the hydroxyl radicals in solution to around 1 µsec, 'snapshots' of the protein as it folded could be obtained.[105-107] Their results confirmed the presence of partially-folded intermediate conformations of barstar during its refolding (possibly a molten globule state), which was

consistent with previous UV absorbance, fluorescence, and CD spectroscopic studies.[108-111] The CL-MS experiments also revealed residue-level information about conformational changes during barstar folding, which was not available from previous spectroscopic measurements. In a similar manner, the same group studied the refolding kinetics of the paramyxovirus fusion protein, which is essential in mediation of virus-host cell fusion and genetic transfer.[112] Besides temperature-triggered folding experiments, Konermann and co-workers established a platform to study short-lived protein folding/unfolding intermediates using continuous-flow mixing and pulsed laser-induced oxidative labeling.[113-116] To examine protein unfolding, a protein of interest is premixed with H_2O_2 in a reaction capillary and unfolding is initiated by rapid mixing with acid or denaturant.[113, 115] To study protein folding, a denatured protein is prepared by acid or denaturant, and the folding is then initiated via a pH jump or rapid denaturant dilution.[114, 116] To study conformations at different time points after initiating unfolding or folding, a pulsed laser beam is focused at different positions along the mixing capillary, and a certain amount of the resulting oxidized sample is collected for further bottom-up LC-MS/MS analysis (**Figure 1.7b**). Folding/unfolding intermediates of bacteriorhodopsin, apo- and holo-myoglobin have been studied by this method, and the in-depth residue-level information from the rapid pulsed CL, in combination with the global structural features obtained from optical spectroscopy, can provide good spatial and temporal resolution of protein folding/unfolding pathways.[113-116]
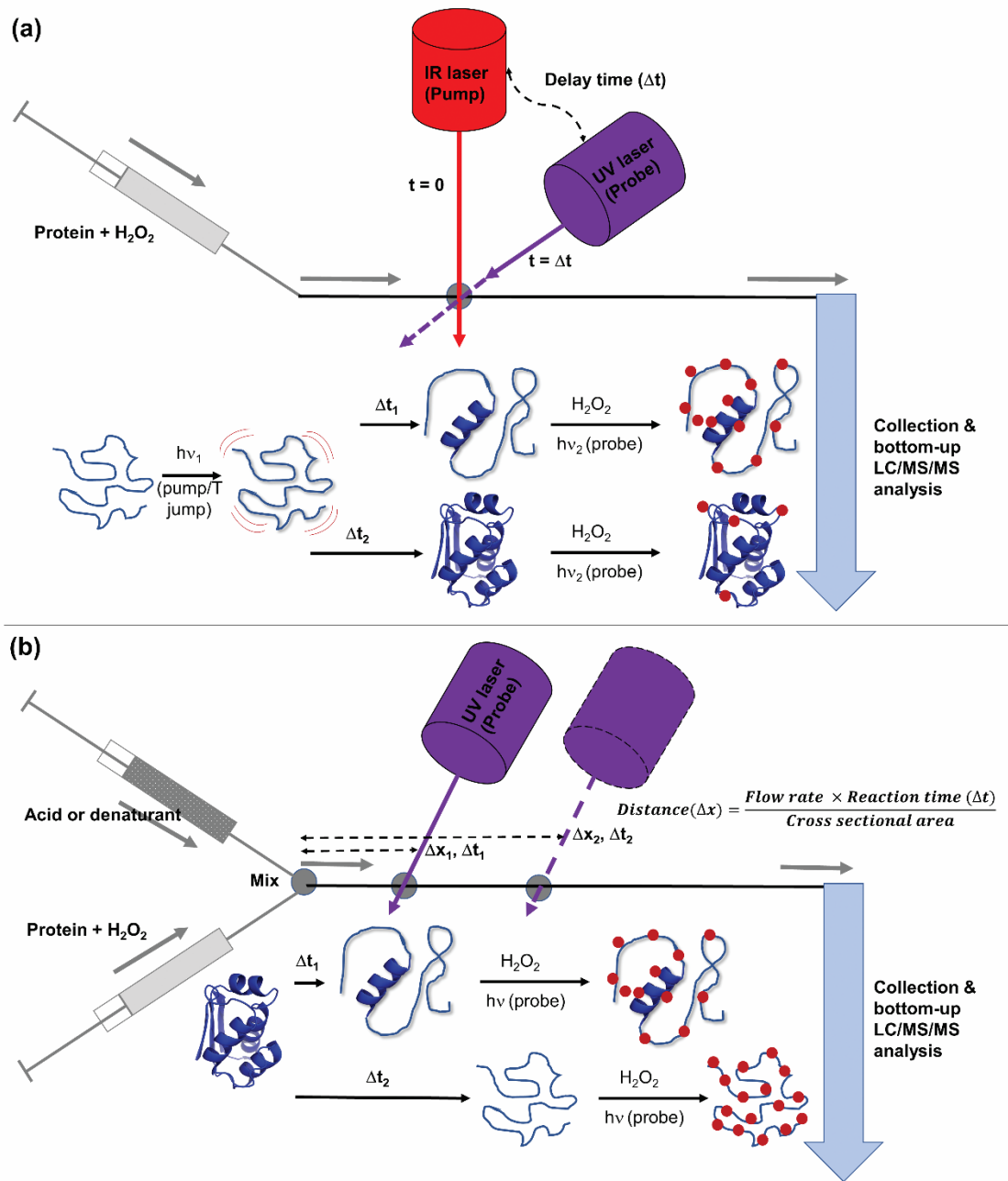
Figure 1.7: An illustration of the experimental setups used for studying protein folding/unfolding using laser-based CL methods.
Schematic diagrams are shown for (a) temperature jump (T jump) and (b) continuous-flow mixing experiments coupled with pulsed laser-induced oxidative labeling. Details about each experiment are provided in the text.

1.4.2 Amyloid-Forming Proteins

Amyloids are insoluble protein aggregates that are associated with several human diseases, including Alzheimer's and Parkinson's.[117] Numerous efforts have been devoted to deciphering the aggregation process *in vitro*, but this information is challenging to obtain because amyloid proteins usually aggregate rapidly.[118] Moreover, pre-amyloid oligomers, which may in some cases be disease causing[119-121] are short-lived and can be difficult to structurally interrogate.[81, 122, 123] As compared to NMR and X-ray crystallography, CL can provide better time resolution for studying transient pre-amyloid structures, and various CL approaches have been used to gain insight into the aggregation interfaces for several amyloid proteins. This structural information can be obtained by comparing the CL ratio of residues before and during oligomer formation to help identify aggregation interfaces. As an example, Chance and coworkers applied synchrotron-based hydroxyl radical CL to Aβ40, which is the protein that aggregates in Alzheimer's disease.[124] Protection factors were used to indicate the degree of change in solvent accessibility. Structural information of full-length Aβ40 monomer in fibrils and the core structure of two- and three-filament conformers were acquired and were compared to the previously reported structures by solid state NMR. Gross and co-workers applied FPOP in a time-dependent manner to study the pre-amyloid formation process of Aβ1-42.[125] By plotting the fraction of modified intact Aβ1-42 as a function of time, a kinetic scheme of the aggregating process was revealed. Similarly, by plotting the labeling percentages of modified peptides over time, the regions with significantly decreased labeling ratios indicated the aggregating core structure. Moreover, because MS/MS can indicate the exact labeled residues, even more detailed structural information could be obtained. DEPC labeling has also been successfully used

to study the pre-amyloid oligomers of β-2-microglobulin (β2m), which is the protein that forms amyloids in dialysis-related amyloidosis.[126] By conducting DEPC labeling at different time points after initiating amyloid formation with Cu(II), the structural changes to the monomer caused by Cu(II) binding,[63] as well as structural models for the pre-amyloid dimer[78] and tetramer[79] could be determined.

1.4.3 Protein-Ligand Systems

CL is also well suited for studying protein-ligand complex binding sites when traditional methods, such as NMR and X-ray crystallography, are too time-consuming or fail due to limited sample amounts, protein instability, sample heterogeneity, protein molecular weight limitations, or difficulties with crystallization. In its most straightforward implementation, CL-MS can report on ligand binding sites via decreases in labeling extents at specific residues in the protein. Ligand binding protects residues on the protein surface from being labeled. Many examples of CL used for protein-ligand binding studies are available.[81, 95, 127-132] One example by Manzi *et al*. demonstrated that carbene-based labeling is capable of reporting the binding site of the lysozyme-NAG complex.[52] Multiple residues showed a significant decrease in labeling extent in the presence of the ligand, clearly correlating those residues to the well-studied binding pocket of the complex. Similarly, our group has applied DEPC labeling together with two other site-specific labeling reagents to study β2m binding to small molecules that inhibit its amyloid forming. The simplicity and speed of the labeling chemistry allows for ligand binding site information on a fast aggregating protein whose ligand binding site could not otherwise be determined.[81] Combining CL-MS with computational docking further refines the structural information that is acquired. CL-MS can test the pose predicted by docking[81, 131, 132] or

restrain the region of protein to which the ligand is bound, but CL-MS is usually unable to show the orientation of the ligand, which is information that computational docking can provide. In one example reported by Chance and co-workers,[131] a hybrid structural approach that include synchrotron-based hydroxyl radical CL, homology modeling, computational modeling, and protein ligand docking was used to predict the structure of the membrane protein Serotonin Type 4 Receptor (5-HT4R) and bound ligands. In addition to providing protein-ligand binding site information, CL has been shown to provide information about conformational changes upon ligand binding.[95, 129, 130] Gross and co-workers applied FPOP to study heme binding and conformational changes in myoglobin.[129] Several residues that are close to the heme binding site show significant differences in their ability to be modified in apo or holo form of myoglobin. In another case, the structural changes caused by the binding of the peptides melittin and mastoparan to calmodulin were compared to the known structure of the calmodulin-M13 complex.[130] The CL results indicated statistically similar patterns in labeling extents, leading to the conclusion that ligand binding changed the structure of calmodulin but the ligand bound calmodulin retained a similar structure when bound to M13, melittin, or mastoparan. It should be noted that distinguishing between ligand binding and ligand-induced conformational changes must be done carefully. For protein-ligand complexes with unknown binding sites, a decrease in labeling of a given residue (or a set of residues) might be due to protection from ligand binding or it could be due to allosteric changes that decrease the solvent accessibility of residues not directly involved in ligand binding. Further application of CL to protein-ligand binding should reveal whether such allosteric changes are likely to significantly change labeling patterns.

1.4.4 Covalent Labeling with Computational Modeling

The structural resolution provided by CL-MS is lower than that provided by NMR or X-ray crystallography because only information about the solvent accessible side chains is obtained, and this information is only a fraction of the amino acids in a protein. Even with relatively limited information, CL-MS has been successfully used to confirm or rule out proposed structural models.[133, 134] CL-MS has the potential to provide even higher resolution structural information when it is combined with computational modeling. Several groups have explored the combination of CL-MS and computational modeling techniques. For example, Gerega and Downard developed an algorithm, known as PROXIMO, that is specially designed to model protein-peptide complex structures based on oxidative labeling data from radical probe MS.[135] Similarly, Chance and co-workers have developed the ClusPro program for predicting protein complexes using the restraints from hydroxyl radical labeling to facilitate homology modeling and associated protein structural prediction for protein complexes with no NMR or X-ray crystal structure.[136] In addition, Sharp and co-workers have developed a strategy that correlates oxidative labeling levels with SASA at each modified residue, which can be utilized to assess the quality of protein structural models[65] or model protein complex structures.[137] While each of these above approaches have helped refine the protein structural information obtained by CL-MS, they have not been adopted broadly. Recent work has shown the promise of combining multiple complementary experimental methods to improve structural modeling. For example, methods like limited proteolysis,[138] small-angle X-ray scattering,[139] chemical crosslinking combined with native MS[140] have been combined with CL-MS and computational methods to arrive at protein structural models. It appears that CL-MS

research is in the very early stages of identifying the best methods to efficiently combine constraints from CL-MS, complementary methods, and computational modeling to obtain higher resolution protein structural information.

1.4.5 Membrane Protein Footprinting

Another promising area in which CL-MS has the potential to have an impact is in the study of membrane proteins. Membrane proteins are inherently more difficult to study as compared to water-soluble proteins. CL-MS would appear to be a promising tool for structural analysis of membrane proteins because CL can often be done under conditions that preserve conformation and topologic information. Several groups have demonstrated the promise of CL in studying membrane proteins.[55, 141-147] Hydroxyl radical labeling has been shown to map the topology of intrinsic membrane proteins in their natural lipid environment,[141] in nanodiscs,[145] within amphipols,[146] and even within living cells.[147] New hydrophobic-based CL labeling reagents, e.g. based on triflinate,[55] might be able to further improve the information accessible by CL-MS by allowing the labeling of protein regions that are inserted into membranes themselves.

1.4.6 In-Cell/In-Organism Covalent Labeling

Beyond live-cell membrane protein CL,[147] Jones and co-workers recently performed in-cell FPOP of live Vero cells. They demonstrated that proteins in different cellular compartments can be labeled and residue-level oxidative modification extents can be measured, and these labeling extents correlate well with the SASA of the residues [148]. It is reasonable to assume that this application of CL-MS could be expanded for studying protein interactions and conformational changes in living cells.

1.5 Structural Analysis of Protein Therapeutics Using CL-MS

Protein therapeutics are the fastest growing pharmaceuticals on the market and have important roles in almost every field of medicine.[149, 150] In 2016, more than 50% of new drug approvals were biologics,[151] of which the vast majority are protein biopharmaceuticals, and this number is estimated to rise to around 70% in 2025.[152] Protein therapeutics are classified as enzymes, hormones, growth factors, fusion proteins, cytokines, anticoagulants, antibodies, antibody-drug conjugates, protein vaccines, and protein diagnostics.[153, 154] Monoclonal antibodies (mAbs) belong to the fastest growing segment.[155, 156] In contrast to small molecule drugs, the three-dimensional structure of therapeutic proteins contributes to the higher binding specificity towards drug targets, leading to greater therapeutic efficacy and less side effects.[153] Growing concerns on the quality, efficacy, and safety of protein biopharmaceuticals have led the developments of many analytical methods to investigate quality attributes of therapeutic proteins at different levels, from primary structure (e.g. amino acid sequence and post-translational modifications) to higher-order structure (e.g. protein conformation) [**Figure 1.8**].[152, 156, 157] During manufacturing, biopharmaceutical products must be well characterized and lot-by-lot comparable. In the case of biosimilars (i.e. follow-on versions), in addition to their lot-by-lot comparability, their quality attributes must be proven to be comparable to a reference product (i.e. an innovator product).[152, 156]
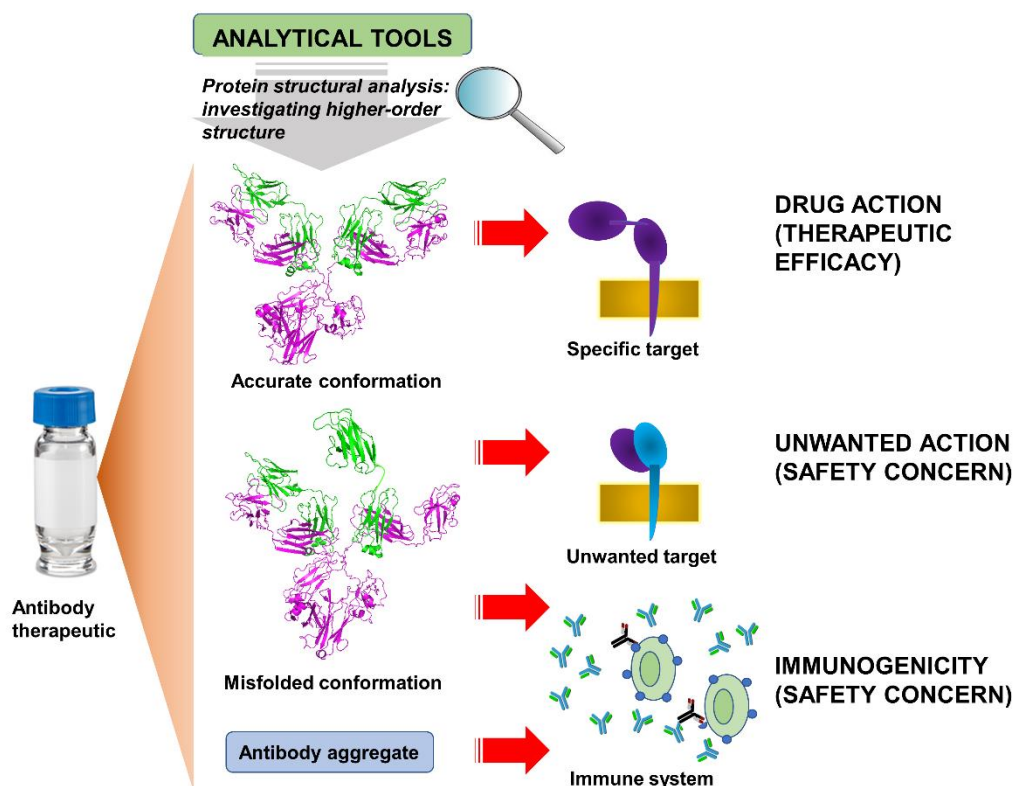
Figure 1.8: Higher-order structure is one of quality attributes of therapeutic proteins that require proper analytical tools to characterize in order to ensure quality, efficacy, and safety of biopharmaceutical products.

Changes in HOS, e.g. protein misfolding and aggregation, can reduce drug efficacy, enhance unwanted drug actions, and/or induce immunogenicity.[158, 159] Reliable analytical tools are currently needed for characterizing the HOS of biopharmaceuticals to ensure the quality, efficacy and safety throughout a product life cycle – from manufacturing to dose administration.[158, 160-163] Several low-resolution spectroscopic methods are still commonly used to characterize HOS,[158, 164] but information about small localized conformation changes, some of which may potentially affect drug efficacy and safety, cannot be provided. Hence, there is a growing need for novel analytical tools with good sensitivity and moderate structural resolution to characterize HOS at the amino acid level.

CL-MS techniques have the ability to probe the structure of protein therapeutics with moderate structural resolution and could be used in a semi-high throughput manner to monitor numerous therapeutic preparations.[73, 165, 166] FPOP along with MS detection has been shown by Watson and Sharp to be sensitive enough to detect subtle structural changes of expired and mishandled granulocyte colony-stimulating factor (GCSF) samples.[165] Deperalta *et al.* demonstrated the ability of hydroxyl radical labeling to characterize the dimer interface of an antibody dimer commonly found in formulated therapeutic IgG1 mAbs. Hydroxyl radical labeling data obtained using synchrotron radiolysis of water has shown that the F_{ab} domain displays decreased oxidation rates and is likely involved in dimer interface.[166] Our group has also demonstrated that DEPC labeling along with MS can identify minor conformation changes in thermally-stressed human growth hormone (hGH) and IgG1 mAb samples. Heat-stressed hGH undergoes very subtle structural change that other biophysical techniques (CD, fluorescence spectroscopy, dynamic light scattering) are not sensitive enough to detect, but DEPC labeling along with MS detection can detect with good structural resolution. We have also found that DEPC labeling can reveal the aggregation interface in a heat-stressed IgG1 mAb sample.[73] Epitope/paratope mapping of the specific residues involved in antibody-antigen and other protein-receptor interactions is particularly valuable for assessing epitope novelty, predicting immunogenicity, evaluating binding characteristics, and for optimizing binding.[167] Recently, Gross and co-workers used FPOP techniques to identify the interface of anti-thrombin mAb/thrombin,[96] anti-vascular endothelial growth factor (VEGF) F_{ab}-1 fragment/VEGF,[97] anti-interleukin-23 (IL-23) mAb/IL-23,[90] and IL-6/IL-6 receptor interactions.[168]

1.6 Summary

In this chapter, we have provided an overview of CL-MS methods that use reagents that react with a wide range of amino acid residues, including hydroxyl and trifluoromethyl radicals, carbenes, and DEPC. These broadly reactive reagents are making CL-MS methods of increasing interest for the structural analysis of macromolecules. CL-MS fills in the gap among current structural characterization techniques, with higher resolution than CD and fluorescence spectroscopies, and faster measurement times than NMR and X-ray crystallography. At the same time, it retains the limited sample consumption common to most MS-based methods and can be applied to protein systems that are difficult to study by other techniques. CL-MS has very limited label loss and label scrambling, allowing greater sensitivity to structural changes in some cases and better reproducibility in most cases than HDX-MS. The fact that only solvent accessible amino acid resides are modified makes the technique particularly well suited for studying protein interactions with ligands and other proteins, including aggregating proteins. Radical and carbene-based labeling can monitor protein structural changes on the μsec timescale, allowing the study of protein folding and unfolding.

While significant progress has been made with CL methods to address protein structure, some technique limitations remain. Unlike most spectroscopic methods, CL-MS relies on changing a protein in order to be probe its structure. Covalent modification of amino acid residues is needed to encode a protein's structural properties into its measured mass, and this necessarily changes the protein. Thus, every labeling procedure has to be carefully controlled and understood to ensure that accurate structural information is obtained. Our understanding of the underlying chemistry in most cases is still incomplete,

and this lack of understanding can give rise to ambiguous structural information in some cases. Another limitation of CL is that not every amino acid in a protein can be probed, and this inherently limits the structural resolution obtainable by this approach. No one CL reagent has been found to react with every type of amino acid side chain. A related issue is that every amino acid residue reacts to a different extent with a given CL reagent. This fact complicates attempts to find correlations between reactivity and higher order structure. Despite these limitations, and in some cases because of these limitations, several opportunities still exist for the future development of CL-MS.

This dissertation focuses on the use of DEPC-based CL-MS techniques to address the need in pharmaceutical industry, specifically the protein therapeutics field. Characterizing HOS of mAbs is challenging given their size and the multidomain nature. As one of the non-specific reagents, DEPC can modify a wide range of amino acid side chains simultaneously, providing good structural resolution to probe site-specific HOS changes of therapeutic mAbs.

In **Chapter 2**, we demonstrated the effect of residue microenvironment on DEPC labeling reactivity. We find that in intact proteins weakly nucleophilic side chains (Ser, Thr, and Tyr) can be modified by DEPC in addition to other residues such as His and Lys, which allows for higher structural resolution to obtain insight into conformational changes of therapeutic proteins. From a close examination of the structural features near the reactive residues, we find that nearby hydrophobic residues are essential to the enhanced reactivity of certain Ser, Thr, and Tyr residues.

Chapter 3 describes the use of DEPC-based CL-MS to detect conformational changes of a therapeutic mAb caused by heat stress. DEPC directly probes solvent

39

accessibility and the microenvironment of side chains, and any changes in DEPC reactivity are indicative of changes in protein conformation. Using rituximab as a model mAb therapeutic, we demonstrate the ability to site-specifically detect subtle HOS changes at the temperatures far below the melting point of the mAb therapeutic, changes that are not detected by common characterization techniques. At higher stress temperature, conformational changes and aggregation sites are also identified.

In **Chapter 4**, we evaluated applicability of DEPC reaction kinetics to ensure site-specific structural integrity of mAbs during labeling as reliable information about HOS can be obtained only when the protein's structural integrity is preserved. We find that multi-domain mAbs can withstand many more than one label, which contrasts to previously studied single-domain proteins. This more extensive labeling provides a more sensitive measure of structure, making DEPC-based CL-MS suitable for the HOS analyses of mAbs.

Finally, **Chapter 5** contains conclusions and future directions of this research.

1.7 References

1.   Katta, V.; Chait, B. T.; Carr, S., Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry. *Rapid Communications in Mass Spectrometry* **1991,** *5* (4), 214-217.
2.   Engen, J. R.; Smith, D. L., Peer Reviewed: Investigating Protein Structure and Dynamics by Hydrogen Exchange MS. *Analytical Chemistry* **2001,** *73* (9), 256 A-265 A.
3.   Kaltashov, I. A.; Eyles, S. J., Studies of biomolecular conformations and conformational dynamics by mass spectrometry. *Mass Spectrometry Reviews* **2002,** *21* (1), 37-71.
4.   Wales, T. E.; Engen, J. R., Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrometry Reviews* **2006,** *25* (1), 158-170.
5.   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.
6.   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.

7.   Pirrone, G. F.; Iacob, R. E.; Engen, J. R., Applications of Hydrogen/Deuterium Exchange MS from 2012 to 2014. *Analytical Chemistry* **2015,** *87* (1), 99-118.

8.   Sinz, A., Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein–protein interactions. *Mass Spectrometry Reviews* **2006,** *25* (4), 663-682.

9.   Holding, A. N., XL-MS: Protein cross-linking coupled with mass spectrometry. *Methods* **2015,** *89*, 54-63.

10.   Yu, C.; Huang, L., Cross-Linking Mass Spectrometry: An Emerging Technology for Interactomics and Structural Biology. *Analytical Chemistry* **2018,** *90* (1), 144-165.

11.   Xu, G.; Chance, M. R., Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics. *Chem. Rev.* **2007,** *107* (8), 3514-3543.

12.   Mendoza, V. L.; Vachet, R. W., Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spectrom Rev* **2009,** *28* (5), 785-815.

13.   Wang, L.; Chance, M. R., Structural Mass Spectrometry of Proteins Using Hydroxyl Radical Based Protein Footprinting. *Anal. Chem.* **2011,** *83* (19), 7234-7241.

14.   Baslé, E.; Joubert, N.; Pucheault, M., Protein Chemical Modification on Endogenous Amino Acids. *Chemistry & Biology* **2010,** *17* (3), 213-227.

15.   Shannon, D. A.; Weerapana, E., Covalent protein modification: the current landscape of residue-specific electrophiles. *Current Opinion in Chemical Biology* **2015,** *24* (Supplement C), 18-26.

16.   Maleknia, S. D.; Brenowitz, M.; Chance, M. R., Millisecond Radiolytic Modification of Peptides by Synchrotron X-rays Identified by Mass Spectrometry. *Analytical Chemistry* **1999,** *71* (18), 3965-3973.

17.   Takamoto, K.; Chance, M. R., Radiolytic protein footprinting with mass spectrometry to probe the structure of macromolecular complexes. *Annual Review of Biophysics and Biomolecular Structure* **2006,** *35* (1), 251-276.

18.   Maleknia, S. D.; Ralston, C. Y.; Brenowitz, M. D.; Downard, K. M.; Chance, M. R., Determination of Macromolecular Folding and Structure by Synchrotron X-Ray Radiolysis Techniques. *Analytical Biochemistry* **2001,** *289* (2), 103-115.

19.   Maleknia, S. D.; Wong, J. W. H.; Downard, K. M., Photochemical and electrophysical production of radicals on millisecond timescales to probe the structure, dynamics and interactions of proteins. *Photochemical & Photobiological Sciences* **2004,** *3* (8), 741-748.

20.   Naganathan, A. N.; Muñoz, V., Scaling of Folding Times with Protein Size. *Journal of the American Chemical Society* **2005,** *127* (2), 480-481.

21.   Sharp, J. S.; Becker, J. M.; Hettich, R. L., Analysis of Protein Solvent Accessible Surfaces by Photochemical Oxidation and Mass Spectrometry. *Analytical Chemistry* **2004,** *76* (3), 672-683.

22.   Aye, T. T.; Low, T. Y.; Sze, S. K., Nanosecond Laser-Induced Photochemical Oxidation Method for Protein Surface Mapping with Mass Spectrometry. *Analytical Chemistry* **2005,** *77* (18), 5814-5822.

23.   Hambly, D. M.; Gross, M. L., Laser Flash Photolysis of Hydrogen Peroxide to Oxidize Protein Solvent-Accessible Residues on the Microsecond Timescale. *Journal of the American Society for Mass Spectrometry* **2005,** *16* (12), 2057-2063.

24.   Gau, B. C.; Sharp, J. S.; Rempel, D. L.; Gross, M. L., Fast Photochemical Oxidation of Protein Footprints Faster than Protein Unfolding. *Analytical Chemistry* **2009,** *81* (16), 6563-6571.

25.   Watson, C.; Janik, I.; Zhuang, T.; Charvátová, O.; Woods, R. J.; Sharp, J. S., Pulsed Electron Beam Water Radiolysis for Submicrosecond Hydroxyl Radical Protein Footprinting. *Analytical Chemistry* **2009,** *81* (7), 2496-2505.

26.   Bridgewater, J. D.; Lim, J.; Vachet, R. W., Transition Metal−Peptide Binding Studied by Metal-Catalyzed Oxidation Reactions and Mass Spectrometry. *Analytical Chemistry* **2006,** *78* (7), 2432-2438.

27.   Bridgewater, J. D.; Lim, J.; Vachet, R. W., Using metal-catalyzed oxidation reactions and mass spectrometry to identify amino acid residues within 10 Å of the metal in Cu-binding proteins. *Journal of the American Society for Mass Spectrometry* **2006,** *17* (11), 1552-1559.

28.   Srikanth, R.; Wilson, J.; Burns, C. S.; Vachet, R. W., Identification of the Copper(II) Coordinating Residues in the Prion Protein by Metal-Catalyzed Oxidation Mass Spectrometry: Evidence for Multiple Isomers at Low Copper(II) Loadings. *Biochemistry* **2008,** *47* (35), 9258-9268.

29.   Hampel, K. J.; Burke, J. M., Time-Resolved Hydroxyl-Radical Footprinting of RNA Using Fe(II)-EDTA. *Methods* **2001,** *23* (3), 233-239.

30.   Tullius, T. D.; Greenbaum, J. A., Mapping nucleic acid structure by hydroxyl radical cleavage. *Current Opinion in Chemical Biology* **2005,** *9* (2), 127-134.

31.   Sharp, J. S.; Becker, J. M.; Hettich, R. L., Protein surface mapping by chemical oxidation: Structural analysis by mass spectrometry. *Analytical Biochemistry* **2003,** *313* (2), 216-225.

32.   Götte, M.; Marquet, R.; Isel, C.; Anderson, V. E.; Keith, G.; Gross, H. J.; Ehresmann, C.; Ehresmann, B.; Heumann, H., Probing the higher order structure of RNA with peroxonitrous acid. *FEBS Letters* **1996,** *390* (2), 226-228.

33.   King, P. A.; Jamison, E.; Strahs, D.; Anderson, V. E.; Brenowitz, M., 'Footprinting' proteins on DNA with peroxonitrous acid. *Nucleic Acids Research* **1993,** *21* (10), 2473-2478.

34.   Maleknia, S. D.; Chance, M. R.; Downard, K. M., Electrospray-assisted modification of proteins: a radical probe of protein structure. *Rapid Communications in Mass Spectrometry* **1999,** *13* (23), 2352-2358.

35.   Wong, J. W. H.; Maleknia, S. D.; Downard, K. M., Study of the Ribonuclease−S-Protein−Peptide Complex Using a Radical Probe and Electrospray Ionization Mass Spectrometry. *Analytical Chemistry* **2003,** *75* (7), 1557-1563.

36.   Wong, J. W. H.; Maleknia, S. D.; Downard, K. M., Hydroxyl radical probe of the calmodulin-melittin complex interface by electrospray ionization mass spectrometry. *Journal of the American Society for Mass Spectrometry* **2005,** *16* (2), 225-233.

37.   McClintock, C.; Kertesz, V.; Hettich, R. L., Development of an Electrochemical Oxidation Method for Probing Higher Order Protein Structure with Mass Spectrometry. *Analytical Chemistry* **2008,** *80* (9), 3304-3317.

38.   Minkoff, B. B.; Blatz, J. M.; Choudhury, F. A.; Benjamin, D.; Shohet, J. L.; Sussman, M. R., Plasma-Generated OH Radical Production for Analyzing Three-Dimensional Structure in Protein Therapeutics. *Scientific Reports* **2017,** *7* (1), 12946.

39.   Sharp, J. S.; Tomer, K. B., Analysis of the Oxidative Damage-Induced Conformational Changes of Apo- and Holocalmodulin by Dose-Dependent Protein Oxidative Surface Mapping. *Biophysical Journal* **2007,** *92* (5), 1682-1692.

40.   Xu, G.; Kiselar, J.; He, Q.; Chance, M. R., Secondary Reactions and Strategies To Improve Quantitative Protein Footprinting. *Analytical Chemistry* **2005,** *77* (10), 3029-3037.

41.   Vahidi, S.; Konermann, L., Probing the Time Scale of FPOP (Fast Photochemical Oxidation of Proteins): Radical Reactions Extend Over Tens of Milliseconds. *Journal of The American Society for Mass Spectrometry* **2016,** *27* (7), 1156-1164.

42.   Blencowe, A.; Hayes, W., Development and application of diazirines in biological and synthetic macromolecular systems. *Soft Matter* **2005,** *1* (3), 178-205.

43.   Suchanek, M.; Radzikowska, A.; Thiele, C., Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells. *Nature Methods* **2005,** *2*, 261.

44.   Jumper, C. C.; Schriemer, D. C., Mass Spectrometry of Laser-Initiated Carbene Reactions for Protein Topographic Analysis. *Analytical Chemistry* **2011,** *83* (8), 2913-2920.

45.   Das, J., Aliphatic Diazirines as Photoaffinity Probes for Proteins: Recent Developments. *Chemical Reviews* **2011,** *111* (8), 4405-4417.

46.   Richards, F. M.; Lamed, R.; Wynn, R.; Patel, D.; Olack, G., Methylene as a possible universal footprinting reagent that will include hydrophobic surface areas: overview and feasibility: properties of diazirine as a precursor. *Protein Science* **2000,** *9* (12), 2506-2517.

47.   Craig, P. O.; Ureta, D. B.; Delfino, J. M., Probing protein conformation with a minimal photochemical reagent. *Protein Science : A Publication of the Protein Society* **2002,** *11* (6), 1353-1366.

48.   Gómez, G. E.; Cauerhff, A.; Craig, P. O.; Goldbaum, F. A.; Delfino, J. M., Exploring protein interfaces with a general photochemical reagent. *Protein Science* **2006,** *15* (4), 744-752.

49.   Ureta, D. B.; Craig, P. O.; Gómez, G. E.; Delfino, J. M., Assessing Native and Non-native Conformational States of a Protein by Methylene Carbene Labeling: The Case of Bacillus licheniformis β-Lactamase. *Biochemistry* **2007,** *46* (50), 14567-14577.

50.   Jumper, C. C.; Bomgarden, R.; Rogers, J.; Etienne, C.; Schriemer, D. C., High-Resolution Mapping of Carbene-Based Protein Footprints. *Analytical Chemistry* **2012,** *84* (10), 4411-4418.

51.   Zhang, B.; Rempel, D. L.; Gross, M. L., Protein Footprinting by Carbenes on a Fast Photochemical Oxidation of Proteins (FPOP) Platform. *Journal of The American Society for Mass Spectrometry* **2016,** *27* (3), 552-555.

52.   Manzi, L.; Barrow, A. S.; Scott, D.; Layfield, R.; Wright, T. G.; Moses, J. E.; Oldham, N. J., Carbene footprinting accurately maps binding sites in protein–ligand and protein–protein interactions. *Nat. Commun.* **2016,** *7*, 13288.

53.   Ziemianowicz, D. S.; Bomgarden, R.; Etienne, C.; Schriemer, D. C., Amino Acid Insertion Frequencies Arising from Photoproducts Generated Using Aliphatic Diazirines. *Journal of The American Society for Mass Spectrometry* **2017,** *28* (10), 2011-2021.

54.   Brunner, J.; Senn, H.; Richards, F. M., 3-Trifluoromethyl-3-phenyldiazirine. A new carbene generating group for photolabeling reagents. *Journal of Biological Chemistry* **1980,** *255* (8), 3313-3318.
55.   Cheng, M.; Zhang, B.; Cui, W.; Gross, M. L., Laser-Initiated Radical Trifluoromethylation of Peptides and Proteins: Application to Mass-Spectrometry-Based Protein Footprinting. *Angewandte Chemie International Edition* **2017,** *56* (45), 14007-14010.
56.   Lilie, J.; Behar, D.; Sujdak, R. J.; Schuler, R. H., Lifetime of trifluoromethyl radical in aqueous solution. *The Journal of Physical Chemistry* **1972,** *76* (18), 2517-2520.
57.   Miles, E. W., Modification of histidyl residues in proteins by diethylpyrocarbonate. In *Methods in Enzymology*, Academic Press: Cambridge, Massachusetts, United States, 1977; Vol. 47, pp 431-442.
58.   Mendoza, V. L.; Vachet, R. W., Protein Surface Mapping Using Diethylpyrocarbonate with Mass Spectrometric Detection. *Analytical Chemistry* **2008,** *80* (8), 2895-2904.
59.   Zhou, Y.; Vachet, R. W., Increased Protein Structural Resolution from Diethylpyrocarbonate-based Covalent Labeling and Mass Spectrometric Detection. *Journal of The American Society for Mass Spectrometry* **2012,** *23* (4), 708-717.
60.   Zhou, Y.; Vachet, R. W., Diethylpyrocarbonate Labeling for the Structural Analysis of Proteins: Label Scrambling in Solution and How to Avoid It. *Journal of The American Society for Mass Spectrometry* **2012,** *23* (5), 899-907.
61.   Goldsmith, S. C.; Guan, J.-Q.; Almo, S. C.; Chance, M. R., Synchrotron Protein Footprinting: A Technique to Investigate Protein-Protein Interactions. *Journal of Biomolecular Structure and Dynamics* **2001,** *19* (3), 405-418.
62.   Kiselar, J. G.; Maleknia, S. D.; Sullivan, M.; Downard, K. M.; Chance, M. R., Hydroxyl radical probe of protein surfaces using synchrotron X-ray radiolysis and mass spectrometry. *International Journal of Radiation Biology* **2002,** *78* (2), 101-114.
63.   Srikanth, R.; Mendoza, V. L.; Bridgewater, J. D.; Zhang, G.; Vachet, R. W., Copper Binding to β-2-Microglobulin and Its Pre-Amyloid Oligomers. *Biochemistry* **2009,** *48* (41), 9871-9881.
64.   Huang, W.; Ravikumar, Krishnakumar M.; Chance, Mark R.; Yang, S., Quantitative Mapping of Protein Structure by Hydroxyl Radical Footprinting-Mediated Structural Mass Spectrometry: A Protection Factor Analysis. *Biophysical Journal* **2015,** *108* (1), 107-115.
65.   Xie, B.; Sood, A.; Woods, R. J.; Sharp, J. S., Quantitative Protein Topography Measurements by High Resolution Hydroxyl Radical Protein Footprinting Enable Accurate Molecular Model Selection. *Scientific Reports* **2017,** *7* (1), 4552.
66.   Wang, L.; Chance, M. R., Protein footprinting comes of age: mass spectrometry for biophysical structure assessment. *Molecular & Cellular Proteomics* **2017**.
67.   Xu, G.; Chance, M. R., Radiolytic Modification and Reactivity of Amino Acid Residues Serving as Structural Probes for Protein Footprinting. *Analytical Chemistry* **2005,** *77* (14), 4549-4555.
68.   Mehler, E. L.; Fuxreiter, M.; Simon, I.; Garcia-Moreno E, B., The role of hydrophobic microenvironments in modulating pKa shifts in proteins. *Proteins: Struct., Funct., Bioinf.* **2002,** *48* (2), 283-292.

69.   Harris, T. K.; Turner, G. J., Structural Basis of Perturbed pKa Values of Catalytic Groups in Enzyme Active Sites. *IUBMB Life* **2002,** *53* (2), 85-98.

70.   Isom, D. G.; Castañeda, C. A.; Cannon, B. R.; García-Moreno E, B., Large shifts in pKa values of lysine residues buried inside a protein. *Proceedings of the National Academy of Sciences* **2011,** *108* (13), 5260.

71.   Peck, M. T.; Ortega, G.; De Luca-Johnson, J. N.; Schlessman, J. L.; Robinson, A. C.; García-Moreno E, B., Local Backbone Flexibility as a Determinant of the Apparent pKa Values of Buried Ionizable Groups in Proteins. *Biochemistry* **2017,** *56* (40), 5338-5346.

72.   Zhou, Y.; Wu, Y.; Yao, M.; Liu, Z.; Chen, J.; Chen, J.; Tian, L.; Han, G.; Shen, J.-R.; Wang, F., Probing the Lysine Proximal Microenvironments within Membrane Protein Complexes by Active Dimethyl Labeling and Mass Spectrometry. *Anal. Chem.* **2016,** *88* (24), 12060-12065.

73.   Borotto, N. B.; Zhou, Y.; Hollingsworth, S. R.; Hale, J. E.; Graban, E. M.; Vaughan, R. C.; Vachet, R. W., Investigating Therapeutic Protein Structure with Diethylpyrocarbonate Labeling and Mass Spectrometry. *Analytical Chemistry* **2015,** *87* (20), 10627-10634.

74.   Herrmann, A.; Svangård, E.; Claeson, P.; Gullbo, J.; Bohlin, L.; Göransson, U., Key role of glutamic acid for the cytotoxic activity of the cyclotide cycloviolacin O2. *Cellular and Molecular Life Sciences CMLS* **2006,** *63* (2), 235-245.

75.   Tong, X.; Wren, J. C.; Konermann, L., Effects of Protein Concentration on the Extent of γ-Ray-Mediated Oxidative Labeling Studied by Electrospray Mass Spectrometry. *Analytical Chemistry* **2007,** *79* (16), 6376-6382.

76.   Kiselar, J. G.; Janmey, P. A.; Almo, S. C.; Chance, M. R., Structural Analysis of Gelsolin Using Synchrotron Protein Footprinting. *Molecular & Cellular Proteomics* **2003,** *2* (10), 1120.

77.   Gupta, S.; Celestre, R.; Petzold, C. J.; Chance, M. R.; Ralston, C., Development of a microsecond X-ray protein footprinting facility at the Advanced Light Source. *Journal of Synchrotron Radiation* **2014,** *21* (4), 690-699.

78.   Mendoza, V. L.; Antwi, K.; Barón-Rodríguez, M. A.; Blanco, C.; Vachet, R. W., Structure of the Preamyloid Dimer of β-2-Microglobulin from Covalent Labeling and Mass Spectrometry. *Biochemistry* **2010,** *49* (7), 1522-1532.

79.   Mendoza, V. L.; Barón-Rodríguez, M. A.; Blanco, C.; Vachet, R. W., Structural Insights into the Pre-Amyloid Tetramer of β-2-Microglobulin from Covalent Labeling and Mass Spectrometry. *Biochemistry* **2011,** *50* (31), 6711-6722.

80.   Ivankov, D. N.; Finkelstein, A. V., Prediction of protein folding rates from the amino acid sequence-predicted secondary structure. *Proceedings of the National Academy of Sciences of the United States of America* **2004,** *101* (24), 8942-8944.

81.   Liu, T.; Marcinko, T. M.; Kiefer, P. A.; Vachet, R. W., Using Covalent Labeling and Mass Spectrometry To Study Protein Binding Sites of Amyloid Inhibiting Molecules. *Analytical Chemistry* **2017,** *89* (21), 11583-11591.

82.   Madsen, J. A.; Yin, Y.; Qiao, J.; Gill, V.; Renganathan, K.; Fu, W.-Y.; Smith, S.; Anderson, J., Covalent Labeling Denaturation Mass Spectrometry for Sensitive Localized Higher Order Structure Comparisons. *Analytical Chemistry* **2016,** *88* (4), 2478-2488.

83. Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W., Top Down versus Bottom Up Protein Characterization by Tandem High-Resolution Mass Spectrometry. *Journal of the American Chemical Society* **1999,** *121* (4), 806-812.

84. McLafferty, F. W.; Breuker, K.; Jin, M.; Han, X.; Infusini, G.; Jiang, H.; Kong, X.; Begley, T. P., Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics. *FEBS Journal* **2007,** *274* (24), 6256-6268.

85. Siuti, N.; Kelleher, N. L., Decoding protein modifications using top-down mass spectrometry. *Nature Methods* **2007,** *4*, 817.

86. Meng, F.; Forbes, A. J.; Miller, L. M.; Kelleher, N. L., Detection and localization of protein modifications by high resolution tandem mass spectrometry. *Mass Spectrometry Reviews* **2005,** *24* (2), 126-134.

87. Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M.-C.; Yates, J. R., Protein Analysis by Shotgun/Bottom-up Proteomics. *Chemical reviews* **2013,** *113* (4), 2343-2394.

88. Chen, J.; Cui, W.; Giblin, D.; Gross, M. L., New Protein Footprinting: Fast Photochemical Iodination Combined with Top-down and Bottom-up Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2012,** *23* (8), 1306-1318.

89. Cammarata, M.; Lin, K.-Y.; Pruet, J.; Liu, H.-w.; Brodbelt, J., Probing the Unfolding of Myoglobin and Domain C of PARP-1 with Covalent Labeling and Top-Down Ultraviolet Photodissociation Mass Spectrometry. *Analytical Chemistry* **2014,** *86* (5), 2534-2542.

90. Li, J.; Wei, H.; Krystek, S. R.; Bond, D.; Brender, T. M.; Cohen, D.; Feiner, J.; Hamacher, N.; Harshman, J.; Huang, R. Y. C.; Julien, S. H.; Lin, Z.; Moore, K.; Mueller, L.; Noriega, C.; Sejwal, P.; Sheppard, P.; Stevens, B.; Chen, G.; Tymiak, A. A.; Gross, M. L.; Schneeweis, L. A., Mapping the Energetic Epitope of an Antibody/Interleukin-23 Interaction with Hydrogen/Deuterium Exchange, Fast Photochemical Oxidation of Proteins Mass Spectrometry, and Alanine Shave Mutagenesis. *Analytical Chemistry* **2017,** *89* (4), 2250-2258.

91. Srikanth, R.; Wilson, J.; Bridgewater, J. D.; Numbers, J. R.; Lim, J.; Olbris, M. R.; Kettani, A.; Vachet, R. W., Improved Sequencing of Oxidized Cysteine and Methionine Containing Peptides Using Electron Transfer Dissociation. *Journal of the American Society for Mass Spectrometry* **2007,** *18* (8), 1499-1506.

92. Srikanth, R.; Wilson, J.; Vachet, R. W., Correct identification of oxidized histidine residues using electron-transfer dissociation. *Journal of Mass Spectrometry* **2009,** *44* (5), 755-762.

93. Li, X.; Li, Z.; Xie, B.; Sharp, J. S., Improved Identification and Relative Quantification of Sites of Peptide and Protein Oxidation for Hydroxyl Radical Footprinting. *Journal of The American Society for Mass Spectrometry* **2013,** *24* (11), 1767-1776.

94. Borotto, N. B.; Degraan-Weber, N.; Zhou, Y.; Vachet, R. W., Label Scrambling During CID of Covalently Labeled Peptide Ions. *J. Am. Soc. Mass Spectrom.* **2014,** *25* (10), 1739-1746.

95. Li, Z.; Moniz, H.; Wang, S.; Ramiah, A.; Zhang, F.; Moremen, K. W.; Linhardt, R. J.; Sharp, J. S., High Structural Resolution Hydroxyl Radical Protein Footprinting Reveals an Extended Robo1-Heparin Binding Interface. *Journal of Biological Chemistry* **2015,** *290* (17), 10729-10740.

96.   Jones, L. M.; B. Sperry, J.; A. Carroll, J.; Gross, M. L., Fast Photochemical Oxidation of Proteins for Epitope Mapping. *Analytical Chemistry* **2011,** *83* (20), 7657-7661.

97.   Zhang, Y.; Wecksler, A. T.; Molina, P.; Deperalta, G.; Gross, M. L., Mapping the Binding Interface of VEGF and a Monoclonal Antibody Fab-1 Fragment with Fast Photochemical Oxidation of Proteins (FPOP) and Mass Spectrometry. *Journal of The American Society for Mass Spectrometry* **2017,** *28* (5), 850-858.

98.   Kaur, P.; Kiselar, J. G.; Chance, M. R., Integrated Algorithms for High-Throughput Examination of Covalently Labeled Biomolecules by Structural Mass Spectrometry. *Analytical Chemistry* **2009,** *81* (19), 8141-8149.

99.   Rinas, A.; Espino, J. A.; Jones, L. M., An efficient quantitation strategy for hydroxyl radical-mediated protein footprinting using Proteome Discoverer. *Analytical and Bioanalytical Chemistry* **2016,** *408* (11), 3021-3031.

100.   Dobson, C. M., Protein folding and misfolding. *Nature* **2003,** *426*, 884.

101.   Cohen, F. E.; Kelly, J. W., Therapeutic approaches to protein-misfolding diseases. *Nature* **2003,** *426*, 905.

102.   Luheshi, L. M.; Crowther, D. C.; Dobson, C. M., Protein misfolding and disease: from the test tube to the organism. *Current Opinion in Chemical Biology* **2008,** *12* (1), 25-31.

103.   Konermann, L.; Stocks, B. B.; Pan, Y.; Tong, X., Mass spectrometry combined with oxidative labeling for exploring protein structure and folding. *Mass Spectrometry Reviews* **2010,** *29* (4), 651-667.

104.   Wong, K.-B.; Freund, S. M. V.; Fersht, A. R., Cold Denaturation of Barstar:1H,15N and13C NMR Assignment and Characterisation of Residual Structure. *Journal of Molecular Biology* **1996,** *259* (4), 805-818.

105.   Chen, J.; Rempel, D. L.; Gross, M. L., Temperature Jump and Fast Photochemical Oxidation Probe Submillisecond Protein Folding. *J. Am. Chem. Soc.* **2010,** *132* (44), 15502-15504.

106.   Chen, J.; Rempel, D. L.; Gau, B. C.; Gross, M. L., Fast Photochemical Oxidation of Proteins and Mass Spectrometry Follow Submillisecond Protein Folding at the Amino-Acid Level. *J. Am. Chem. Soc.* **2012,** *134* (45), 18724-18731.

107.   Gau, B. C.; Chen, J.; Gross, M. L., Fast photochemical oxidation of proteins for comparing solvent-accessibility changes accompanying protein folding: Data processing and application to barstar. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2013,** *1834* (6), 1230-1238.

108.   Khurana, R.; Udgaonkar, J. B., Equilibrium unfolding studies of barstar: Evidence for an alternative conformation which resembles a molten globule. *Biochemistry* **1994,** *33* (1), 106-115.

109.   Shastry, M. C. R.; Udgaonkar, J. B., The Folding Mechanism of Barstar: Evidence for Multiple Pathways and Multiple Intermediates. *Journal of Molecular Biology* **1995,** *247* (5), 1013-1027.

110.   Nölting, B.; Golbik, R.; Fersht, A. R., Submillisecond events in protein folding. *Proceedings of the National Academy of Sciences of the United States of America* **1995,** *92* (23), 10668-10672.

111.   Nölting, B.; Golbik, R.; Neira, J. L.; Soler-Gonzalez, A. S.; Schreiber, G.; Fersht, A. R., The folding pathway of a protein at high resolution from microseconds to seconds. *Proceedings of the National Academy of Sciences* **1997,** *94* (3), 826-830.

112.   Poor, T. A.; Jones, L. M.; Sood, A.; Leser, G. P.; Plasencia, M. D.; Rempel, D. L.; Jardetzky, T. S.; Woods, R. J.; Gross, M. L.; Lamb, R. A., Probing the paramyxovirus fusion (F) protein-refolding event from pre- to postfusion by oxidative footprinting. *Proceedings of the National Academy of Sciences* **2014,** *111* (25), E2596-E2605.

113.   Stocks, B. B.; Konermann, L., Structural Characterization of Short-Lived Protein Unfolding Intermediates by Laser-Induced Oxidative Labeling and Mass Spectrometry. *Anal. Chem.* **2009,** *81* (1), 20-27.

114.   Pan, Y.; Brown, L.; Konermann, L., Kinetic Folding Mechanism of an Integral Membrane Protein Examined by Pulsed Oxidative Labeling and Mass Spectrometry. *Journal of Molecular Biology* **2011,** *410* (1), 146-158.

115.   Vahidi, S.; Stocks, B. B.; Liaghati-Mobarhan, Y.; Konermann, L., Mapping pH-Induced Protein Structural Changes Under Equilibrium Conditions by Pulsed Oxidative Labeling and Mass Spectrometry. *Analytical Chemistry* **2012,** *84* (21), 9124-9130.

116.   Vahidi, S.; Stocks, B. B.; Liaghati-Mobarhan, Y.; Konermann, L., Submillisecond Protein Folding Events Monitored by Rapid Mixing and Mass Spectrometry-Based Oxidative Labeling. *Anal. Chem.* **2013,** *85* (18), 8618-8625.

117.   Hazenberg, B. P. C., Amyloidosis: A Clinical Overview. *Rheumatic Disease Clinics of North America* **2013,** *39* (2), 323-345.

118.   Li, H.; Rahimi, F.; Sinha, S.; Maiti, P.; Bitan, G.; Murakami, K., Amyloids and Protein Aggregation—Analytical Methods. In *Encyclopedia of Analytical Chemistry*, John Wiley & Sons, Ltd: 2006.

119.   Verma, M.; Vats, A.; Taneja, V., Toxic species in amyloid disorders: Oligomers or mature fibrils. *Annals of Indian Academy of Neurology* **2015,** *18* (2), 138-145.

120.   Shankar, G. M.; Li, S.; Mehta, T. H.; Garcia-Munoz, A.; Shepardson, N. E.; Smith, I.; Brett, F. M.; Farrell, M. A.; Rowan, M. J.; Lemere, C. A.; Regan, C. M.; Walsh, D. M.; Sabatini, B. L.; Selkoe, D. J., Amyloid-β protein dimers isolated directly from Alzheimer's brains impair synaptic plasticity and memory. *Nature Medicine* **2008,** *14*, 837.

121.   Sengupta, U.; Nilson, A. N.; Kayed, R., The Role of Amyloid-β Oligomers in Toxicity, Propagation, and Immunotherapy. *EBioMedicine* **2016,** *6*, 42-49.

122.   Bitan, G., Structural Study of Metastable Amyloidogenic Protein Oligomers by Photo-Induced Cross-Linking of Unmodified Proteins. In *Methods in Enzymology*, Academic Press: 2006; Vol. 413, pp 217-236.

123.   Marcinko, T. M.; Dong, J.; LeBlanc, R.; Daborowski, K. V.; Vachet, R. W., Small molecule-mediated inhibition of β-2-microglobulin-based amyloid fibril formation. *Journal of Biological Chemistry* **2017,** *292* (25), 10630-10638.

124.   Klinger, A. L.; Kiselar, J.; Ilchenko, S.; Komatsu, H.; Chance, M. R.; Axelsen, P. H., A Synchrotron-Based Hydroxyl Radical Footprinting Analysis of Amyloid Fibrils and Prefibrillar Intermediates with Residue-Specific Resolution. *Biochemistry* **2014,** *53* (49), 7724-7734.

125.   Li, K. S.; Rempel, D. L.; Gross, M. L., Conformational-Sensitive Fast Photochemical Oxidation of Proteins and Mass Spectrometry Characterize Amyloid Beta 1–42 Aggregation. *J. Am. Chem. Soc.* **2016,** *138* (37), 12090-12098.

126.   Floege, J.; Ketteler, M., β2-Microglobulin–derived amyloidosis: An update. *Kidney International* **2001,** *59*, S164-S171.

127.   Guan, J.-Q.; Vorobiev, S.; Almo, S. C.; Chance, M. R., Mapping the G-Actin Binding Surface of Cofilin Using Synchrotron Protein Footprinting. *Biochemistry* **2002,** *41* (18), 5765-5775.

128.   Rashidzadeh, H.; Khrapunov, S.; Chance, M. R.; Brenowitz, M., Solution Structure and Interdomain Interactions of the Saccharomyces cerevisiae "TATA Binding Protein" (TBP) Probed by Radiolytic Protein Footprinting. *Biochemistry* **2003,** *42* (13), 3655-3665.

129.   Hambly, D.; Gross, M., Laser flash photochemical oxidation to locate heme binding and conformational changes in myoglobin. *International Journal of Mass Spectrometry* **2007,** *259* (1), 124-129.

130.   Zhang, H.; Gau, B. C.; Jones, L. M.; Vidavsky, I.; Gross, M. L., Fast Photochemical Oxidation of Proteins for Comparing Structures of Protein−Ligand Complexes: The Calmodulin−Peptide Model System. *Anal. Chem.* **2011,** *83* (1), 311-318.

131.   Padayatti, P. S.; Wang, L.; Gupta, S.; Orban, T.; Sun, W.; Salom, D.; Jordan, S. R.; Palczewski, K.; Chance, M. R., A Hybrid Structural Approach to Analyze Ligand Binding by the Serotonin Type 4 Receptor (5-HT4). *Molecular & Cellular Proteomics : MCP* **2013,** *12* (5), 1259-1271.

132.   Misra, S. K.; Sood, A.; Soares, P. A.; Pomin, V. H.; Woods, R. J.; Sharp, J. S., Mapping of the Fondaparinux Binding Site of JR-FL gp120 by High Resolution Hydroxyl Radical Protein Footprinting and Computational Docking. *bioRxiv* **2017**.

133.   Wang, L.; Qin, Y.; Ilchenko, S.; Bohon, J.; Shi, W.; Cho, M. W.; Takamoto, K.; Chance, M. R., Structural Analysis of a Highly Glycosylated and Unliganded gp120-Based Antigen Using Mass Spectrometry. *Biochemistry* **2010,** *49* (42), 9032-9045.

134.   Zheng, X.; Wintrode, P. L.; Chance, M. R., Complementary Structural Mass Spectrometry Techniques Reveal Local Dynamics in Functionally Important Regions of a Metastable Serpin. *Structure* **2008,** *16* (1), 38-51.

135.   Gerega, S. K.; Downard, K. M., PROXIMO—a new docking algorithm to model protein complexes using data from radical probe mass spectrometry (RP-MS). *Bioinformatics* **2006,** *22* (14), 1702-1709.

136.   Kamal, J. K. A.; Chance, M. R., Modeling of protein binary complexes using structural mass spectrometry data. *Protein Science* **2008,** *17* (1), 79-94.

137.   Charvátová, O.; Foley, B. L.; Bern, M. W.; Sharp, J. S.; Orlando, R.; Woods, R. J., Quantifying Protein Interface Footprinting by Hydroxyl Radical Oxidation and Molecular Dynamics Simulation: Application to Galectin-1. *Journal of the American Society for Mass Spectrometry* **2008,** *19* (11), 1692-1705.

138.   Schorzman, A. N.; Perera, L.; Cutalo-Patterson, J. M.; Pedersen, L. C.; Pedersen, L. G.; Kunkel, T. A.; Tomer, K. B., Modeling of the DNA-binding site of yeast Pms1 by mass spectrometry *DNA repair* **2011,** *10* (5), 454-465.

139.   Huang, W.; Ravikumar, K. M.; Parisien, M.; Yang, S., Theoretical modeling of multiprotein complexes by iSPOT: Integration of small-angle X-ray scattering, hydroxyl

radical footprinting, and computational docking. *Journal of structural biology* **2016,** *196* (3), 340-349.

140. Schmidt, C.; Macpherson, J. A.; Lau, A. M.; Tan, K. W.; Fraternali, F.; Politis, A., Surface Accessibility and Dynamics of Macromolecular Assemblies Probed by Covalent Labeling Mass Spectrometry and Integrative Modeling. *Analytical Chemistry* **2017,** *89* (3), 1459-1468.

141. Pan, Y.; Piyadasa, H.; O'Neil, J. D.; Konermann, L., Conformational Dynamics of a Membrane Transport Protein Probed by H/D Exchange and Covalent Labeling: The Glycerol Facilitator. *Journal of Molecular Biology* **2012,** *416* (3), 400-413.

142. Pan, Y.; Ruan, X.; Valvano, M. A.; Konermann, L., Validation of Membrane Protein Topology Models by Oxidative Labeling and Mass Spectrometry. *Journal of The American Society for Mass Spectrometry* **2012,** *23* (5), 889-898.

143. Konermann, L.; Pan, Y., Exploring membrane protein structural features by oxidative labeling and mass spectrometry. *Expert Review of Proteomics* **2012,** *9* (5), 497-504.

144. Bavro, Vassiliy N.; Gupta, S.; Ralston, C., Oxidative footprinting in the study of structure and function of membrane proteins: current state and perspectives. *Biochemical Society Transactions* **2015,** *43* (5), 983.

145. Lu, Y.; Zhang, H.; Niedzwiedzki, D. M.; Jiang, J.; Blankenship, R. E.; Gross, M. L., Fast Photochemical Oxidation of Proteins Maps the Topology of Intrinsic Membrane Proteins: Light-Harvesting Complex 2 in a Nanodisc. *Analytical Chemistry* **2016,** *88* (17), 8827-8834.

146. Watkinson, T. G.; Calabrese, A. N.; Ault, J. R.; Radford, S. E.; Ashcroft, A. E., FPOP-LC-MS/MS Suggests Differences in Interaction Sites of Amphipols and Detergents with Outer Membrane Proteins. *Journal of The American Society for Mass Spectrometry* **2017,** *28* (1), 50-55.

147. Shen, G.; Li, S.; Cui, W.; Liu, S.; Yang, Y.; Gross, M.; Li, W., Membrane Protein Structure in Live Cells: Methodology for Studying Drug Interaction by Mass Spectrometry-Based Footprinting. *Biochemistry* **2018**, Article ASAP.

148. Espino, J. A.; Mali, V. S.; Jones, L. M., In Cell Footprinting Coupled with Mass Spectrometry for the Structural Analysis of Proteins in Live Cells. *Analytical Chemistry* **2015,** *87* (15), 7971-7978.

149. Leader, B.; Baca, Q. J.; Golan, D. E., Protein therapeutics: a summary and pharmacological classification. *Nature Reviews Drug Discovery* **2008,** *7*, 21.

150. Ecker, D. M.; Jones, S. D.; Levine, H. L., The therapeutic monoclonal antibody market. *mAbs* **2015,** *7* (1), 9-14.

151. 2016 FDA Approved Drugs. http://www.centerwatch.com/drug-information/fda-approved-drugs/year/2016 (accessed Mar 29).

152. Berkowitz, S. A.; Engen, J. R.; Mazzeo, J. R.; Jones, G. B., Analytical tools for characterizing biopharmaceuticals and the implications for biosimilars. *Nat Rev Drug Discov* **2012,** *11* (7), 527-540.

153. Leader, B.; Baca, Q. J.; Golan, D. E., Protein therapeutics: a summary and pharmacological classification. *Nat Rev Drug Discov* **2008,** *7* (1), 21-39.

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

155.   Beck, A.; Wurch, T.; Bailly, C.; Corvaia, N., Strategies and challenges for the next generation of therapeutic antibodies. *Nat Rev Immunol* **2010,** *10* (5), 345-352.

156.   Beck, A.; Diemer, H.; Ayoub, D.; Debaene, F.; Wagner-Rousset, E.; Carapito, C.; Van Dorsselaer, A.; Sanglier-Cianférani, S., Analytical characterization of biosimilar antibodies and Fc-fusion proteins. *Trends Anal Chem* **2013,** *48*, 81-95.

157.   Kaltashov, I. A.; Bobst, C. E.; Abzalimov, R. R.; Wang, G.; Baykal, B.; Wang, S., Advances and challenges in analytical characterization of biotechnology products: Mass spectrometry-based approaches to study properties and behavior of protein therapeutics. *Biotechnol Adv* **2012,** *30* (1), 210-222.

158.   Berkowitz, S. A.; Engen, J. R.; Mazzeo, J. R.; Jones, G. B., Analytical tools for characterizing biopharmaceuticals and the implications for biosimilars. *Nature Reviews Drug Discovery* **2012,** *11*, 527.

159.   Frokjaer, S.; Otzen, D. E., Protein drug stability: a formulation challenge. *Nature Reviews Drug Discovery* **2005,** *4*, 298.

160.   Kaltashov, I. A.; Bobst, C. E.; Abzalimov, R. R.; Berkowitz, S. A.; Houde, D., Conformation and Dynamics of Biopharmaceuticals: Transition of Mass Spectrometry-Based Tools from Academe to Industry. *Journal of the American Society for Mass Spectrometry* **2010,** *21* (3), 323-337.

161.   Kaltashov, I. A.; Bobst, C. E.; Abzalimov, R. R.; Wang, G.; Baykal, B.; Wang, S., Advances and challenges in analytical characterization of biotechnology products: Mass spectrometry-based approaches to study properties and behavior of protein therapeutics. *Biotechnology Advances* **2012,** *30* (1), 210-222.

162.   Weiss, W. F.; Gabrielson, J. P.; Al-Azzam, W.; Chen, G.; Davis, D. L.; Das, T. K.; Hayes, D. B.; Houde, D.; Singh, S. K., Technical Decision Making With Higher Order Structure Data: Perspectives on Higher Order Structure Characterization From the Biopharmaceutical Industry. *Journal of Pharmaceutical Sciences* **2016,** *105* (12), 3465-3470.

163.   Zhang, H.; Cui, W.; Gross, M. L., Mass spectrometry for the biophysical characterization of therapeutic monoclonal antibodies. *FEBS Letters* **2014,** *588* (2), 308-317.

164.   Ambrogelly, A.; Gozo, S.; Katiyar, A.; Dellatore, S.; Kune, Y.; Bhat, R.; Sun, J.; Li, N.; Wang, D.; Nowak, C.; Neill, A.; Ponniah, G.; King, C.; Mason, B.; Beck, A.; Liu, H., Analytical comparability study of recombinant monoclonal antibody therapeutics. *mAbs* **2018,** *10* (4), 513-538.

165.   Watson, C.; Sharp, J. S., Conformational analysis of therapeutic proteins by hydroxyl radical protein footprinting. *AAPS J.* **2012,** *14* (2), 206-217.

166.   Deperalta, G.; Alvarez, M.; Bechtel, C.; Dong, K.; McDonald, R.; Ling, V., Structural analysis of a therapeutic monoclonal antibody dimer by hydroxyl radical footprinting. *mAbs* **2013,** *5* (1), 86-101.

167.   Abbott, W. M.; Damschroder, M. M.; Lowe, D. C., Current approaches to fine mapping of antigen–antibody interactions. *Immunology* **2014,** *142* (4), 526-535.

168.   Li, K. S.; Chen, G.; Mo, J.; Huang, R. Y. C.; Deyanova, E. G.; Beno, B. R.; O'Neil, S. R.; Tymiak, A. A.; Gross, M. L., Orthogonal Mass Spectrometry-Based Footprinting for Epitope Mapping and Structural Characterization: The IL-6 Receptor upon Binding of Protein Therapeutics. *Analytical Chemistry* **2017,** *89* (14), 7742-7749.

CHAPTER 2

COVALENT LABELING WITH DIETHYLPYROCARBONATE: SENSITIVE TO RESIDUE MICROENVIRONMENT, PROVIDING IMPROVED ANALYSIS OF PROTEIN HIGHER ORDER STRUCTURE BY MASS SPECTROMETRY

This chapter is part of a research article published as: Limpikirati, P.; Pan, X.; Vachet, R. W., Covalent Labeling with Diethylpyrocarbonate: Sensitive to the Residue Microenvironment, Providing Improved Analysis of Protein Higher Order Structure by Mass Spectrometry. *Analytical Chemistry* **2019,** *91* (13), 8516-8523.

2.1 Introduction

Covalent labeling (CL) with mass spectrometry (MS) has been increasingly utilized for the higher order structural analysis of proteins and for studying protein-protein interactions. From reactions with labeling reagents, a protein's structural properties or interactions can be encoded into the mass of protein, which can then be read-out in a straightforward site-specific manner via MS-based bottom-up analysis. The differential reactivity of proteins under different conditions can be used to distinguish and deduce protein conformations, topologies, and/or binding sites. Common applications of CL-MS techniques include protein folding/unfolding,[1-4] amyloid-forming proteins,[5-9] protein-ligand systems,[10-13] and structural analysis of protein therapeutics,[14-19] among others. Several review articles are dedicated to these relatively new MS-based structural techniques.[20-26] The wider use of CL-MS techniques as an orthogonal method to probe structures of proteins results from their unique advantages over other analytical methods because of (a) a more rapid, sensitive, and sample-efficient analysis compared to NMR, X-ray crystallography, and cryo-EM, (b) the ability to obtain much higher structural resolution than optical spectroscopy, i.e. residue-level structural information can be obtained from CL-MS, and (c) the limited label loss and scrambling due to irreversible

nature of labeling, compared to hydrogen-deuterium exchange (HDX) – MS. CL reagents modify solvent-exposed amino acid side chains either at specific residues (e.g. Lys labeling, Trp labeling, and carboxylic acid labeling) or a wide range of residues. In the latter case, non-specific reagents, e.g. hydroxyl radicals, carbenes, and diethylpyrocarbonate, can be used to gain greater structural resolution as they can probe a range of side chains simultaneously.

Diethylpyrocarbonate (DEPC) is a simple-to-use, commercially-available reagent that reacts with a range of nucleophilic residues and the N-terminus of proteins.[20, 21] Unlike radical and carbene reagents, DEPC can readily label proteins once added to solution and no specialized instrumentation is needed to generate a reactive species.[20, 21] CL-MS methods based on DEPC have been extensively developed by our group.[27-30] We have shown its capability to obtain insight into protein-ligand and protein-protein interactions and conformational changes of therapeutic proteins with a structural resolution as low as 8 – 10 Å because up to 30% of surface exposed residues on proteins can be labeled.[5-7, 12, 16] In addition to Cys, His and Lys which are good nucleophiles that readily react with DEPC and account for around 10% coverage of the average protein sequence,[31] side chains of weakly nucleophilic residues, such as Ser, Thr, and Tyr, are also found to be modified by DEPC. We find that in many different proteins,[16, 19, 27-29] these weak nucleophiles can be modified by DEPC to a significant extent in intact proteins. The reactivity of these residues is valuable because Ser, Thr, and Tyr are the third, seventh, and sixteenth most commonly-found amino acids in proteins, respectively, and the three of them account for about 17% of the average protein sequence.[31] Indeed, the ability of DEPC to label Ser, Thr, and Tyr

residues in proteins increases the resolution of structural analyses using DEPC-based CL-MS.

While Ser, Thr, and Tyr residues in proteins can be reactive with DEPC, we have found a relatively poor correlation between the solvent accessibility of these residues and the extent of their DEPC reactivity, which contrasts with the more reactive Lys and His residues. This contrast suggests that microenvironment might affect the reactivity of Ser, Thr, and Tyr residues.[16, 19, 27, 29] Previous experimental and computational studies suggest that the reactivity of amino acid side chains in proteins can be tuned by the microenvironment in different ways. Microenvironment can influence the acid/base characteristics and the reactivity of side chains because of polar interactions (i.e. charge-charge, charge-dipole, or dipole-dipole)[32, 33] or hydrophobic effects via Born or desolvation effects.[32, 34-37] Steric hindrance as created by intra/intermolecular interactions of nearby side chains can limit reactivity.[38] Also, certain microenvironments have been suggested to increase local concentrations of reactants and/or a reagent's affinity for specific protein surface sites, thus enhancing the reactivity of some side chains.[33, 39-41] In this work, we have investigated whether microenvironment effects explain the relatively poor correlation between DEPC reactivity and the solvent accessibility of Ser, Thr, and Tyr residues in proteins by comparing the DEPC reactivity of model peptides to proteins containing the same sequences but in a 3-dimensional context. We find that free peptides with Ser, Thr, and Tyr residues rarely react with DEPC under normal labeling conditions, whereas these residues in intact proteins can be very reactive. A careful study of the microenvironment around reactive Ser, Thr, and Tyr residues in proteins indicate that nearby hydrophobic

patches increase the reactivity of these residues, presumably through an increased local concentration effect.

2.2 Experimental Section

2.2.1 Materials

Human full-length beta-2-microglobulin (β2m) (#126-11) and recombinant human growth hormone (hGH) (#4769) were obtained from Lee Biosolutions (Maryland Heights, MO) and BioVision (Milpitas, CA), respectively. Bovine ubiquitin (#U6253), bradykinin (#B3259), diethylpyrocarbonate (DEPC) (#D5758), imidazole (#I5513), iodoacetamide (#I6125), tris(2-carboxyethyl)phosphine (TCEP) (#C4706), 3-morpholinopropane-1-sulfonic acid (MOPS) (#M1254), and MOPS sodium salt (#M9381) were all purchased from Sigma-Aldrich (St. Louis, MO). Preproenkephalin was ordered from American Peptide Company (Sunnyvale, CA). Model peptides (Fmoc-DGXGG-amide, where X = H, K, Y, S, or T) were custom synthesized by GenScript USA Inc (Piscataway, NJ). Pierce™ sulfo-NHS-Acetate (#26777) was obtained from Thermo Scientific (Waltham, MA). Sodium phosphate monobasic monohydrate (#S0710) was purchased from EM Science (Darmstadt, Hesse, Germany). Sodium phosphate dibasic anhydrous (#S374), N,N-Dimethylformamide (#D119), formic acid (#A117), acetonitrile (#A998), and water (#W7) were obtained from Fisher Scientific (Fair Lawn, NJ). Immobilized trypsin (#EN-251) and chymotrypsin (#EN-261) were purchased from Princeton Separations (Adelphia, NJ). Amicon® centrifugal filters (#UFC5010 and #UFC5003) were bought from EMD Millipore (Burlington, MA). All reagents used in this study have no known potential hazards.

2.2.2 Sample Preparation and DEPC Labeling Reactions

β-2-microglobulin (β2m) and ubiquitin were prepared in 10 mM MOPS buffer (pH 7.4), while human growth hormone (hGH) was prepared in 10 mM phosphate buffer (pH 8.0). Stock solutions of DEPC at a concentration of 69 mM were freshly prepared in acetonitrile. DEPC labeling of β2m (50 μM) was initiated by adding DEPC in a molar excess of 4, and the solution was reacted for 1 min at 37 °C. DEPC labeling of ubiquitin (10 μM) was performed at 37 °C for 5 min at a DEPC to protein molar ratio of 4 to 1. DEPC labeling of hGH (30 μM) was performed at 5:1 (DEPC:protein) molar ratio for 1 min at 37 °C. In all cases, the reaction was quenched by the addition of imidazole at a 1:50 DEPC to imidazole molar ratio, and the final amount of acetonitrile present was 1%, which does not perturb the structure of protein during the labeling. Labeling experiments with these intact proteins were performed in triplicate or quadruplicate.

For the DEPC reactions of free peptides, peptides from the same proteins were first produced via proteolytic digestion (see **Section 2.2.3**). The N-termini of the peptides were acetylated before labeling them with DEPC. N-terminal blocking was achieved by reacting sulfo-NHS-acetate in a molar excess of 1,000 for 1 h at room temperature. The N-terminally blocked peptides were labeled and quenched under the same conditions as those used for the intact proteins. It was important to ensure that the digested peptide concentrations were as close as possible to the intact protein concentration so that the DEPC reactivity for the free peptide and intact protein could be appropriately compared. To ensure that the concentrations were as close as possible, the proteolytic digest was prepared using a stock solution with a higher protein concentration, so that upon the various dilutions steps associated with digestion and acetylation, the final concentration of digested

peptides was identical to the intact protein concentration. Three or four replicate reactions and analyses were conducted on each protein digest.

2.2.3 Proteolytic Digestion

After the DEPC reactions, labeled β2m and hGH samples were desalted and preconcentrated in 10 mM MOPS buffer (pH 7.4) and 10 mM phosphate buffer (pH 8.0), respectively, using a centrifugal filter with a 10 kDa molecular weight cutoff (MWCO). The resulting samples were incubated with 10% (v/v) acetonitrile at 50°C for 45 min to denature the intact protein. TCEP was added in a 40-fold molar excess to reduce the disulfide bonds, and iodoacetamide was simultaneously added in an 80-fold molar excess to alkylate the reduced Cys residues. The samples were kept in the dark at room temperature for 20 min. Subsequently, the proteolytic enzyme was added to the resulting samples at a 1:10 (w/w) enzyme to substrate, and the protein was digested at 37 °C. Digestion was performed with immobilized chymotrypsin and trypsin for β2m (3 h) and immobilized trypsin for hGH (5 h). For ubiquitin, the labeled samples were desalted and preconcentrated in 25 mM ammonium bicarbonate (pH 7.8) using a centrifugal filter with a 3 kDa MWCO. 80% (v/v) acetonitrile was added to the desalted samples to denature the protein for 45 min at 50 °C. Overnight digestion was then performed with immobilized trypsin at a 1:10 (w/w) enzyme to substrate. Following digestion, the immobilized enzyme was separated from the protein digest by centrifugation. The supernatant was collected for LC-MS/MS analysis.

To prepare peptides for DEPC labeling reactions, free peptides from the same proteins were produced via the same digestion conditions, except for ubiquitin

preconcentration where 10 mM MOPS buffer (pH 7.4) was used instead of ammonium bicarbonate in order not to suppress the subsequent N-terminal blocking and DEPC labeling reactions.

2.2.4 Liquid Chromatography (LC)

For on-line LC-MS/MS analyses, a labeled sample containing approximately 0.5 µg ubiquitin peptides or 2 µg β2m or hGH peptides was loaded on a Thermo Scientific Easy-NanoLC 1000 system (Waltham, MA). Samples were loaded, trapped, and desalted on a Thermo Scientific Acclaim™ PepMap™ C18 trap column (2 cm x 75 µm, 3 µm particle size). Separation of peptides was performed using an Acclaim™ PepMap™ RSLC C18 nanocolumn (15 cm x 75 µm, 3 µm particle size; Thermo Scientific) with a flow rate of 300 nL/min. LC/MS-grade water (solvent A) and acetonitrile (solvent B), each containing 0.1% formic acid, were used as mobile phases. A linear gradient for separation of the peptides consisted of 0% B to 50% B over 60 min (β2m), 0% B to 50% B over 30 min (ubiquitin), and 0% B to 45% B over 35 min (hGH).

2.2.5 Mass Spectrometry (MS)

Mass spectra were acquired on a Thermo Scientific Orbitrap Fusion mass spectrometer (Waltham, MA). The nano-electrospray ionization (ESI) was operated using a positive mode at a needle voltage of 2100 V, and the ion transfer tube temperature was set to 325 °C. Mass spectra were acquired on an Orbitrap analyzer, with a resolution of 60,000. Tandem mass spectrometry (MS/MS) was conducted on a linear quadrupole ion trap using collision induced dissociation (CID) with a normalized collision energy of 35 %. Tandem mass spectra were collected for most intense peptides with ion abundances

58

above 5000 from each mass spectral scan. Dynamic exclusion of 30 sec was activated after 3 spectra were acquired for any given precursor ion within 5 sec. Mass detection during MS and MS/MS was performed in centroid mode to simplify data analysis.

2.2.6 Peptide Identification and Peak Quantification

Thermo Scientific Xcalibur™ software was used to visualize total ion chromatograms (TICs) from raw mass spectral data files obtained from the LC-MS/MS analyses using Thermo Scientific Orbitrap Fusion. Extracted ion chromatograms (XICs) of specific masses of unmodified and modified peptides were generated. Peptide sequencing and labeling site identification were achieved using CID tandem mass spectra. Assignments of b and y ions were performed with a mass tolerance of 0.5 Da. A custom software pipeline described previously[16, 19] and specifically designed for protein CL-MS studies with DEPC was used to identify labeling sites and quantify peak areas. The search parameters were set as follows: a precursor mass tolerance of 10 ppm, carbamidomethylation of Cys as a variable modification, DEPC modification of His, Lys, Ser, Thr, Tyr, and N-terminus as a user variable modification. DEPC modification results in mass addition of 72.0211 Da (**Figure 2.1**). For N-terminally blocked peptides, acetylation of N-termini was set as a fixed modification, while its side reaction at Lys was set as a variable modification. Acetylation of N-terminus or Lys leads to mass addition of 42.0106 Da (**Figure 2.2**).
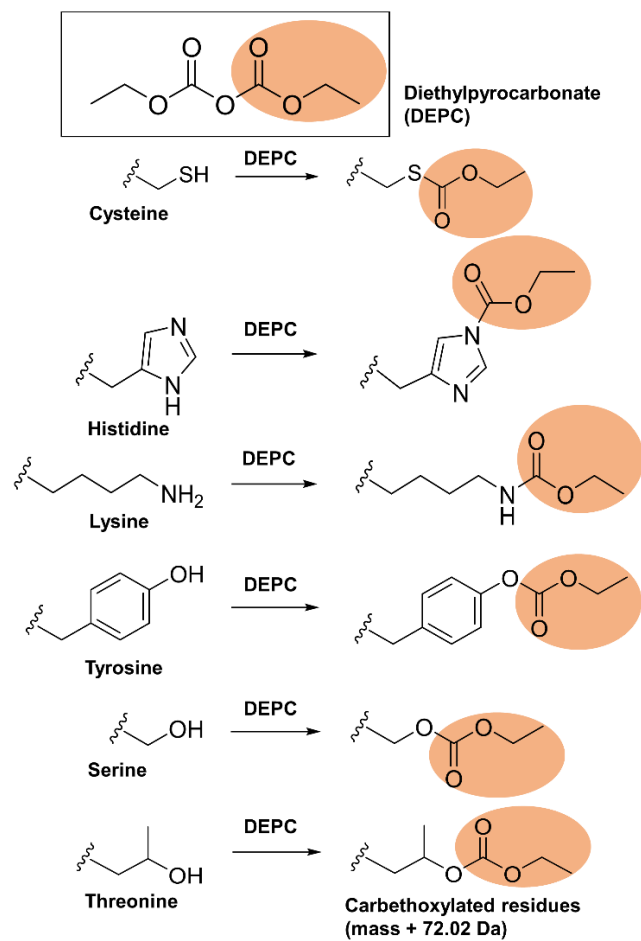
Figure 2.1: DEPC labeling reactions of Cys, His, Lys, Tyr, Ser, and Thr
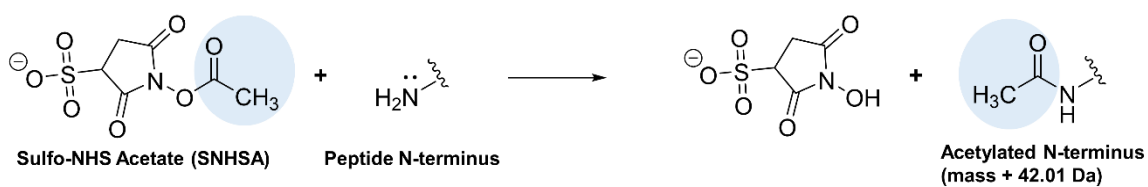
Figure 2.2: Acetylation reaction of peptide N-terminus

(a) Total ion chromatogram (TIC)

$$\% \textbf{\textit{DEPC Labeling}} = \frac{\sum Peak\ area\ of\ \text{Labeled}}{\sum \text{Labeled} + \sum \text{Unlabeled}} \times 100$$

(b) Extracted ion chromatogram (XIC): Unmodified peptide

Unmodified IQVYSR (m/z 383.2163, z=2)

(c) XIC: DEPC-modified peptides

IQVYS*R (m/z 419.2269, z=2)

IQVY*SR (m/z 419.2269, z=2)

(d) CID MS/MS: β2m (7-12) IQVYSR

NH2-I Q V Y S R-COOH

(e) CID MS/MS: β2m (7-12) IQVYS*R

NH2-I Q V Y S R-COOH

(f) CID MS/MS: β2m (7-12) IQVY*SR

NH2-I Q V Y S R-COOH

Figure 2.3: Illustration of how the DEPC modification levels are calculated.
After DEPC labeling and proteolytic digestion, (a) LC-MS analysis of the digested protein is performed. Peak areas of (b) unlabeled and (c) labeled peptides in a chromatogram are used to calculate the labeling percentage. During LC-MS, peptides are subjected to CID MS/MS. Tandem mass spectra of (d) unlabeled and (e) & (f) labeled peptides obtained at specific retention time are used for peptide sequencing and/or identification of DEPC labeled site.

61

2.2.7 Determination of Modification Percentages

DEPC modification levels (%) of each labeled residue were calculated as follows. (**Equation 2.1**)

$$\% \, DEPC \, labeling = \frac{\sum_{i=1}^{n} \sum_{z=1}^{m} A_{i,z}^{modi}}{\sum_{i=1}^{n} \sum_{z=1}^{m} A_{i,z}^{modi} + \sum_{i=1}^{n} \sum_{z=1}^{m} A_{i,z}^{unmodi}} \times 100 \qquad (2.1)$$

where $A_{i,z}^{unmodi}$ is the peak area of DEPC-unmodified peptide whose sequence (i) contains the residue of interest and possesses a certain charge state (z), and $A_{i,z}^{modi}$ is the peak area of peptide in which the residue of interest is DEPC-modified. An illustration of this calculation can be found in **Figure 2.3**.

Note that DEPC modification percentages calculated by **Equation 2.1** are only used for the relative quantitation, i.e. comparing the DEPC labeling at the same residue under different conditions (labeling on intact protein vs. protein digest). The modification levels do not reflect the absolute quantitation of modified species as the addition of a carbethoxyl group to the modified peptide and different LC solvent conditions during gradient elution of peptides result in different ionization efficiencies of the unmodified and modified peptides.

2.2.8 Solvent Accessible Surface Area (SASA) calculation

SASA of amino acid side chain was calculated from the Protein Data Bank (PDB) atomic coordinates using GETAREA 1.0 beta.[42] A probe radius of 1.4 Å which represents the van der Waals sphere of water was used in a calculation. Atomic coordinates of the 3D structures of β2m (PDB ID: 1JNJ),[43] ubiquitin (PDB ID: 1UBQ),[44] and hGH (PDB ID: 1HGU)[45] were submitted to calculate individual side-chain SASA. Note that human and

bovine ubiquitin have the same amino acid sequence. Even though bovine ubiquitin was used in experiments, the SASA and structural features considered in this study were obtained from the PDB structure of human ubiquitin (PDB ID: 1UBQ). The calculated SASA was compared to the surface area of the side chain in a Gly-X-Gly tripeptide, where X is a side chain of interest, to generate a percent ratio (%SASA).

2.2.9 DEPC Labeling and LC-MS/MS Analyses of Model Peptides

2.2.9.1 DEPC Labeling Reactions

A lyophilized powder of model peptide (Fmoc-DGXGG-amide, where X = H, K, Y, S, or T) was first reconstituted in water (His and Lys) or dimethylformamide (Ser, Thr, and Tyr), and the resulting solution of peptide was then diluted in 10 mM MOPS buffer (pH 7.4) to make a final solution. Each model peptide (50 µM) was reacted with DEPC at 37 °C for 5 min at a DEPC to protein molar ratio of 5 to 1 (His, Lys, and Tyr) or 50 to 1 (Ser and Thr). The reaction was then quenched by the addition of imidazole at a 1:50 DEPC to imidazole molar ratio.

Bradykinin and preproenkephalin peptides were reconstituted in water. A peptide mixture was prepared in 10 mM MOPS buffer (pH 7.4) containing 10 µM of each peptide. The N-termini of the peptides were acetylated before labeling them with DEPC. The N-terminal blocking was initiated by adding sulfo-NHS-acetate in a molar excess of 600, and the solution was reacted for 1 h at room temperature. Subsequently, the peptide mixture was reacted with DEPC at 37 °C for 5 min at a molar excess of 50. The reaction was then quenched by the addition of imidazole.

2.2.9.2 Liquid Chromatography (LC)

Following quenching online LC-MS/MS analyses were conducted on a Thermo Scientific Ultimate 3000 HPLC (Waltham, MA). LC separation was performed on a Thermo Scientific Acclaim™ PepMap™ RSLC C18 column (15 cm x 300 μm, 2 μm particle size). LC/MS-grade water (solvent A) and acetonitrile (solvent B), each containing 0.1% formic acid, were used as mobile phases. A flow rate of 4 μL/min was used, and 1 μL of sample was first loaded and desalted at 5%B during the first 5 min. An isocratic elution mode at 40%B over 15 min was then applied to separate unmodified and modified peptides.

For a mixture of bradykinin and preproenkephalin peptides, 5 μL of sample was first loaded and desalted during the first 5 min. After that a linear gradient was increased from 5%B to 50%B over 50 min to separate unmodified and modified peptides.

2.2.9.3 Mass Spectrometry (MS)

Mass spectra from the online LC-MS/MS were acquired on a Bruker AmaZon quadrupole ion trap (Billerica, MA). The electrospray needle voltage was operated using a positive mode at ~4 kV, and the capillary temperature was set to 300 ^{O}C. Tandem spectra were collected for the top intense species with ion abundances above 1,000 from each mass spectrum. Collisional-induced dissociation (CID) was conducted with a ramp energy 60% to 180% of 1.5 V amplitude.

2.2.9.4 Peptide Identification and Peak Quantification

Bruker Compass™ Data Analysis software was used to reconstruct extracted ion chromatograms (XICs) of unmodified and modified peptides. Peptide identification and

64

peak quantification were performed in a manual manner using tandem spectra and mass spectral peak areas, respectively. DEPC modification levels (%CL) of each labeled residue were calculated using **Equation 2.1** (see **Section 2.2.7**).

2.2.10 LC-MS/MS Analyses of Chymotryptic Protein Digests

Chymotryptic digests obtained from the DEPC reactions on β2m intact proteins underwent LC-MS/MS analyses using a Thermo Scientific Dionex Ultimate 3000 HPLC (Waltham, MA) and a Bruker AmaZon quadrupole ion trap mass spectrometer (Billerica, MA).

2.2.10.1 Liquid Chromatography (LC)

Chymotryptic digests obtained from the DEPC reactions on β2m intact proteins were analyzed using a Thermo Scientific Dionex Ultimate 3000 HPLC (Waltham, MA). 5 μL of sample was loaded and separation of peptides was performed on a Thermo Scientific Acclaim™ PepMap™ RSLC C18 column (15 cm x 300 μm, 2 μm particle size). A flow rate of 4 μL/min was used, and desalting was performed at 5%B during the first 5 min after sample injection. A 50-min linear gradient was applied with %B increased from 5%B to 50%B to separate peptides. The gradient was finally elevated to and held at 95%B to flush a column.

2.2.10.2 Mass Spectrometry (MS)

Mass spectra were acquired on a Bruker AmaZon quadrupole ion trap (Billerica, MA). The electrospray needle voltage was kept at ~4 kV (positive mode), and the capillary temperature was set to 300°C. The top intense peptides with ion abundances above 1,000

from each mass spectrum were selected for MS/MS acquisition. CID was conducted at 1.5 V amplitude with a ramp energy 60% to 180%. Because of the large number of measured peaks, active exclusion of 0.5 min was activated after 2 spectra were acquired for any given precursor ion.

2.2.10.3 Peptide Identification and Peak Quantification

See **Section 2.2.9.4**

2.3 Results and Discussion

2.3.1 DEPC Labeling Reactivity of Weakly Nucleophilic Residues

Solvent accessible nucleophilic side chains (Cys, His, Lys, Thr, Tyr, Ser) and N-termini of proteins can be modified by DEPC via a nucleophilic substitution at a carbonyl group (**Figure 2.1**) to yield a mass addition of +72.021 Da. The specific modification sites in proteins can be identified and semi-quantified using MS-based bottom-up sequence analysis (**Figure 2.3**). As indicated in the introduction, Ser, Thr, and Tyr residues in proteins are readily modified by DEPC. Indeed, from four proteins studied previously, including equine myoglobin,[29] human β2m,[16, 29] hGH,[16] murine IgG1,[16] and rituximab,[19] which have a total of 1,770 amino acids, 154 out of 198 Ser residues, 100 out of 136 Thr residues, and 48 out of 71 Tyr residues were labeled in intact proteins at relative levels ranging from 0.01 to 90%. In stark contrast, these residues in small peptides are relatively unreactive after 5 min with DEPC at a DEPC:peptide molar ratio of 5:1, which is the commonly used time and labeling ratio for intact proteins. As an example, Ser and Thr residues in the model peptide sequence of Fmoc-DGXGG-amide, where X = Ser or Thr, are completely unreactive (**Figure 2.4a and b**), whereas the same peptide with Tyr (i.e.

Fmoc-DGYGG-amide) produces a relative modification extent of 5% (**Figure 2.4c**). By comparison, the same peptide sequence with His (Fmoc-DGHGG-amide) and Lys (Fmoc-DGKGG-amide) are much more reactive, having modification percentages of 88% and 43%, respectively, under the same labeling conditions (**Figure 2.4d and e**).
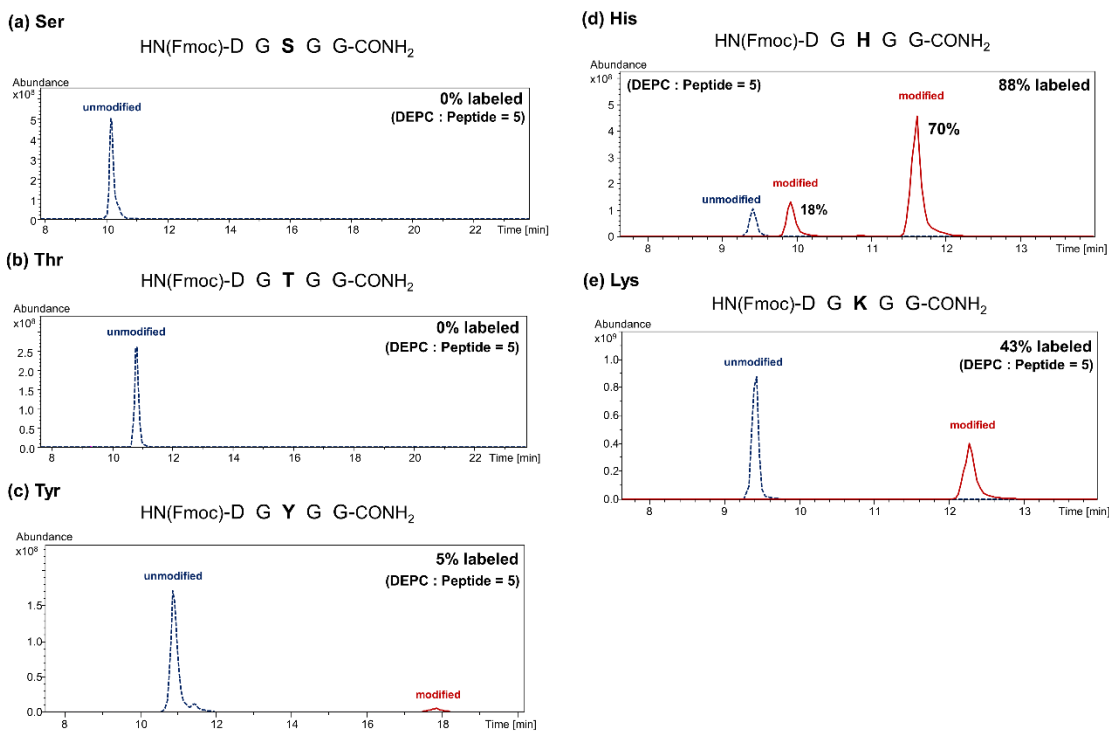
Figure 2.4: DEPC labeling of model peptides.
Extracted ion chromatograms (XICs, left) of the +1 ions of the unmodified and DEPC-modified peptides Fmoc-DGXGG-amide, where X is a DEPC modifiable residue (a) Ser, (b) Thr, (c) Tyr, (d) His, or (e) Lys, after allowing the peptide to react with DEPC at a DEPC:peptide molar ratio of 5:1. Modification percentages are calculated from peak areas in XICs. Tandem mass spectra acquired after CID of the labeled precursor ions are used to confirm the site of modification. MS/MS assignments of the modified peptides are shown in **Figure 2.5**. Note that two chromatographic peaks for the labeled peptide in (d) are observed because the His side chain has two nitrogens that are separately labeled to produce isomers that can be separated by LC.
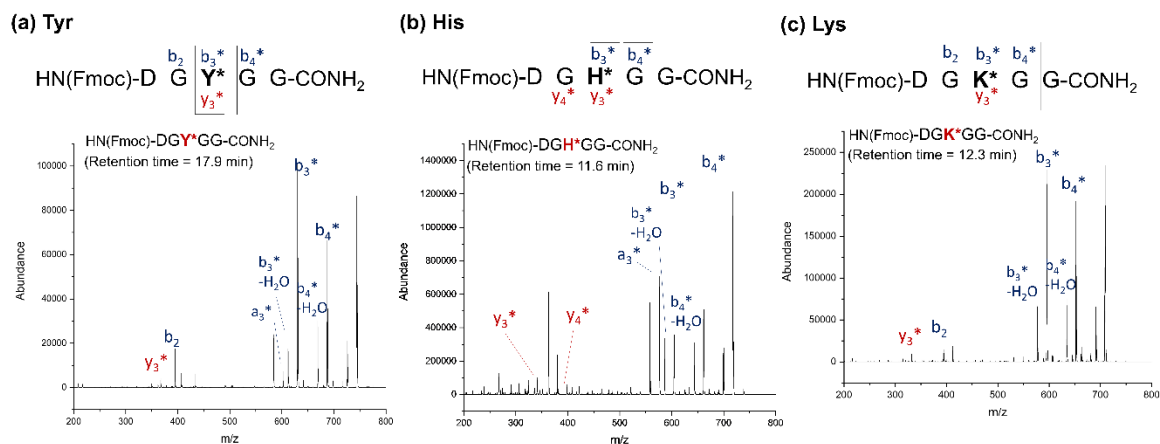
Figure 2.5: DEPC labeling of model peptides.
MS/MS assignments for nucleophilic residues that are modified in model peptides Fmoc-DGXGG-amide, where X is a DEPC modifiable residue (a) Tyr, (b) His, or (c) Lys, after allowing the peptide to react with DEPC at a molar ratio of 5:1 (DEPC:peptide).
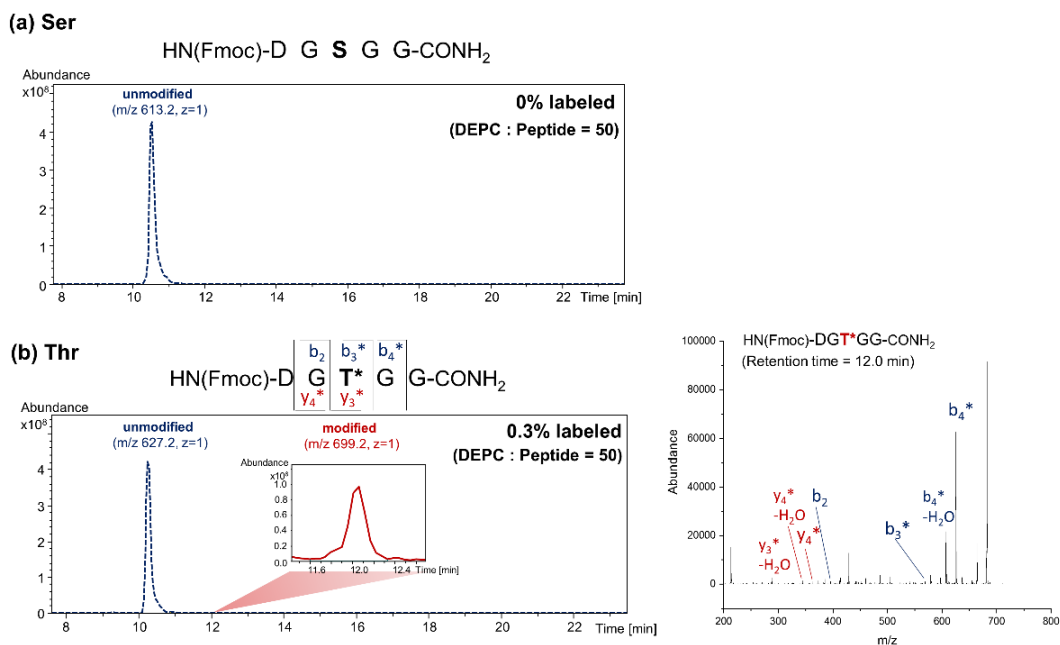
Figure 2.6: DEPC labeling of model peptides.
Extracted ion chromatograms (XICs, left) of the +1 ions of the unmodified and DEPC-modified peptides Fmoc-DGXGG-amide, where X is a DEPC modifiable (a) Ser or (b) Thr, after allowing the peptide to react with DEPC at a DEPC:peptide molar ratio of 50:1. Modification percentages are calculated from peak areas in XIC. Tandem mass spectra (right) are used to confirm the site of modification. MS/MS assignment of the modified peptide is shown above the XICs.

Modification of Ser and Thr residues in this peptide is still negligible (0% and 0.3%, respectively) even with a 50-fold molar excess of DEPC (**Figure 2.6**). The same lack of reactivity at Ser, Thr, and Tyr residues is observed in other peptides too, including bradykinin and preproenkephalin, even when the N-terminus is blocked to reduce competitive reaction with DEPC (**Figure 2.7**).
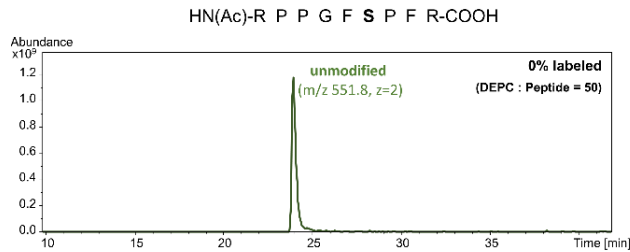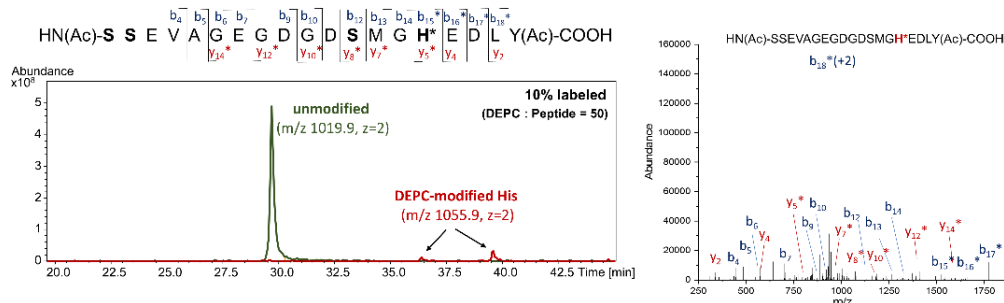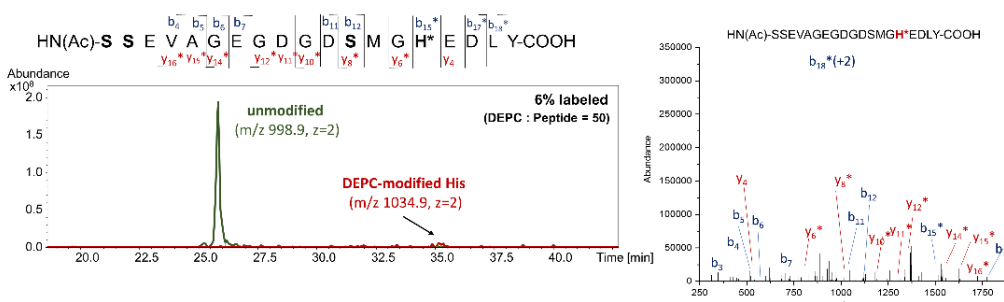
Figure 2.7: DEPC labeling of bradykinin and preproenkephalin peptides.
Extracted ion chromatograms (XICs, left) of the +2 ions of the unmodified and DEPC-modified versions of the N-terminally blocked peptides (a) bradykinin, (b) & (c) preproenkephalin, after reacting the peptide mixture with DEPC at a DEPC to peptide molar ratio of 50 to 1. Tyr and Ser residues in these peptides are unreactive with DEPC even when the other reactive site (N-terminus) are blocked via acetylation with sulfo-NHS-acetate in these peptides (**Figure 2.2**). Tandem mass spectra (right) were used to confirm the site of modification. MS/MS assignments of the modified peptides are shown above the XICs.

2.3.2 Influence of Higher-Order Structure on the Covalent Labeling Based Structural Analysis of Proteins

One possible reason for the significant reactivity of Ser, Thr, and Tyr residues in proteins but not in free peptides could be due to local sequence effects that are not fully accounted for in the model peptides. To test this possibility, we compared the labeling reactivity of intact proteins with their peptide fragments as produced by proteolysis (**Figure 2.8**). In the first scenario (**Figure 2.8a**), three proteins, β2m, ubiquitin, and hGH, were reacted with DEPC under native conditions (37 °C, pH ~ 7.5) using a DEPC molar excess of 4 or 5, which has previously been found to prevent labeling-induced structural perturbations to proteins during the labeling reaction.[12, 16, 27] In the second scenario (**Figure 2.8b**), peptide fragments from the same three proteins were first produced via the same digestion conditions and then reacted with DEPC under the same labeling conditions. In this second set of experiments, the N-termini of the peptide fragments were blocked via acetylation with sulfo-NHS-acetate (**Figure 2.2**) before labeling them with DEPC because many new N-termini are produced upon protein digestion and these newly-formed N-termini could potentially outcompete Ser, Thr, and Tyr residues.

Upon comparing the modification extents of Ser, Thr, and Tyr residues under the two DEPC reaction scenarios shown in **Figure 2.8**, we find that these residues are rarely modified in free peptides but are modified in the intact proteins (**Table 2.1**). When intact proteins are reacted with DEPC, 18 out of 67 Ser, Thr, and Tyr residues (9 in β2m, 2 in ubiquitin, and 7 in hGH) are labeled at levels ranging from 0.02% to 64%. When the same proteins are first digested into peptide fragments, N-terminally blocked, and then reacted with DEPC under the same labeling conditions, only two residues in the peptide fragments are found to be modified, and these are modified at very low levels (< 3%) (see MS/MS

data in **Figure 2.9**). Moreover, only one of these residues (Tyr136 in hGH) is modified in the intact protein. All modifiable Ser, Thr, and Tyr residues along with all the modified His and Lys sites (His, Lys, Ser, Thr, Tyr) and their labeling levels in β2m, ubiquitin, and hGH are listed in **Tables A.1, A.2, and A.3 in Appendix A**, respectively. Together, the results indicate that the weak nucleophilic side chains, without the influence of protein higher-order structure, are poorly reactive with DEPC. These observations imply that the microenvironment around specific Ser, Thr, and Tyr residues in proteins tune their reactivity such that they can only be labeled in the 3D context of the protein.
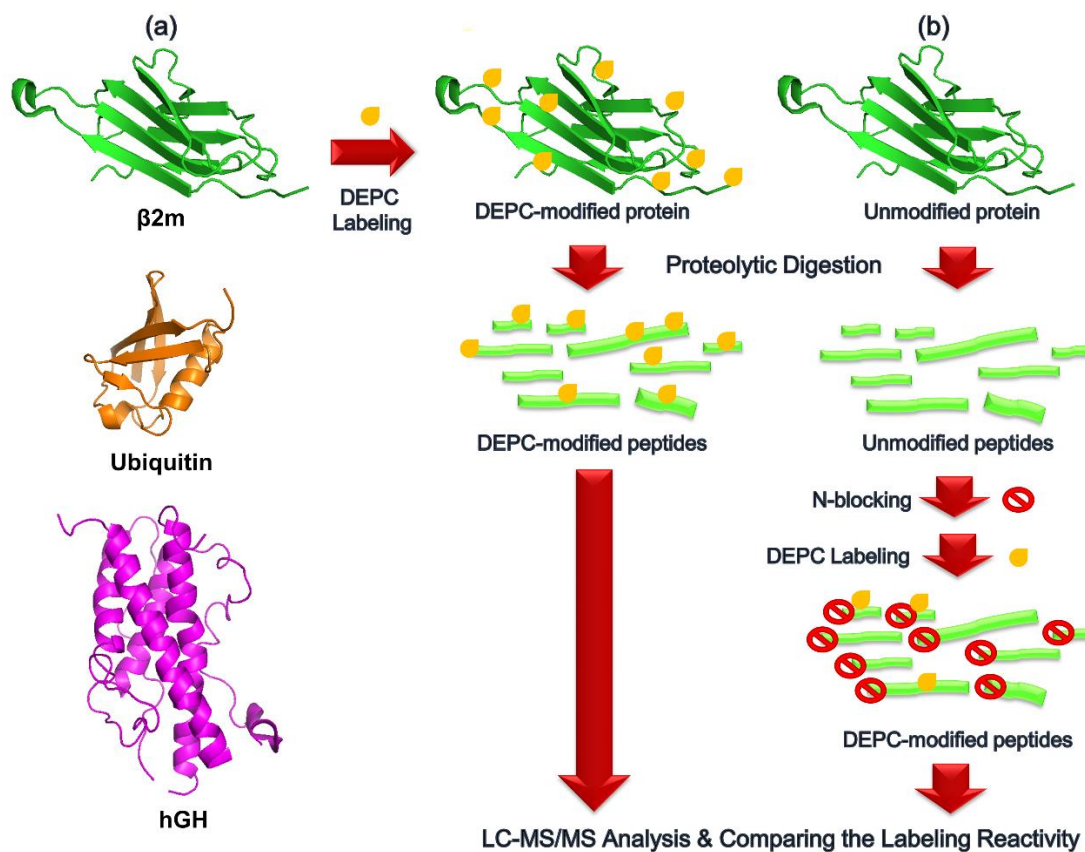
Figure 2.8: Experimental workflow to evaluate the influence of higher-order structure (i.e. microenvironment effects) on covalent labeling by comparing the DEPC labeling reactivity of (a) intact proteins with that of (b) their N-terminally acetylated peptide fragments.

71

Table 2.1: DEPC modification percentages of Ser, Thr, and Tyr residues in intact proteins and peptide fragments for β2m, ubiquitin, and hGH.

Each experiment was performed in triplicate or quadruplicate (n = 3 or 4). Error bars shown in a table are standard deviations. Listed are side-chain solvent accessible surface areas (SASA) and brief details of the microenvironment around each residue. The presence (✔) or absence (✘) of nearby charged polar contact(s) and hydrophobic residue(s) within 4 Å and 6 Å, respectively, are presented in a table. (+) and (-) represent positively charged and negatively charged polar contacts, respectively.

Residue	DEPC modification (%)		SASA (%)*	Microenvironment	
	Intact protein	Peptide fragments		Charged polar contact(s) [within 4 Å]	Hydrophobic residue(s) [within 6 Å]
β-2-microglobulin (β2m)					
T4	59 ± 5 (N-term & T4)	N.D.[‡]	81.5	✘	✔
Y10	15 ± 2	N.D.	41.6	✘	✔
S11	3.0 ± 0.5	N.D.	10.3	✘	✔
S20	N.D.	2.6 ± 0.6	61.6	H-bonding with Glu (-)	✘
S33	1.6 ± 0.4	N.D.	70.1	✘	✔
S55	0.6 ± 0.3	N.D.	55.3	✘	✔
S57	1.8 ± 0.9	N.D.	34.9	H-bonding with Asp (-)	✔
S61	10 ± 1	N.D.	28.1	H-bonding with Arg (+)	✔
Y63	0.3 ± 0.2	N.D.	27.7	✘	✔
Y66	0.3 ± 0.2	N.D.	7.5	✘	✔
Ubiquitin					
T7	13 ± 2	N.D.	14.2	✘	✔
S65	2.4 ± 0.2	N.D.	9.2	✘	✔
Human growth hormone (hGH)					
T28	1.3 ± 0.6	N.D.	10.7	H-bonding with Asp (-)	✔
Y29	0.2 ± 0.1	N.D.	4.1	H-bonding with Lys (+)	✔
Y36	0.4 ± 0.3	N.D.	37.2	✘	✔
S56	0.02 ± 0.01	N.D.	20.8	✘	✔
S58	64 ± 4	N.D.	3.3	✘	✔
T136	0.15 ± 0.09	0.1 ± 0.1	35.9	H-bonding with Arg (+)	✔
Y144	0.66 ± 0.05	N.D.	31.1	✘	✔

* SASA calculation is explained in detail in **Section 2.2.8**.
‡ N.D. = not detected (less than DEPC labeling threshold of 0.01%[19])
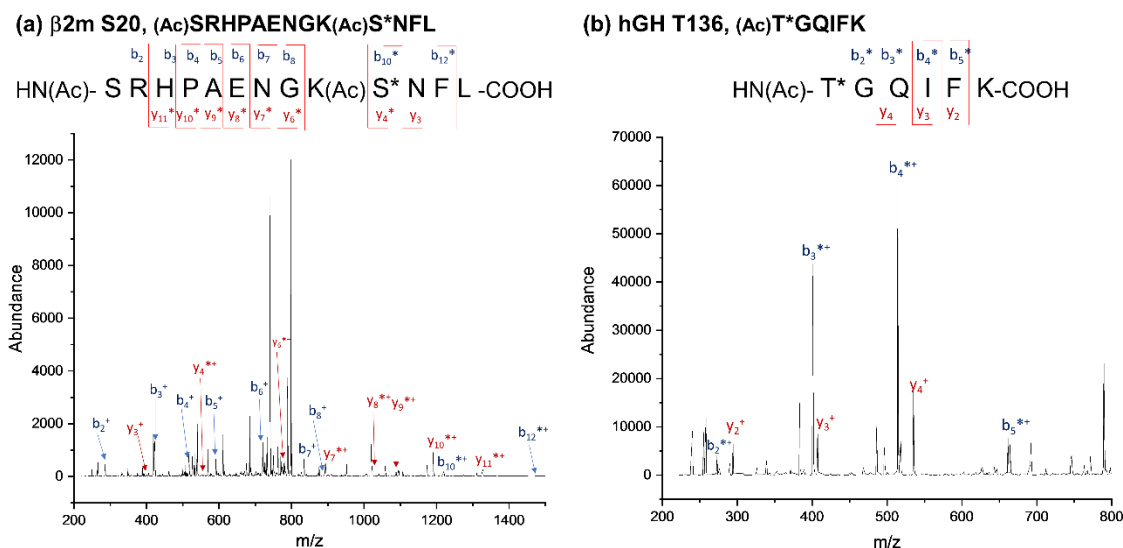
Figure 2.9: MS/MS assignments for weakly nucleophilic residues that are modified at peptide level.

When reacting the N-terminally blocked peptides at DEPC to peptide molar ratios of 4 to 1 or 5 to 1, the reactivity of the Ser, Thr, and Tyr residues in free peptides are found to be lower than in the intact proteins. Only two residues in three model proteins are found to be modified. This Figure S-hows tandem mass spectra acquired after CID of (a) HN(Ac)-SRHPAENGK(Ac)S*NFL-COOH peptide from chymotryptic digest of β-2-microglobulin (β2m), where Ser20 is the DEPC modification site, and (b) HN(Ac)-T*GQIFK-COOH peptide from tryptic digest of human growth hormone (hGH), where Thr136 is the modification site. Product ions with DEPC-modified residue have mass addition of 72.02 Da while acetylated N-terminus or Lys contributes to mass addition of 42.01 Da.

2.3.3 Identifying Structural Features that Tune the Reactivity of Weak Nucleophiles

To identify the protein structural factors that affect the reactivity of Ser, Thr, and Tyr residues, we considered SASA and microenvironment effects such as polar and hydrophobic interactions. SASAs of the side chains were calculated using GETAREA,[42] and insight into the local microenvironment around each side chain was obtained from the protein's 3D structure using the molecular visualization program PyMOL.[46] SASA is typically considered an important criteria for residue reactivity in CL-MS, but we find a relatively poor correlation between SASA and Ser, Thr, and Tyr reactivity. Typically, side chains with %SASA below 20% are considered buried;[42, 47] however, we do not see a

significant difference in the number of Ser, Thr, and Tyr residues having %SASA values above 20% between the sets of labeled and unlabeled residues in the three proteins. Indeed, of the 18 Ser, Thr, and Tyr residues labeled in the proteins, only 11 (or 61%) have SASA percentages above 20%, whereas in the unlabeled Ser, Thr, and Tyr residues, 34 of the 44 (or 77%) have SASA percentages above 20% (**Figure 2.10**). Hence, solvent accessibility does not significantly influence the reactivity of these weakly nucleophilic residues, which is consistent with the previously observed poor correlation between SASA and reactivity of these residues.[16,19,27,29] We also considered if protein secondary structure affects the reactivity of these side chains, but we do not find a preference for α-helix, β-strand, or random coil structure among the residues that are found to be labeled (see **Tables A.1** to **A.3**).
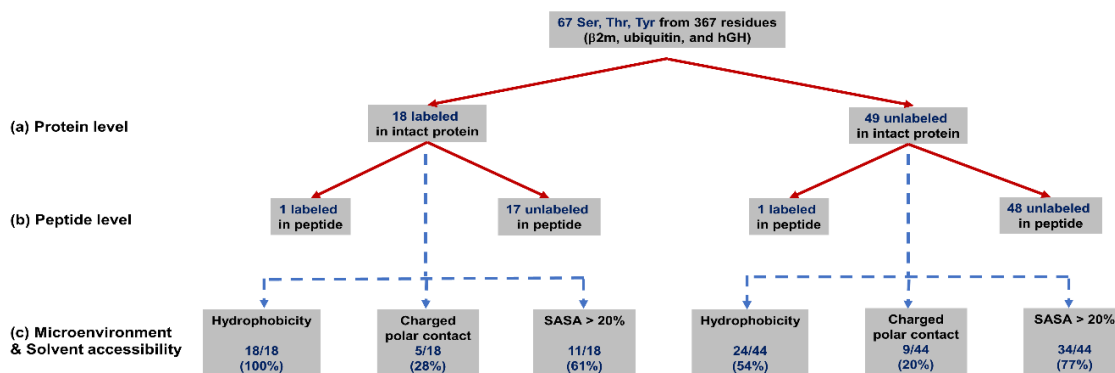
Figure 2.10: Flow chart summarizing the covalent labeling results and the structural features of weakly nucleophilic residues in β2m, ubiquitin, and hGH.
The numbers of these residues that are found to be labeled or unlabeled in (a) intact proteins and (b) peptide fragments are shown in the chart. Structural features that could affect the labeling reactivity of these residues, such as the presence of nearby hydrophobic residues or charged residues and the solvent accessible surface area (SASA) above 20%, are indicated in (c). The fraction of labeled or unlabeled residues that have each specific structural feature is shown in the chart. For the unlabeled Ser, Thr, and Tyr residues, structural features are investigated for 44 out of 49 residues as 5 of the residues in hGH are not resolved in the crystal structure (PDB ID: 1HGU).[45]

The microenvironment of a side chain can influence the acid/base characteristics of side chains because of polar interactions (charge-charge interactions, charge-dipole interactions, dipole-dipole interactions)[32, 33] or hydrophobic effects (i.e. Born or desolvation effect).[32, 34-37] In chemical labeling experiments conducted with monofunctional NHS esters, O-acylation of Ser, Thr, and Tyr was suggested to be dependent on protein/peptide conformation and intramolecular interactions.[33, 48-50] Hydrogen bonding and other non-covalent interactions can cause changes in a side chain's protonation state and thus affect its nucleophilicity.[33, 51, 52] To explore this possible effect, we examined adjacent charged residues (Arg, Lys, Asp, Glu) and hydrogen bonding interactions between these residues and Ser, Thr, and Tyr. Nearby positively-charged residues (Arg, Lys) can presumably stabilize deprotonated forms of these weak nucleophiles, increasing their nucleophilicity, and negatively-charged residues (Asp, Glu) that form an ionic hydrogen bond with the hydroxyl group of Ser, Thr, and Tyr side chains might also increase their nucleophilicity (**Figure 2.11**). Only interactions between these acidic and basic groups and Ser, Thr, or Tyr residues within 4 Å were considered as this is a common distance over which such interactions typically have an effect.[53-56] From our results, though, there seems to be very little correlation between nearby charged residues and Ser, Thr, and Tyr labeling in proteins. Only 28% of the labeled Ser, Thr, and Tyr residues have nearby charged polar contacts, which is very similar to the 20% of the unlabeled residues that have nearby charged residues (**Figure 2.10**). Perhaps, the charged polar contacts cannot sufficiently tune the nucleophilicity of Ser, Thr, and Tyr residues to a degree that it can enhance reactivity of these side chains.
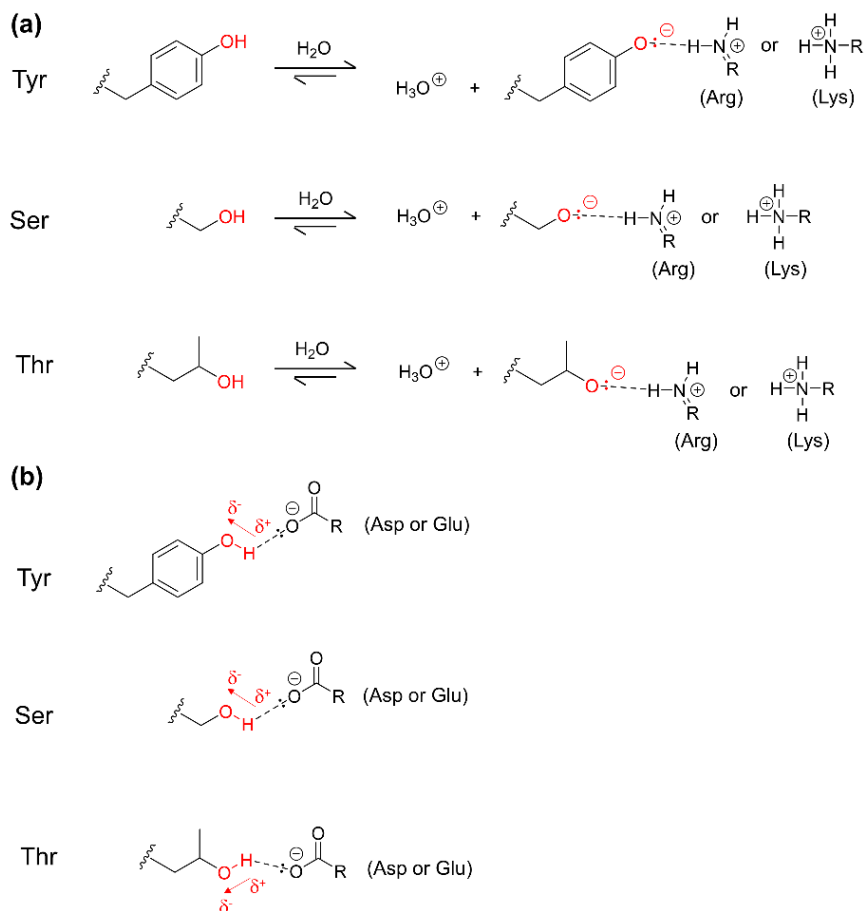
Figure 2.11: Molecular schemes showing how both positively-charged and negatively-charged residues could affect Ser, Thr, and Tyr.
(a) Deprotonated forms of these weak nucleophiles can presumably be stabilized by nearby positively-charged residues (Arg, Lys), increasing their nucleophilicity, and (b) the hydroxyl group of Ser, Thr, and Tyr side chains can form an ionic hydrogen bond with nearby negatively-charged residues (Asp, Glu), which might also increase their nucleophilicity.

The presence of nearby hydrophobic residues was also considered as such groups might increase the local concentration of DEPC, thereby enhancing the reactivity of Ser, Thr, and Tyr residues towards chemical labeling.[33, 39] Even though DEPC is readily soluble in aqueous solutions at moderate concentrations (up to 40 mM) which is useful for labeling proteins,[57] this molecule is quite hydrophobic and has limited solubility in water (~ 0.1%). We hypothesized that perhaps the hydrophobicity of a protein's local structure may

enhance Ser, Thr, and Tyr reactivity by increasing DEPC's local concentration. To study

the potential influence of a hydrophobic microenvironment, we searched for the presence

of hydrophobic side chains (Ala, Ile, Leu, Met, Pro, Val) and aromatic sidechains (Phe,

Trp, Tyr) within 6 Å. A distance of 6 Å was chosen because of the molecular dimensions

of DEPC, which is a symmetrical molecule having two electrophilic carbonyls, one of

which could be positioned close enough to the hydroxyl group of Ser, Thr, or Tyr if it

interacted with a hydrophobic group 6 Å away (**Figure 2.12**). Interestingly, each of the

Ser, Thr, and Tyr residues that are found to be labeled in these model proteins sits close to

at least one hydrophobic side chain (**Figures 2.10** and **2.13**). In contrast, only half of the

unlabeled Ser, Thr, and Tyr residues have nearby hydrophobic residues **(Figure 2.10)**.

These results strongly suggest that hydrophobicity is an important factor that enhances the

labeling reactivity of Ser, Thr, and Tyr residues in intact proteins. We propose that elevated

local DEPC concentrations near these hydrophobic groups explain this enhanced reactivity.

An analogous idea was suggested by Zenobi *et al*. to explain the enhanced cross-linking

reactivity of Ser, Thr, and Tyr with bifunctional NHS esters to nearby Lys or His residues.[33]

They argued that the high local concentration of the NHS cross-linker after formation of

an N-acylated intermediate at His or Lys would allow the O-acylation of weak nucleophiles

to occur at higher reaction yield. Similarly, a recent study performed by Tetin and

coworkers suggested that a fluorescent label AlexaFluor488 was predominantly conjugated

to Lys residues located in the proximity of Trp and Tyr residues.[39]
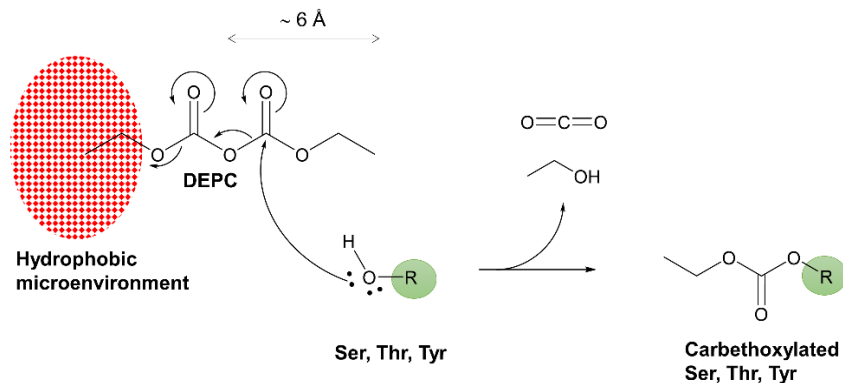
Figure 2.12: Mechanism of the nucleophilic substitution at the carbonyl group for DEPC covalent labeling.

Our study indicates that hydrophobicity contributes to the higher labeling reactivity of certain Ser, Thr, and Tyr residues, due to the increase in local concentration of DEPC near these residues.
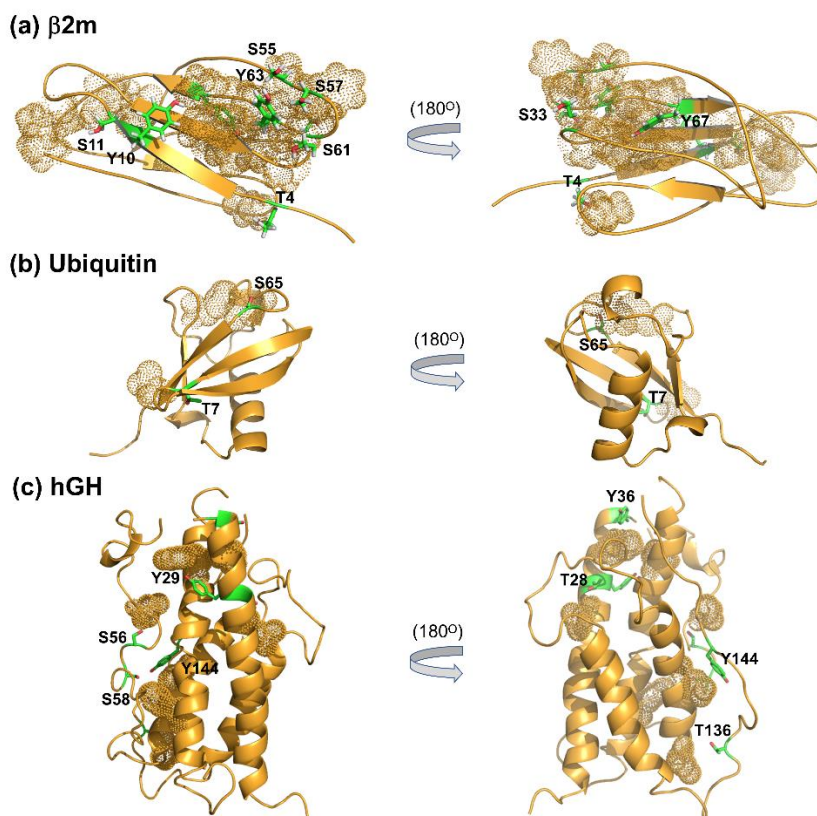
Figure 2.13: Ser, Thr, and Tyr residues that are found to be labeled by DEPC in the proteins (a) β2m, (b) ubiquitin, and (c) hGH.

Side chains of the labeled residues are shown in green. Dotted spheres represent the nearby hydrophobic side chains that contribute to the higher DEPC labeling reactivity of these residues.

To further investigate the idea that the increased local DEPC concentration can drive the reactivity of Ser, Thr, and Tyr residues, we labeled β2m, ubiquitin, and hGH peptides at 10-fold higher DEPC concentrations (i.e. 50:1 DEPC:peptide). Whereas only two Ser, Thr, or Tyr residues out of 67 are labeled at low DEPC concentration (i.e. 4:1 or 5:1 DEPC:peptide), we find seven residues are modified at the higher DEPC concentration (**Table 2**). Obviously, this number of labeled Ser, Thr, and Tyr residues is still lower than in intact proteins, suggesting that the environment of the folded protein is still probably important for forming the necessary hydrophobic microenvironment to increase the DEPC local concentration sufficiently.

The importance of nearby hydrophobic groups for enhancing Ser, Thr, or Tyr reactivity may explain why the SASA of side chains does not correlate well with extent of reactivity for these residues. This fact can be illustrated by considering the distribution of %SASA values for both labeled and unlabeled Ser, Thr, and Tyr residues in the three proteins (**Figure 2.14**). Most of the labeled residues have SASA percentages between 5% and 45%, and only 17% of the residues have SASA values above 45%. In contrast, almost half of the unlabeled residues have SASA above 45%. As expected, we see a decrease in the average number of nearby hydrophobic sites when SASA increases (red bars in **Figure 2.14**). Between SASA values of 25% and 45%, there is a 'sweet spot' where many Ser, Thr, and Tyr residues are found to be labeled. In this range, the Ser, Thr, and Tyr residues have enough nearby hydrophobic sites, yet are still solvent accessible enough for DEPC to react with them. The unlabeled side chains are too exposed to the aqueous environment, and lack nearby hydrophobic groups to enhance their reactivity. Together, our findings indicate that a hydrophobic microenvironment influences Ser, Thr, and Tyr reactivity

towards DEPC labeling in intact proteins more than SASA. Additional studies are needed to further verify this effect, but the reactivity of specific Ser, Thr, and Tyr residues could be used to provide insight into the local microenvironment around these residues and could be useful in CL-MS based protein structural prediction. For example, the reactivity of specific Ser, Thr, and Tyr residues could indicate the presence of nearby hydrophobic residues, and this information could be used to improve confidence in certain structural models.
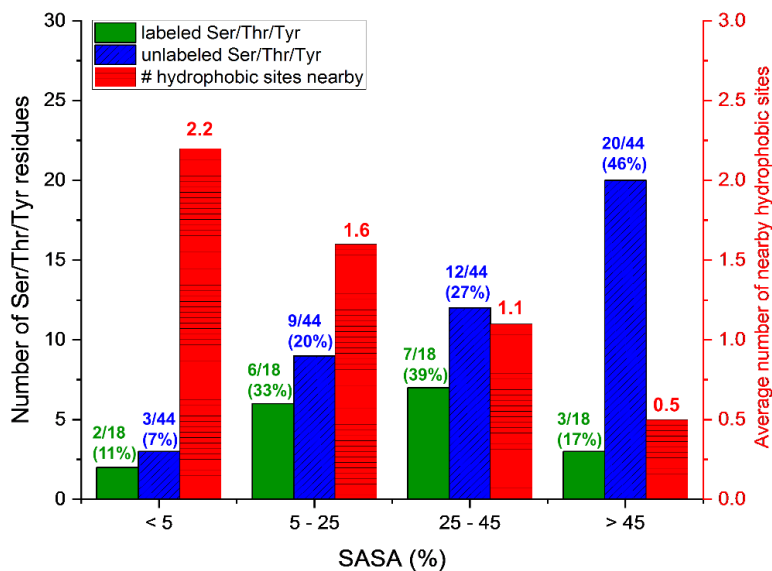
Figure 2.14: %SASA values and the average number of nearby hydrophobic sites for both labeled and unlabeled Ser, Thr, and Tyr residues in the three proteins
Histograms illustrating the distribution of %SASA values for both labeled (green) and unlabeled (blue) Ser, Thr, and Tyr residues in the proteins β2m, ubiquitin, and hGH. (See left axis.) Also included is the distribution of the average number of nearby hydrophobic sites (red) per Ser/Thr/Tyr residue in each bin of SASA values. (See right axis.)

Table 2.2: DEPC modification percentages of weakly nucleophilic residues in intact proteins and peptide fragments of β2m, ubiquitin, and hGH.
DEPC to protein/peptide molar ratios used in the study are specified in the column headers. Each experiment was performed in triplicate or quadruplicate (n = 3 or 4). Error bars shown in a table are standard deviations.

Residue	DEPC modification (%)		
	Intact protein (CL at 4X/5X molar excess)	Peptide fragments (CL at 4X or 5X molar excess)	Peptide fragments (CL at 50X molar excess)
β-2-microglobulin (β2m)			
T4	59 ± 5 (N-term & T4)	N.D. [‡]	0.02 ± 0.01
Y10	15 ± 2	N.D.	3.9 ± 0.8
S11	3.0 ± 0.5	N.D.	0.02 ± 0.01
S20	N.D.	2.6 ± 0.6	N.D.
S33	1.6 ± 0.4	N.D.	N.D.
S55	0.6 ± 0.3	N.D.	N.D.
S57	1.8 ± 0.9	N.D.	N.D.
S61	10 ± 1	N.D.	N.D.
Y63	0.3 ± 0.2	N.D.	N.D.
Y66	0.3 ± 0.2	N.D.	N.D.
Y67	N.D.	N.D.	1.3 ± 0.6
Ubiquitin			
T7	13 ± 2	N.D.	N.D.
S65	2.4 ± 0.2	N.D.	N.D.
Human growth hormone (hGH)			
T28	1.3 ± 0.6	N.D.	N.D.
Y29	0.2 ± 0.1	N.D.	N.D.
Y36	0.4 ± 0.3	N.D.	N.D.
S56	0.02 ± 0.01	N.D.	N.D.
S58	64 ± 4	N.D.	N.D.
S107	N.D.	N.D.	0.2 ± 0.1
T136	0.15 ± 0.09	0.1 ± 0.1	0.08 ± 0.02
Y144	0.66 ± 0.05	N.D.	0.02 ± 0.03

[‡] N.D. = not detected (less than DEPC labeling threshold of 0.01% [19])

2.4 Conclusion

DEPC is capable of modifying weakly nucleophilic residues such as Ser, Thr, and Tyr in intact proteins, allowing for higher structural resolution. The same reagent, however, is essentially unreactive with Ser, Thr, and Tyr residues in small peptides, suggesting that the 3D context of the protein's structure is able to tune the reactivity of these residues with DEPC. By comparing the DEPC reactivity of three intact proteins and their proteolytic fragments, we are able to identify the chemical environments that most influence the reactivity of these side chains in intact proteins. Interestingly, we find evidence that the presence of nearby hydrophobic residues increases the DEPC reactivity of Ser, Thr, and Tyr residues in intact proteins. This finding suggests that the reactivity of certain Ser, Thr, and Tyr residues could be used to indicate the presence of nearby hydrophobic groups and thus could be used as constraints in protein structure prediction.

2.5 References

1.   Stocks, B. B.; Konermann, L., Structural Characterization of Short-Lived Protein Unfolding Intermediates by Laser-Induced Oxidative Labeling and Mass Spectrometry. *Anal. Chem.* **2009,** *81* (1), 20-27.
2.   Vahidi, S.; Stocks, B. B.; Liaghati-Mobarhan, Y.; Konermann, L., Submillisecond Protein Folding Events Monitored by Rapid Mixing and Mass Spectrometry-Based Oxidative Labeling. *Anal. Chem.* **2013,** *85* (18), 8618-8625.
3.   Chen, J.; Rempel, D. L.; Gross, M. L., Temperature Jump and Fast Photochemical Oxidation Probe Submillisecond Protein Folding. *J. Am. Chem. Soc.* **2010,** *132* (44), 15502-15504.
4.   Chen, J.; Rempel, D. L.; Gau, B. C.; Gross, M. L., Fast Photochemical Oxidation of Proteins and Mass Spectrometry Follow Submillisecond Protein Folding at the Amino-Acid Level. *J. Am. Chem. Soc.* **2012,** *134* (45), 18724-18731.
5.   Srikanth, R.; Mendoza, V. L.; Bridgewater, J. D.; Zhang, G.; Vachet, R. W., Copper Binding to β-2-Microglobulin and Its Pre-Amyloid Oligomers. *Biochemistry* **2009,** *48* (41), 9871-9881.
6.   Mendoza, V. L.; Antwi, K.; Barón-Rodríguez, M. A.; Blanco, C.; Vachet, R. W., Structure of the Preamyloid Dimer of β-2-Microglobulin from Covalent Labeling and Mass Spectrometry. *Biochemistry* **2010,** *49* (7), 1522-1532.

7.   Mendoza, V. L.; Barón-Rodríguez, M. A.; Blanco, C.; Vachet, R. W., Structural Insights into the Pre-Amyloid Tetramer of β-2-Microglobulin from Covalent Labeling and Mass Spectrometry. *Biochemistry* **2011,** *50* (31), 6711-6722.

8.   Klinger, A. L.; Kiselar, J.; Ilchenko, S.; Komatsu, H.; Chance, M. R.; Axelsen, P. H., A Synchrotron-Based Hydroxyl Radical Footprinting Analysis of Amyloid Fibrils and Prefibrillar Intermediates with Residue-Specific Resolution. *Biochemistry* **2014,** *53* (49), 7724-7734.

9.   Li, K. S.; Rempel, D. L.; Gross, M. L., Conformational-Sensitive Fast Photochemical Oxidation of Proteins and Mass Spectrometry Characterize Amyloid Beta 1–42 Aggregation. *J. Am. Chem. Soc.* **2016,** *138* (37), 12090-12098.

10.   Zhang, H.; Gau, B. C.; Jones, L. M.; Vidavsky, I.; Gross, M. L., Fast Photochemical Oxidation of Proteins for Comparing Structures of Protein−Ligand Complexes: The Calmodulin−Peptide Model System. *Anal. Chem.* **2011,** *83* (1), 311-318.

11.   Manzi, L.; Barrow, A. S.; Scott, D.; Layfield, R.; Wright, T. G.; Moses, J. E.; Oldham, N. J., Carbene footprinting accurately maps binding sites in protein–ligand and protein–protein interactions. *Nat. Commun.* **2016,** *7*, 13288.

12.   Liu, T.; Marcinko, T. M.; Kiefer, P. A.; Vachet, R. W., Using Covalent Labeling and Mass Spectrometry To Study Protein Binding Sites of Amyloid Inhibiting Molecules. *Analytical Chemistry* **2017,** *89* (21), 11583-11591.

13.   Guo, C.; Cheng, M.; Gross, M. L., Protein-Metal-Ion Interactions Studied by Mass Spectrometry-Based Footprinting with Isotope-Encoded Benzhydrazide. *Anal. Chem.* **2019,** *91* (2), 1416-1423.

14.   Watson, C.; Sharp, J. S., Conformational analysis of therapeutic proteins by hydroxyl radical protein footprinting. *AAPS J.* **2012,** *14* (2), 206-217.

15.   Deperalta, G.; Alvarez, M.; Bechtel, C.; Dong, K.; McDonald, R.; Ling, V., Structural analysis of a therapeutic monoclonal antibody dimer by hydroxyl radical footprinting. *mAbs* **2013,** *5* (1), 86-101.

16.   Borotto, N. B.; Zhou, Y.; Hollingsworth, S. R.; Hale, J. E.; Graban, E. M.; Vaughan, R. C.; Vachet, R. W., Investigating Therapeutic Protein Structure with Diethylpyrocarbonate Labeling and Mass Spectrometry. *Analytical Chemistry* **2015,** *87* (20), 10627-10634.

17.   Zhang, Y.; Wecksler, A. T.; Molina, P.; Deperalta, G.; Gross, M. L., Mapping the Binding Interface of VEGF and a Monoclonal Antibody Fab-1 Fragment with Fast Photochemical Oxidation of Proteins (FPOP) and Mass Spectrometry. *Journal of The American Society for Mass Spectrometry* **2017,** *28* (5), 850-858.

18.   Pan, L. Y.; Salas-Solano, O.; Valliere-Douglass, J. F., Localized conformational interrogation of antibody and antibody-drug conjugates by site-specific carboxyl group footprinting. *mAbs* **2017,** *9* (2), 307-318.

19.   Limpikirati, P.; Hale, J. E.; Hazelbaker, M.; Huang, Y.; Jia, Z.; Yazdani, M.; Graban, E. M.; Vaughan, R. C.; Vachet, R. W., Covalent labeling and mass spectrometry reveal subtle higher order structural changes for antibody therapeutics. *mAbs* **2019,** *11* (3), 463-476.

20.   Limpikirati, P.; Liu, T.; Vachet, R. W., Covalent labeling-mass spectrometry with non-specific reagents for studying protein structure and interactions. *Methods* **2018,** *144*, 79-93.

21.   Mendoza, V. L.; Vachet, R. W., Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spectrom. Rev.* **2008,** *28* (5), 785-815.

22.   Xu, G.; Chance, M. R., Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics. *Chem. Rev.* **2007,** *107* (8), 3514-3543.

23.   Wang, L.; Chance, M. R., Structural Mass Spectrometry of Proteins Using Hydroxyl Radical Based Protein Footprinting. *Anal. Chem.* **2011,** *83* (19), 7234-7241.

24.   Kiselar, J.; Chance, M. R., High-Resolution Hydroxyl Radical Protein Footprinting: Biophysics Tool for Drug Discovery. *Annu. Rev. Biophys.* **2018,** *47* (1), 315-333.

25.   Li, K. S.; Shi, L.; Gross, M. L., Mass Spectrometry-Based Fast Photochemical Oxidation of Proteins (FPOP) for Higher Order Structure Characterization. *Acc. Chem. Res.* **2018,** *51* (3), 736-744.

26.   Zhang, B.; Cheng, M.; Rempel, D.; Gross, M. L., Implementing fast photochemical oxidation of proteins (FPOP) as a footprinting approach to solve diverse problems in structural biology. *Methods* **2018,** *144*, 94-103.

27.   Mendoza, V. L.; Vachet, R. W., Protein Surface Mapping Using Diethylpyrocarbonate with Mass Spectrometric Detection. *Analytical Chemistry* **2008,** *80* (8), 2895-2904.

28.   Zhou, Y.; Vachet, R. W., Diethylpyrocarbonate Labeling for the Structural Analysis of Proteins: Label Scrambling in Solution and How to Avoid It. *Journal of The American Society for Mass Spectrometry* **2012,** *23* (5), 899-907.

29.   Zhou, Y.; Vachet, R. W., Increased Protein Structural Resolution from Diethylpyrocarbonate-based Covalent Labeling and Mass Spectrometric Detection. *Journal of The American Society for Mass Spectrometry* **2012,** *23* (4), 708-717.

30.   Borotto, N. B.; Degraan-Weber, N.; Zhou, Y.; Vachet, R. W., Label Scrambling During CID of Covalently Labeled Peptide Ions. *J. Am. Soc. Mass Spectrom.* **2014,** *25* (10), 1739-1746.

31.   Trinquier, G.; Sanejouand, Y. H., Which effective property of amino acids is best preserved by the genetic code? *Protein Engineering, Design and Selection* **1998,** *11* (3), 153-169.

32.   Harris, T. K.; Turner, G. J., Structural Basis of Perturbed pKa Values of Catalytic Groups in Enzyme Active Sites. *IUBMB Life* **2002,** *53* (2), 85-98.

33.   Mädler, S.; Bich, C.; Touboul, D.; Zenobi, R., Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. *J. Mass Spectrom.* **2009,** *44* (5), 694-706.

34.   Mehler, E. L.; Fuxreiter, M.; Simon, I.; Garcia-Moreno E, B., The role of hydrophobic microenvironments in modulating pKa shifts in proteins. *Proteins: Struct., Funct., Bioinf.* **2002,** *48* (2), 283-292.

35.   Isom, D. G.; Castañeda, C. A.; Cannon, B. R.; García-Moreno E, B., Large shifts in pKa values of lysine residues buried inside a protein. *Proceedings of the National Academy of Sciences* **2011,** *108* (13), 5260.

36.   Peck, M. T.; Ortega, G.; De Luca-Johnson, J. N.; Schlessman, J. L.; Robinson, A. C.; García-Moreno E, B., Local Backbone Flexibility as a Determinant of the Apparent pKa Values of Buried Ionizable Groups in Proteins. *Biochemistry* **2017,** *56* (40), 5338-5346.

37.   Robinson, A. C.; Majumdar, A.; Schlessman, J. L.; García-Moreno E, B., Charges in Hydrophobic Environments: A Strategy for Identifying Alternative States in Proteins. *Biochemistry* **2017,** *56* (1), 212-218.

38.   Zhou, Y.; Wu, Y.; Yao, M.; Liu, Z.; Chen, J.; Chen, J.; Tian, L.; Han, G.; Shen, J.-R.; Wang, F., Probing the Lysine Proximal Microenvironments within Membrane Protein Complexes by Active Dimethyl Labeling and Mass Spectrometry. *Anal. Chem.* **2016,** *88* (24), 12060-12065.

39.   Ruan, Q.; Zhao, C.; Ramsay, C. S.; Tetin, S. Y., Characterization of Fluorescently Labeled Protein with Electrospray Ionization-MS and Fluorescence Spectroscopy: How Random is Random Labeling? *Anal. Chem.* **2018,** *90* (16), 9695-9699.

40.   Jumper, C. C.; Schriemer, D. C., Mass Spectrometry of Laser-Initiated Carbene Reactions for Protein Topographic Analysis. *Analytical Chemistry* **2011,** *83* (8), 2913-2920.

41.   Jumper, C. C.; Bomgarden, R.; Rogers, J.; Etienne, C.; Schriemer, D. C., High-Resolution Mapping of Carbene-Based Protein Footprints. *Analytical Chemistry* **2012,** *84* (10), 4411-4418.

42.   Fraczkiewicz, R.; Braun, W., Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J. Comput. Chem.* **1998,** *19* (3), 319-333.

43.   Verdone, G.; Corazza, A.; Viglino, P.; Pettirossi, F.; Giorgetti, S.; Mangione, P.; Andreola, A.; Stoppini, M.; Bellotti, V.; Esposito, G., The solution structure of human β2-microglobulin reveals the prodromes of its amyloid transition. *Protein Sci.* **2002,** *11* (3), 487-499.

44.   Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J., Structure of ubiquitin refined at 1.8 Å resolution. *J. Mol. Biol.* **1987,** *194* (3), 531-544.

45.   Chantalat, L.; Jones, N. D.; Korber, F.; Navaza, J.; Pavlovsky, A. G., The Crystal Structure of Wild-type Growth Hormone at 2.5 Å Resolution. *Protein Pept. Lett.* **1995,** *2* (2), 333-340.

46.   Schrodinger, LLC *The PyMOL Molecular Graphics System, Version 1.8*, 2015.

47.   Negi, S.; Zhu, H.; Fraczkiewicz, R.; Braun, W. Solvent Accessible Surface Areas, Atomic Solvation Energies, and Their Gradients for Macromolecules. 2015 April 17 [accessed 2019 Feb 9]. http://curie.utmb.edu/area_man.html.

48.   Miller, B. T.; Collins, T. J.; Nagle, G. T.; Kurosky, A., The occurrence of O-acylation during biotinylation of gonadotropin-releasing hormone and analogs. Evidence for a reactive serine. *J. Biol. Chem.* **1992,** *267* (8), 5060-9.

49.   Miller, B. T.; Kurosky, A., Elevated Intrinsic Reactivity of Seryl Hydroxyl Groups within the Linear Peptide Triads His-Xaa-Ser or Ser-Xaa-His. *Biochem. Biophys. Res. Commun.* **1993,** *196* (1), 461-467.

50.   Miller, B. T.; Collins, T. J.; Rogers, M. E.; Kurosky, A., Peptide Biotinylation with Amine-Reactive Esters: Differential Side Chain Reactivity. *Peptides* **1997,** *18* (10), 1585-1595.

51.   Guo, X.; Bandyopadhyay, P.; Schilling, B.; Young, M. M.; Fujii, N.; Aynechi, T.; Guy, R. K.; Kuntz, I. D.; Gibson, B. W., Partial Acetylation of Lysine Residues Improves Intraprotein Cross-Linking. *Anal. Chem.* **2008,** *80* (4), 951-960.

52.   Thurlkill, R. L.; Grimsley, G. R.; Scholtz, J. M.; Pace, C. N., pK values of the ionizable groups of proteins. *Protein Sci.* **2006,** *15* (5), 1214-1218.

53.   Zhang, Z.; Browne, S. J.; Vachet, R. W., Exploring Salt Bridge Structures of Gas-Phase Protein Ions using Multiple Stages of Electron Transfer and Collision Induced Dissociation. *J. Am. Soc. Mass Spectrom.* **2014,** *25* (4), 604-613.

54.   Petsko, G. A.; Ringe, D., *Protein Structure and Function*. New Science Press Ltd: London, England, United Kingdom, 2004; pp 10-11.

55.   Whitford, D., *Proteins: Structure and Function*. John Wiley & Sons Ltd: West Sussex, England, United Kingdom, 2005; pp 53-58.

56.   Finkelstein, A. V.; Ptitsyn, O. B., *Protein Physics: A Course of Lectures*. Academic Press: London, England, United Kingdom, 2002; pp 33-42.

57.   Miles, E. W., Modification of histidyl residues in proteins by diethylpyrocarbonate. In *Methods in Enzymology*, Academic Press: Cambridge, Massachusetts, United States, 1977; Vol. 47, pp 431-442.

COVALENT LABELING AND MASS SPECTROMETRY REVEAL SUBTLE

HIGHER ORDER STRUCTURAL CHANGES FOR ANTIBODY

THERAPEUTICS

This chapter is part of a research article published as: Limpikirati, P.; Hale, J. E.; Hazelbaker, M.; Huang, Y.; Jia, Z.; Yazdani, M.; Graban, E. M.; Vaughan, R. C.; Vachet, R. W., Covalent labeling and mass spectrometry reveal subtle higher order structural changes for antibody therapeutics. *mAbs* **2019,** *11* (3), 463-476.

3.1 Introduction

Monoclonal antibodies (mAbs) are among the fastest growing categories of therapeutics in the pharmaceutical industry.[1-3] By 2020, around 70 mAb products are anticipated to be available on the market and their global sales are predicted to be nearly $125 billion, representing 15% of total pharmaceutical sales.[1, 4] Unlike small molecule drugs, the higher-order structure (HOS) of mAbs contributes to the greater binding specificity towards drug targets, resulting in higher therapeutic efficacy and less adverse effects. Changes in HOS upon storage or mishandling, e.g., protein misfolding and aggregation, however, can lead to reduced stability, loss of efficacy, unwanted actions, or possible immunogenicity.[5, 6] Monitoring HOS is thus essential to ensure efficacy and safety of mAb therapeutics throughout a product life cycle – from drug manufacturing to dose administration.[6-9] Any new, structurally-informative method could be useful for biologics license applications (BLAs) because HOS characterization is required in the stability, lot-to-lot comparability, and biosimilar studies of antibody therapeutics.[6, 9, 10]

Detecting HOS changes of mAbs is challenging given their size and the multidomain nature. The current toolbox for HOS analysis of protein therapeutics has

limitations.[5, 6, 8]   X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy can provide atomic-level resolution of protein structure, but these methods are time- and sample-consuming, and not amenable to all proteins. In contrast, biophysical techniques such as differential scanning calorimetry (DSC), dynamic light scattering (DLS), fluorescence spectroscopy, infrared (IR) spectroscopy, and circular dichroism (CD) spectroscopy are rapid, but provide only low-resolution ensemble averages of protein global conformation. These low-resolution methods still are commonly used to characterize HOS of mAb therapeutics,[11-19] even though they do not provide information about small localized conformational changes, some of which may be potentially significant for drug efficacy and safety. Hence, there is a growing need for rapid and sample-efficient analytical tools with moderate resolution that can characterize HOS of therapeutic mAbs at the amino acid level. Current regulatory guidelines do not specify what method(s) should be used in characterizing HOS of biologics,[20-26] allowing for the development of novel analytical techniques.

Mass spectrometry (MS) has become one of the most powerful methods for the analysis of proteins. From a recent evaluation of BLAs, the use of MS to characterize primary structure and HOS of protein therapeutics has increased in recent years.[27] Characterizing HOS using MS requires that a protein's structural information is encoded into the mass of that protein. Commonly-used MS-based approaches for studying protein structure include hydrogen/deuterium exchange (HDX) and covalent labeling.[6, 8, 10] In HDX-MS, information regarding solvent accessibility and dynamics of backbone amides can be obtained from the exchange of hydrogens by deuteriums, thereby increasing the mass in a structurally informative manner. HDX-MS has been successfully used to

investigate structural changes and identify aggregation sites in mAbs obtained from different storage or stress conditions.[28-30] Although this technique has been commonly used in HOS analysis of mAb therapeutics,[31-33] an analytical challenge is the accuracy of HDX measurements due to the transient nature of deuterium labeling that can lead to back exchange and scrambling. In addition, specialized robotic equipment and software are required to obtain optimal results.

Covalent labeling (CL) can also be used with MS to study protein HOS, but unlike HDX, CL is generally not subject to back exchange and scrambling. CL approaches use reagents to irreversibly modify solvent-exposed amino acid side chains, encoding structural information into the mass of the protein. CL along with MS detection (CL-MS), especially when used with bottom-up tandem MS (MS/MS), provides information about solvent accessibility of amino acid side chains, making it complementary to HDX.[34-38] As an example, site-specific carboxyl group footprinting has been applied for the structural characterization of glycosylated therapeutic mAbs[39, 40] and for epitope-paratope mapping.[41, 42] A localized conformation change of mAbs as induced by deglycosylation was also identified using this technique.[43] However, the reagent pair of 1-ethyl-3-(3-(dimethylamino)propyl)-carbodiimide hydrochloride (EDC) and glycine ethyl ester (GEE) that was used in the study can monitor only Glu and Asp residues, limiting its structural resolution. Using a more non-specific reagent capable of modifying a range of amino acid side chains simultaneously can allow protein structure to be probed with greater resolution.[38] Hydroxyl radical footprinting (HRF), in which hydroxyl radicals ($^\bullet$OH) are generated and non-selectively label solvent accessible residues,[44] has been used for the structural analysis of mAbs.[40] HRF with MS detection has been used to identify epitopes

89

and paratopes for antigen-antibody interactions[41, 45-47] and to determine dimer interfaces in formulated mAbs.[48] However, specialized instrumentation such as a laser or a synchrotron source is needed for $^\bullet$OH radical production. Moreover, data interpretation is complicated because $^\bullet$OH radicals can generate over 50 different modification types on proteins.[44] For day-to-day structural analysis of biotherapeutics, a less expensive and simpler experimental design and workflow is needed.

Diethylpyrocarbonate (DEPC) is a commercially-available reagent that can react with a range of nucleophilic side chains and N-termini of proteins, resulting in excellent protein structural coverage.[38, 49] Unlike radical reagents, no specialized instrumentation is required because DEPC can readily modify proteins once added to solution. DEPC labeling reactions also result in only one type of product, allowing very low levels of labeling to be straightforwardly identified and further increasing analytical sensitivity and structural coverage.[34, 38] CL-MS methods based on DEPC have been extensively developed by our group,[35, 50, 51] and have been used to obtain insight into protein-ligand and protein-protein interactions as changes in side chain solvent accessibility at binding sites can be probed.[35, 52-56]

Here, we extend the applicability of DEPC-based CL-MS to study protein conformational changes that occur upon heating. Using rituximab as a model mAb therapeutic, we demonstrate the ability to site-specifically detect subtle HOS changes at the temperatures far below the melting point of the mAb therapeutic, changes that are not detected by common biophysical methods. Sites of these structural changes are revealed at the amino-acid level using MS/MS and are validated by activity assays. Overall, given the

simplicity, sensitivity, and straightforwardness of DEPC CL-MS, we predict that this method will be amenable to the structural investigations of other antibody therapeutics.

3.2 Experimental Section

3.2.1 Materials

Rituximab formulation (Rituxan® 100 mg/10 mL vial, lot# 3209283, Genentech) was ordered from Myoderm. Diethylpyrocarbonate (DEPC) (#D5758), imidazole (#I5513), iodoacetamide (#I6125), tris(2-carboxyethyl)phosphine (TCEP) (#C4706), and trypsin (#T1426) were obtained from Sigma-Aldrich. Urea (#AC424581000) was purchased from Acros Organics. Sodium phosphate monobasic monohydrate (#S0710-1) was obtained from EM Science. Sodium phosphate dibasic anhydrous (#S374-500), LC/MS-grade formic acid (#A117-50), acetonitrile (#A998-4), and water (#W7-4) were purchased from Fisher Scientific.

3.2.2 Heat Treatments

Samples were aliquoted (5 μL) from a rituximab formulation stored at 4 °C. Control samples were incubated at 37 °C for 5 min prior to DEPC labeling. For thermal treatments, the samples were incubated at 45 °C, 55 °C, and 65 °C for 4 h in a temperature-controlled water bath. The samples were cooled to 37 °C prior to labeling.

3.2.3 DEPC Labeling Reactions

Rituximab samples (10 mg/mL, 69.5 μM) in formulation with only minor dilution (to 58 μM) were reacted with DEPC. The reagent was first diluted in acetonitrile to make an intermediate solution, and the final solution of DEPC was then prepared in water.

Labeling of rituximab was performed at 37 °C for 5 min at a DEPC to protein molar ratio of 4 to 1. The reaction was quenched by the addition of imidazole at a 1:50 DEPC to imidazole molar ratio. For each thermal treatment, at least three replicates were performed on the rituximab samples.

3.2.4 Proteolytic Digestion

Following quenching the labeled samples were added into a urea-containing tube and diluted in 50 mM phosphate buffer at pH 7.4. The resulting mixture contained 8 M urea for protein denaturation. L-methionine was added at the final concentration of 2 mg/mL to minimize oxidation during protein digestion. To achieve complete digestion of rituximab, TCEP was added at the final concentration of 25 mM to reduce the disulfide bonds. Iodoacetamide (25 mM) was simultaneously added to alkylate the reduced Cys residues. The samples were kept in the dark at room temperature for 20 min. Tween® 80 was then removed from the samples using DetergentOUT™ Tween® Micro spin columns (#786-214, G-Biosciences). Subsequently, trypsin was added to the resulting samples at a 1:10 (w/w) enzyme to substrate, and the labeled protein was digested overnight at 37 °C. Following digestion, an Amicon® centrifugal filter with a 10 kDa molecular weight cutoff (#UFC501096, Millipore) was used to remove trypsin from the resulting rituximab peptides. The flow-through was collected, flash-frozen in liquid nitrogen, and stored at -20 °C until LC-MS/MS analysis.

3.2.5 HPLC Separation

Online LC-MS/MS analyses were performed on all rituximab digests. A sample containing approximately 2 μg rituximab peptides was loaded on an Easy-NanoLC 1000

system (Thermo Scientific). A flow rate of 300 nL/min was used. Samples were first trapped and desalted on an Acclaim™ PepMap™ C18 trap column (2 cm x 75 µm ID, 3 µm; Thermo Scientific). Separation of rituximab peptides was then performed using a FortisBIO C18 nanocolumn (15 cm x 75 µm ID, 1.7 µm; Fortis Technologies). LC/MS-grade water (solvent A) and acetonitrile (solvent B), each containing 0.1% formic acid, were used as mobile phases. A shallow gradient was utilized to achieve sufficient separation of peptides on analytical column. A linear gradient of solvent B was increased from 0% B to 50% B over 90 min. The LC column was then flushed by elevating the mobile phase composition to 95% B over 15 min and holding at 95% B for another 20 min.

3.2.6 Mass Spectrometry

Mass spectra were acquired on a Thermo Scientific Orbitrap Fusion mass spectrometer. The nano-electrospray ionization source was operated in the positive mode using a needle voltage of 2,000 V. The ion transfer tube temperature was set to 300 °C. The resolution of Orbitrap was set to 60,000 and the MS1 AGC target and maximum injection time were optimized and set to 1×10^6 ions and 100 msec, respectively. Tandem mass spectrometry (MS/MS) was performed on linear quadrupole ion trap for the most abundant peptide ions, with ion abundances above 5,000. The precursor ions were selected using a quadrupole mass filter at an isolation width of 2.0, and the MS2 AGC target and maximum injection time were set to 5×10^4 ions and 100 msec, respectively. Tandem mass spectra were generated using collisional-induced dissociation with a normalized collision energy of 35%. To avoid a biased selection of high-abundance ions, a dynamic exclusion of 60 sec was activated after 5 spectra were acquired for any given precursor ion within 5

sec. Mass detection during MS and MS/MS was done in centroid mode to ease the data analysis.

3.2.7 Peptide Identification and Peak Quantification

A custom software pipeline described previously[56] and specifically designed for protein CL-MS studies was used to identify and quantify labeling sites. Raw mass spectral data files from the LC-MS/MS analyses were firstly converted to .mgf format using MSConvertGUI software.[57, 58] SearchGUI was used to analyze the .mgf files for the peptide identification.[59] The sequence of rituximab's light and heavy chains were added to a sequence database constructed from the common Repository of Adventitious Proteins (cRAP database).[60] Tandem mass spectra were searched against the custom database and its reverse, the decoy database. Several search engines (X!Tandem,[61, 62] MS-GF+,[63] OMSSA,[64] and MyriMatch[65]) were all used. The search parameters were set as follows: a precursor mass tolerance of 10 ppm, carbamidomethylation of Cys as a fixed modification, oxidation of Met as a variable modification, and DEPC modification of His, Lys, Ser, Thr, Tyr, and N-terminus as a user variable modification (mass addition of 72.0211). Nonspecific enzyme cleavage was selected to improve the sequence coverage of searches, accounting for nonspecific proteolytic digestion and degradation of rituximab peptides during sample preparation. Next, PeptideShaker was used to visualize the results from multiple search engines.[66] False discovery rates were set at 1% and post-translational modifications were scored using the PhosphoRS algorithm. The peptide identification results were reformatted and exported to a .csv file and processed to contain only rituximab features, which were to be used as a custom database for LC-MS peak identification. MZmine was used to analyze the raw LC-MS data files for peptide peak quantification.[67,

[68] Mass detection was performed at the MS1 level. Chromatograms were reconstructed and deconvoluted using m/z and ion abundances data, and peak areas were quantified. Deconvoluted chromatograms from multiple samples were aligned using the Join Aligner algorithm. The quantified, identified, and aligned data were finally exported in .csv format. DEPC modification levels (L) of each labeled residue were calculated as follows. (**Equation 3.1**)

$$L = \% \; DEPC \; labeling = \frac{\sum_{i=1}^{n} \sum_{z=1}^{m} A_{i,z}^{modi}}{\sum_{i=1}^{n} \sum_{z=1}^{m} A_{i,z}^{modi} + \sum_{i=1}^{n} \sum_{z=1}^{m} A_{i,z}^{unmodi}} \times 100 \qquad (3.1)$$

where $A_{i,z}^{unmodi}$ is the peak area of unmodified peptide that has a sequence containing the residue of interest (i) and possesses a certain charge state (z) in the mass spectrum, and $A_{i,z}^{modi}$ is the peak area of peptide in which the residue of interest is modified. A pictorial representation of this calculation can be found in **Figure 3.1**.

Note that the DEPC modification extents calculated by **Equation 3.1** are only used for the relative quantitation, i.e., comparing the protein of interest under different conditions (control vs. stressed). The label levels do not reflect the absolute quantitation of modified species as the addition of a carbethoxyl group to the modified peptide and different LC solvent conditions during elution of peptides lead to different ionization efficiency of the unmodified and modified peptides.

(a) Total ion chromatogram (TIC)

$$\% \ DEPC \ Labeling = \frac{\sum Peak \ area \ of \ Labeled}{\sum Labeled + \sum Unlabeled} \times 100$$

(b) Extracted ion chromatogram (XIC)
m/z 437.5666, z=(+3)

(c) XIC m/z 461.5736, z=(+3)

(d) CID MS/MS
HC (311-321) TVLHQDWLNGK

(e) CID MS/MS
HC (311-321) TVLH*$_{(DEPC)}$QDWLNGK
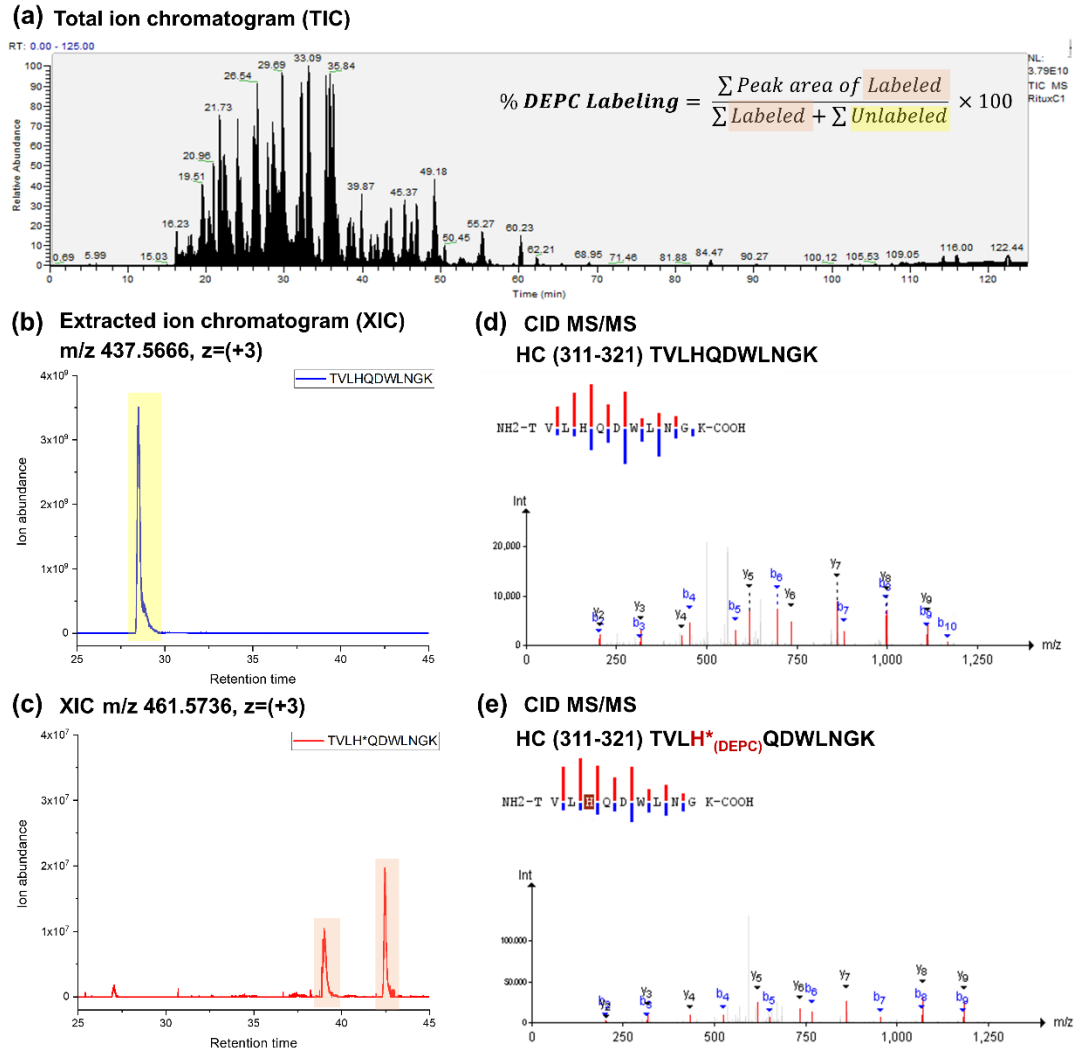
Figure 3.1: Illustration of how the DEPC modification levels are calculated.
After DEPC labeling and proteolytic digestion, (a) LC-MS analysis of the digested rituximab is performed. During LC-MS, peptides are subjected to CID MS/MS in a linear quadrupole ion trap for identication, while peptide ion abundances measured by the Orbitrap are used for peak area quantification. Peak areas of (b) unlabeled and (c) labeled peptides in a chromatogram are used to calculate the labeling percent (see **Equation 3.1**). CID tandem spectra of (d) unlabeled and (e) labeled peptides obtained at specific retention time are used for peptide sequencing and identification of DEPC labeled site. Two LC peaks for the labeled peptide in (c) are observed because the His side chain has two nitrogens that are separately labeled to produce isomers that can be separated by LC.

96

3.2.8 Biophysical Characterization

3.2.8.1 Intrinsic Fluorescence Spectroscopy

Tryptophan fluorescence measurements were performed on a Photon Technology International Quantamaster-4SE spectrofluorometer. Rituximab samples were diluted to 1 μM in 50 mM phosphate buffer at pH 7.4 prior to analysis. Heated samples were then cooled to 37 °C prior to a measurement. A 200-μL solution of 1 μM rituximab sample was transferred to a quartz cuvette. Fluorescence spectra were acquired at room temperature using an excitation wavelength of 285 nm with a slid width of 1 nm and an emission scan range of 310 – 440 nm with a slit width of 0.2 nm.

3.2.8.2 Circular Dichroism (CD) Spectroscopy

Far-UV CD analyses were performed on a Jasco J-1500 spectropolarimeter. CD spectra were recorded at room temperature over a scan range of 250 to 195 nm. Rituximab samples were diluted to 1 μM in 50 mM phosphate buffer at pH 7.4 prior to analysis. Heated samples were cooled to 37 °C prior to a measurement. A 200-μL solution of 1 μM rituximab sample was then transferred to a quartz cuvette. The CD spectrometric parameters were set as follows: a scan resolution (data pitch) of 0.5 nm, a scan rate of 20 nm/min, a band width of 2 nm, and a digital integration time of 1 sec. Triplicate measurements were performed for each sample at room temperature. After background subtraction, raw CD outputs (θ, degree) were converted into mean residue ellipticity using **Equation 3.2**.[69]

$$[\Theta]^{MR} \ (degree \ cm^2 dmol^{-1}) \ = \frac{100 \times \theta \ (millidegree)}{C \ (mol \ L^{-1}) \times N \times l \ (cm)} \qquad (3.2)$$

where $[\theta]^{MR}$ is mean residue ellipticity in deg cm^2 dmol^{-1}, θ is raw signal output in mdeg, C is rituximab concentration in molar, N is the number of amino acid residues in a protein, and l is path length of a cuvette in cm.

3.2.8.3 Dynamic Light Scattering (DLS)

Hydrodynamic radii of native and thermally-stressed rituximab were measured at room temperature using a Malvern Zetasizer ZSP – DLS system. Rituximab samples were diluted to 1 μM in 50 mM phosphate buffer at pH 7.4 prior to analysis. Heated samples were cooled to 37 °C prior to a measurement. A 1-mL solution of 1 μM rituximab sample was transferred to a plastic cuvette. Back scattering was detected at a measurement angle of 173°, and volume particle size distribution of the sample was recorded. Five replicate measurements were performed for each sample at room temperature. Measurement duration was set according to the preset levels (automatic mode).

3.2.8.4 Size-Exclusion Chromatography (SEC)

SEC separation of rituximab after heating was performed at room temperature on an Agilent 1260 Infinity HPLC system using a TSKgel SuperSW3000 column (30 cm x 7.8 mm ID, 5 μm particle size; Tosoh Bioscience LLC). Rituximab samples were diluted to 1 mg/mL in 50 mM phosphate buffer at pH 8.0 prior to analysis. Heated samples were cooled to 37 °C prior to a chromatographic run. A sample containing approximately 50 μg rituximab was loaded on an SEC column. The mobile phase (pH 6.5) comprises 50 mM sodium phosphate, 400 mM sodium perchlorate, and 10% isopropanol, and an isocratic flow rate of 0.5 mL/min was used. A variable wavelength UV detector set at 280 nm was used for detection.

3.2.9 Determination of the Labeling Limits of Quantitation (LOQ) through Peptide Spiking Experiments

Model peptide H$_2$N-VVSVLTVLHQDWLNGK, which comes from the rituximab sequence (HC peptide with amino acids 306-321), was custom synthesized (AnaSpec Inc.). A variety of DEPC-modified species (with different modification sites) were generated from a covalent labeling of the model peptide (10 μM) at a DEPC to protein molar ratio of 10 to 1 at 37°C for 5 min, which resulted in labels at N-terminus, H314, and K321. Labeled sites were identified and modification percentages were determined by LC-MS/MS. A mixture of DEPC-labeled peptides were spiked into a digest of unlabeled rituximab (2.57 μM matrix) at varying concentrations, and the resulting solutions were subsequently analyzed by LC-MS/MS. Signals of the spiked peptide, as measured by LC-MS, were used to estimate the LOQ based on signal-to-noise ratio of response.

3.2.10 Activity Assays

3.2.10.1 Alamar Blue Assay

A range of stress temperatures were used in this study. Rituximab samples were incubated at 37°C (control), 50°C, 60°C, and 68°C prior to an assay. The conditions used for this assay were derived from Zhang *et al*[70]. Briefly, Raji cells (ATCC CLL-86) were grown in a T75 flask using RPMI media supplemented with fetal bovine serum (FBS). Once confluent, cells were washed with phosphate-buffered saline (PBS) and diluted to 1 x 10^6 cells/mL in RPMI media. 90 μL of cells were delivered into a 96 well plate, and 10 μL of control or heat-treated rituximab diluted in RPMI media was added at the concentrations indicated (4 ng/mL rituximab). Plates were then incubated for 30 min at 37 °C, then each well was supplemented with 10% Invitrogen^{TM} normal human serum

(#31876, Thermo Fisher Scientific). Plates were returned to 37 °C for 4 h, then 11 µL of Invitrogen™ Alamar Blue reagent (#DAL1025, Thermo Fisher Scientific) was added per manufacturer's specifications. After 1 hour, fluorescence signal was read using an excitation wavelength of 560 nm and emission of 590 nm. Data was generated on a Synergy H1 microplate reader (BioTek) and results were exported to Microsoft Excel for analysis. Each sample was generated in triplicate, and independently generated to confirm trends were the same.

3.2.10.2 Rituximab Bridging ELISA

Custom rituximab ELISA plates were generated and used as previously described in Cragg *et al*[71] and Hampson *et al*[72] with some modifications. Briefly, anti-human capture antibody (clone SB2H2 recognizing the F_c region of human antibody, #MCA2531, Biorad) was diluted 1 to 1000 in coating buffer (15 mM sodium carbonate, 28.5 mM sodium bicarbonate, pH 9.6), and 100 µL was added to each well of a 96 well NUNC MaxiSorp flat-bottom plate (#44-2404-21, Thermo Fisher Scientific). Plates were then incubated overnight at 4 °C, then blocked with 200 µL of 1% bovine serum albumin (BSA) in PBS for 2 h. Plates were washed three times with PBST prior to use. For the rituximab quantitation and standard curve generation, samples were diluted in phosphate-buffered saline with Tween® 20 (PBST) at the concentrations indicated (1 to 500 ng/mL rituximab). For the binding assay, 10 ng/mL rituximab was used in the experiment. Rituximab samples were incubated at 37°C (control), 45°C, 55°C, and 65°C prior to an assay. 100 µL of sample was added to each well and incubated for 1 h at room temperature. Wells were then washed five times with PBST, and the horseradish peroxidase (HRP) – labeled anti-rituximab detection antibody (clone MB2A4 anti-idiotypic antibody, #MCA2260P, Bio-Rad) was

added at a 1: 60,000 dilution in blocking buffer. Samples were incubated for 90 minutes, then washed five times with PBST. Plates were then developed using 100 μL of HRP substrate for 45 min and stopped with the addition of 50 μL of 3 M sulfuric acid. Data was generated on a Synergy H1 microplate reader (BioTex) measuring absorbance at 450 nm and normalized against the absorbance at 630 nm to clear up any background signal, and the results were exported to Microsoft Excel for analysis. Each sample was generated independently in triplicate.

3.2.10.3 Raji Cell Pull-Down Assay

A range of stress temperatures were used in this study. Rituximab samples were preheated at 37°C (control), 50°C, 60°C, and 68°C. Control and heat-treated rituximab was diluted to 200 ng/mL in PBS and mixed 1:1 with either Raji cells (1 x 10^6 cells/mL) in PBS, or PBS alone was used as a control. The total concentration of rituximab in the experiment was 100 ng/mL. Free rituximab was then quantified using the ELISA plates, and normalized against the control wells. Absolute quantitation was performed using standard curves generated from each heat-treated sample independently to confirm quantitation was in the linear part of the standard curve.

3.3 Results

3.3.1 Biophysical Characterization of Rituximab HOS after Storage at Mild to Moderate Stress

Before studying heat-stressed rituximab by covalent labeling, we used CD spectroscopy, fluorescence spectroscopy, and dynamic light scattering (DLS) to identify any structural perturbations upon heating.
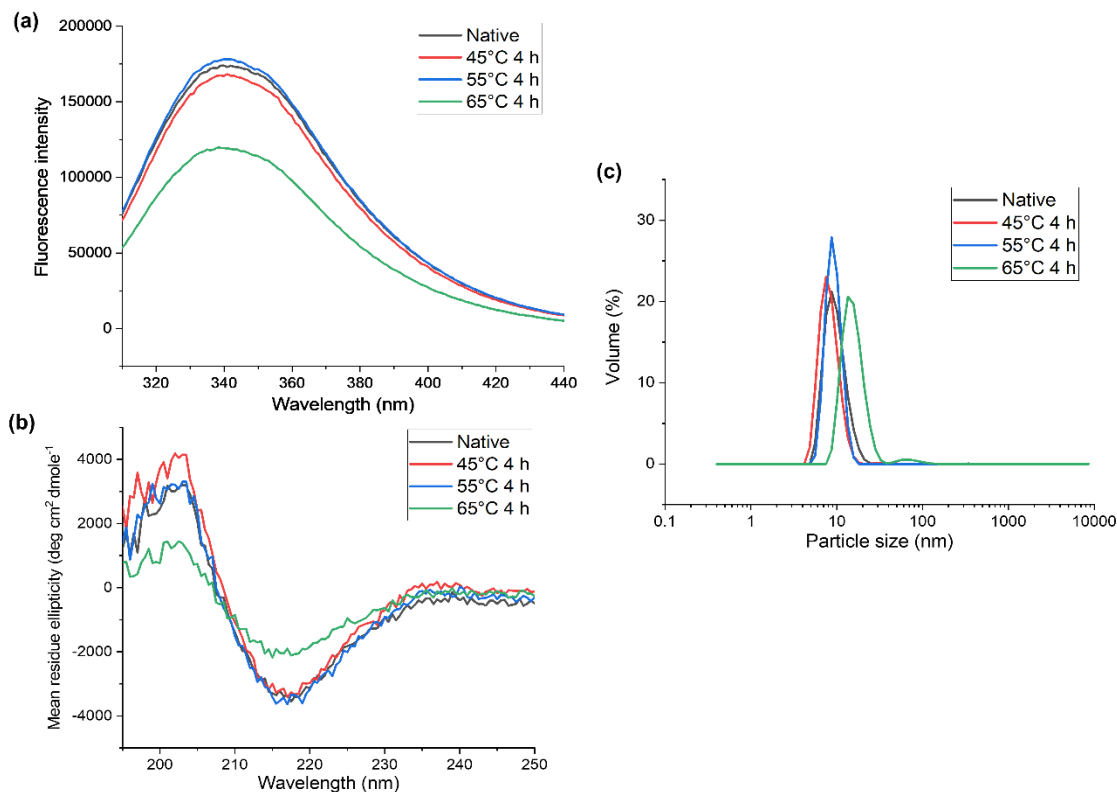
Figure 3.2: Biophysical characterization of rituximab at 37 °C (native, black) and thermally-stressed conditions, after incubation of the rituximab formulation at 45 °C for 4 h (red), 55 °C for 4 h (blue), and 65 °C for 4 h (green).
Techniques used in structural characterization were (a) tryptophan fluorescence spectroscopy, (b) far-UV circular dichroism spectroscopy, and (c) dynamic light scattering. The essentially identical overlap of the spectra of the native and 45 °C and 55°C heat-stressed samples indicate that these techniques do not detect any structural changes to rituximab after preheating to these temperatures. Upon heating at 65 °C, however, structural changes and aggregation occur.

When heating rituximab for 4 h at temperatures below its melting point, we find that the three techniques are not able to detect any significant structural changes at 45 °C or 55 °C and reveal only mild changes at 65 °C (**Figure 3.2**). Upon heating at 65 °C, rituximab undergoes changes in its HOS, as indicated by intrinsic fluorescence and CD spectroscopy (**Figure 3.2a and b**), and unfolds to some extent as indicated by DLS (**Figure 3.2c**). These classical biophysical techniques provide only the weighted average structure of a global conformation, so the locations of any structural changes are unknown.

3.3.2 DEPC Labeling with MS Detection as a Tool for HOS Analysis of Rituximab

DEPC-based covalent labeling together with MS detection was used to identify any structural changes undergone by rituximab that could not be detected by CD, fluorescence, or DLS. DEPC is a reactive electrophile that can modify solvent accessible nucleophilic side chains (Cys, His, Lys, Thr, Tyr, Ser) and N-termini of proteins **(Figure 3.3a and b)**. The resulting carbethoxylated products of these residues have a mass shift of +72.021 Da, and the specific protein modification sites can be identified and semi-quantified after proteolytic digestion, liquid chromatographic separation of the resulting peptides, and tandem MS analysis (**Figure 3.4**). Any changes to the extent of covalent labeling at particular residues can be used to probe HOS changes to proteins upon comparing one condition (e.g. native) to another (e.g., heated). During the labeling conditions, the DEPC to protein molar ratio is limited at 4 to 1 to minimize labeling-induced structural perturbations to a protein during the labeling reaction, while at the same time providing sufficient labeling extents to identify the modified sites.[35] CD and intrinsic fluorescence spectroscopy confirm that the DEPC-modified rituximab undergoes no significant structural changes, as compared to the spectra of unlabeled rituximab **(Figure 3.5)**.
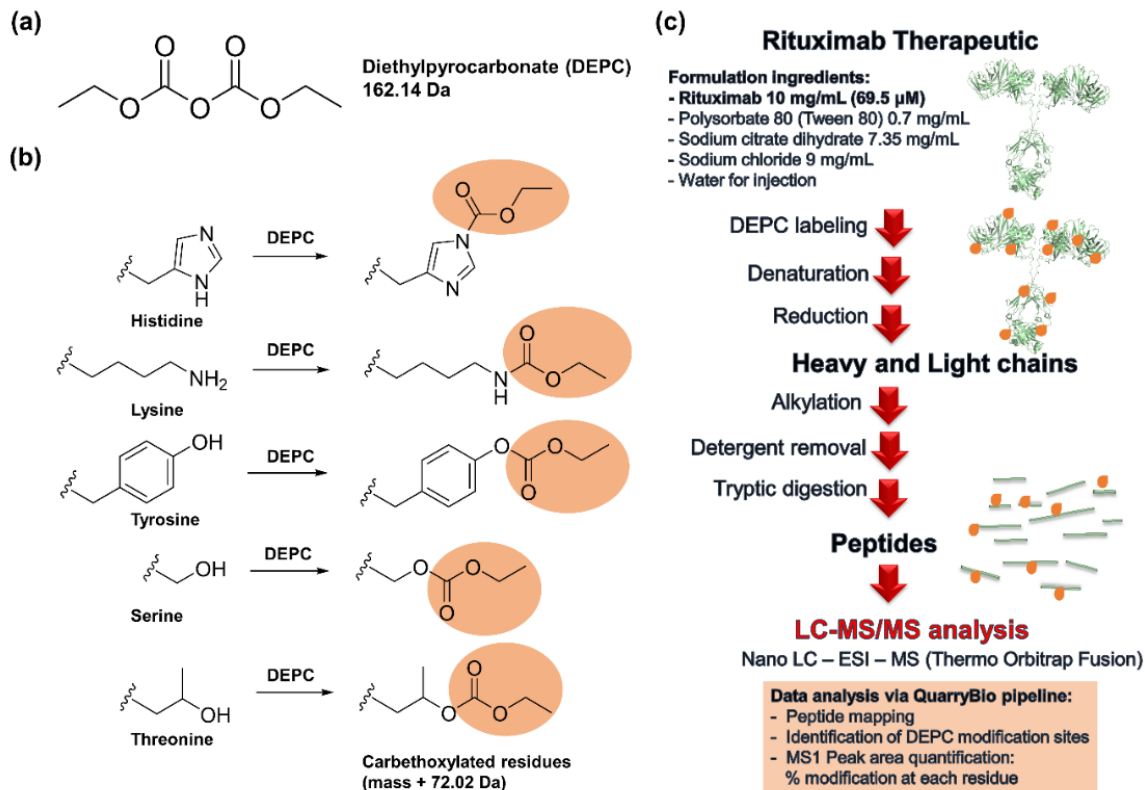
Figure 3.3: DEPC CL-MS as a tool for structural analysis of rituximab.
In an MS-based approach for structural analysis, a protein's structural properties are encoded into the mass of protein via DEPC covalent labeling. This figure shows (a) chemical structure of DEPC, (b) reactions of amino acid residues that are modified in covalent labeling with DEPC, and (c) workflow for DEPC covalent labeling combined with bottom-up MS analysis via proteolytic digestion and LC-MS/MS to identify labeled sites and determine label levels at each modified residue. Peptide identification and peak area quantification were performed using a custom software pipeline developed by QuarryBio Inc.[56]
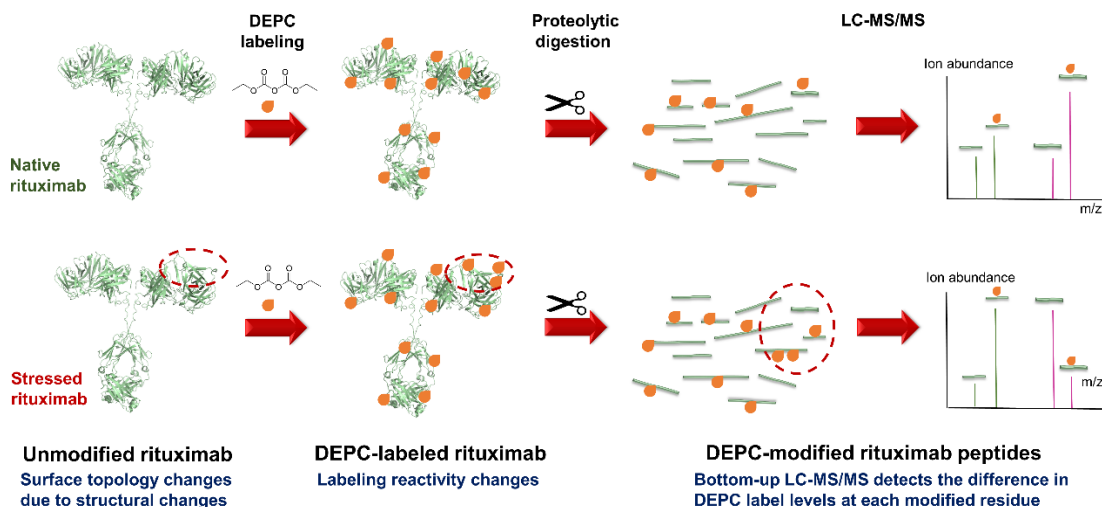
Figure 3.4: Scheme showing DEPC labeling with MS detection for the structural analysis of antibody therapeutics.
DEPC reagent modifies solvent accessible amino acids. The modified protein is subjected to proteolytic digestion and the modified peptides are analyzed using LC–MS/MS. Sites of protein conformational changes can be revealed by changes in the extent of labeling at specific residues
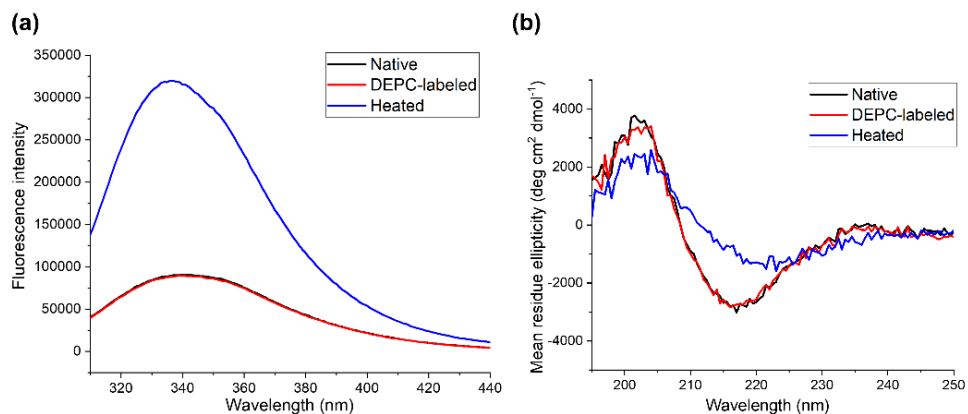
Figure 3.5: Controlling DEPC to rituximab molar ratio to minimize structural perturbations of rituximab during the labeling reaction.
The DEPC to rituximab molar ratio is limited at 4 to 1 in order to minimize structural perturbations of rituximab during the labeling reaction. Previous studies on a variety of proteins indicate that a 4 to 1 ratio is appropriate to minimize labeling-induced structural changes.[35, 55, 56] The structural integrity of rituximab after covalent labeling can be confirmed via: (a) tryptophan fluorescence spectra and (b) circular dichroism spectra obtained from rituximab samples under native (black), DEPC-labeled (red), and heated at 70 °C for 15 min (blue) conditions. Heated rituximab was used to indicate spectral changes due to the structural change. The essentially identical overlap between the spectra of the native and DEPC-labeled rituximab samples indicates that covalent labeling has only little or no effect on the structural perturbation of rituximab.

105

To obtain the extent of labeling at a residue level, the DEPC-labeled rituximab was subjected to "bottom-up" analysis via proteolytic digestion and LC-MS/MS (**Figures 3.1 and 3.3c**). Peptide identification and peak area determination were performed using custom-designed software described previously,[56] as described in **Section 3.2.7**. A rituximab sequence coverage of over 90% was obtained in the tryptic peptide mapping, allowing almost the entire structure of the protein to be probed. Upon DEPC labeling, up to 47 residues are found to be modified in each light chain (LC) while 107 residues are modified in each heavy chain (HC), which together corresponds to 23% of the residues in rituximab. Considering the average distance between adjacent DEPC modification sites,[50] the effective resolution for probing rituximab's structure using DEPC is ~ 8 Å. Some DEPC-modified residues could be reliably detected down to levels as low as 0.001%, however a more conservative labeling threshold of 0.01% was used for all peptides. Detection of peptides at these labeling levels together with the excellent structural coverage allows us to monitor subtle structure changes to the protein upon heat treatment. The limits of quantitation (LOQ) were experimentally determined by spiking experiments of a model synthetic tryptic peptide (rituximab HC peptide with amino acids 306-321; DEPC-labeled at different sites), into a digest of unlabeled rituximab at varying concentrations (**Figures 3.6 and 3.7**).
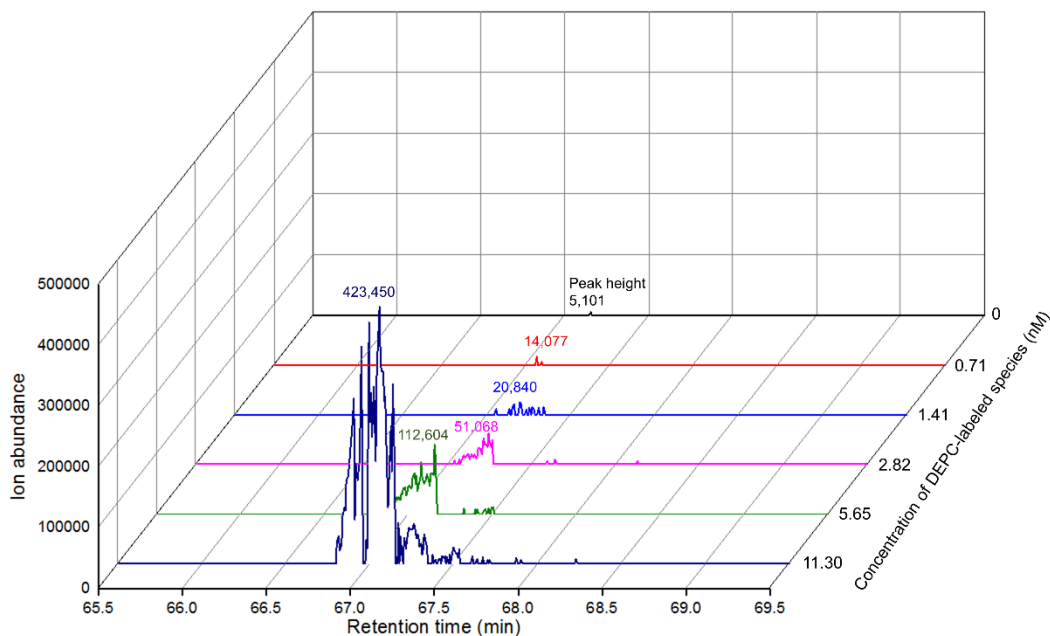
Figure 3.6: Representative extracted ion chromatograms (XICs) used to estimate the labeling LOQ for a DEPC-labeled model peptide (H₂N-VVSVLTVLHQDWLNGK*). This peptide, which comes from the rituximab sequence, was synthesized and DEPC labeled before spiking into a digest of unlabeled rituximab at varying concentrations. Shown above are the XICs for peptide labeled at the Lys residue. Peak heights of the spiked peptide, as measured by LC-MS, are used to estimate the the LOQ based on signal-to-noise ratio of response. When the blank analysis gives a result with a nonzero signal-to-noise ratio, the LOQ is defined as the analyte concentration corresponding to a signal-to-noise ratio that is 10 times the blank signal-to-noise ratio (see **Equation 3.3** below).[73, 74] From the data in this figure, we find that the LOQ is 2.82 nM, which is equivalent to a labeling level of 0.001%. These experiments were repeated three times.

Calculation of LOQ for the K labeled peptide (H₂N-VVSVLTVLHQDWLNGK*)

Blank measurement
Peak height $H_{bl} = 5{,}101$ (noise)

LOQ determination for H₂N-VVSVLTVLHQDWLNGK*
LOQ is defined as the analyte concentration corresponding to the signal-to-noise ratio of 10, hence the peak height $H_{LOQ} = 10 H_{bl}$ (3.3)
$$H_{LOQ} = (10 \times 5{,}101) = 51{,}010$$
The closest data point that yield signal above H_{LOQ} from our surrogate peptide experiment is at $X_{LOQ} = 2.82\ nM$
Using peak areas to calculate a labeling level:
Peak area of this DEPC-modified peptide at 2.82 nM spiked = 20,250.23
Total peak area of unmodified and modified peptides at 2.82 nM spiked = 1,640,538,555.33
$\therefore$ The LOQ is at a labeling level $= \dfrac{20{,}250.23}{1{,}640{,}538{,}555.33} \times 100 = 0.001\%$
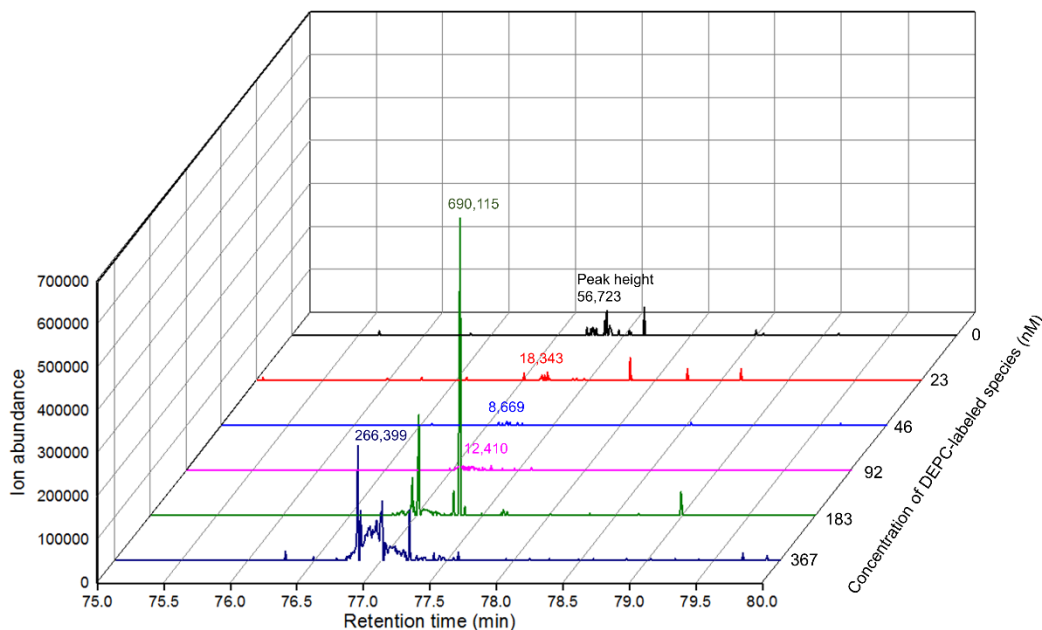
Figure 3.7: Representative extracted ion chromatograms (XICs) used to estimate the labeling LOQ for a DEPC-labeled model peptide (H$_2$N*-VVSVLTVLHQDWLNGK). From the data in this figure, we find that the LOQ is 183 nM, which is equivalent to a labeling level of 0.03%. These experiments were repeated three times.

Calculation of LOQ for the N-terminally labeled peptide (H$_2$N*-VVSVLTVLHQDWLNGK)

Blank measurement
Peak height $H_{bl} = 56,723$ (noise)

LOQ determination for H$_2$N*-VVSVLTVLHQDWLNGK
LOQ is defined as the analyte concentration corresponding to the signal-to-noise ratio of 10, hence the peak height $H_{LOQ} = 10H_{bl}$

$$H_{LOQ} = (10 \times 56,723) = 567,230$$

The closest data point that yield signal above H_{LOQ} from our surrogate peptide experiment is at $X_{LOQ} = 183\ nM$
Using peak areas to calculate a labeling level:
Peak area of this DEPC-modified peptide at 183 nM spiked = 433,880.64
Total peak area of unmodified and modified peptides at 183 nM spiked = 1,644,219,196.06
$\therefore$ The LOQ is at a labeling level $= \dfrac{433,880.64}{1,644,219,196.06} \times 100 = 0.03\%$

To identify regions of rituximab that undergo subtle structural changes upon heating, DEPC modification levels at individual amino acid residues of heat-treated rituximab were compared to the modification levels of the un-treated protein. For each stress condition, unpaired student t-tests were used to determine if the labeling levels are significantly different at a 95% confidence level. For strongly nucleophilic residues, like His and Lys, changes in labeling extent are determined by the absolute difference in DEPC label levels (**Equation 3.4**).

$$Absolute\ difference = \Delta L = L_{stressed} - L_{native} \qquad (3.4)$$

where $L_{stressed}$ and L_{native} are the average % DEPC modification of a given residue in the stressed and native proteins, respectively. (See **Equation 3.1** for a calculation of % DEPC modification.) For weakly nucleophilic residues, such as Tyr, Ser, and Thr, modification extents are much smaller than for His and Lys residues due to their lower intrinsic reactivity with DEPC. Thus, a ratio difference, which is more sensitive to a small modification changes, is used to determine labeling changes for Tyr, Ser, and Thr residues (**Equation 3.5**).

$$Ratio\ difference\ (\%) = \frac{L_{stressed} - L_{native}}{L_{native}} \times 100 \qquad (3.5)$$

3.3.3 DEPC CL-MS for Probing Subtle Structural Changes of Rituximab

In **Table 3.1** the number of DEPC-modifiable residues in each rituximab domain that undergo significant changes in modification extent after heating for 4 h at 45 °C is graphically depicted. The sites that undergo statistically significant labeling changes are binned into three groups – low, medium, and high – based on the extent of the labeling

change. These bins were obtained upon plotting the DEPC modification extents, from all stress temperatures, on a single histogram and finding that they distribute into three ranges (see **Table 3.1** for details). Relatively few changes in DEPC labeling levels are found in rituximab samples stressed at 45 °C. Most of these changes are increases in labeling, as might be expected when the protein is heated and undergoes unfolding. In addition, the variable regions have far fewer changes than the constant domains. More than 70% of the labeling changes occur at Tyr, Ser, and Thr residues, and the labeling changes that occur at His and Lys residues are always less than 20%. While the extent of labeling of both sets of residues are influenced by changes in solvent accessibility, labeling of the weakly nucleophilic Tyr, Ser, and Thr residues is more sensitive to changes in local microenvironment. Thus, labeling changes predominantly in these residues suggests that the structural changes at 45 °C are primarily changes in the local microenvironment instead of large structural changes that lead to significant differences in solvent accessibility. Consistent with this idea is the fact that the labeling changes are scattered throughout the protein structure rather than clustered in certain regions of the protein (**Figure 3.8a**). DEPC labeling sites and levels for rituximab under native and the lower temperature thermally-stressed conditions (45 °C and 55 °C) are listed in full in **Tables B.1 and B.2 in Appendix B**, respectively.

Table 3.1: Changes in DEPC modification extents after the heat stress at 45 °C for 4 h.

The pie charts indicate the fraction of modified residues that undergo statistically significant labeling changes within each domain of rituximab. Red represents labeling increases while blue represents decreases. The bar charts for each domain indicate the number of residues whose extent of covalent labeling (CL) change falls within low (L), medium (M), and high (H) bins.

Domain	# Labeled residues with CL change / # Total modifiable residues		Extent of CL change	
(45 °C 4 h)	His, Lys	Ser, Thr, Tyr	His, Lys	Ser, Thr, Tyr
V_L	(0/6)	(0/33)		
C_L	(0/10)	(4/28)		M = 1, H = 3
V_H	(0/9)	(2/36)		M = 2
C_H1	(1/10)	(3/32)	L = 1	H = 2, 1
C_H2	(2/17)	(1/20)	M = 1, H = 1	H = 1
C_H3	(2/10)	(2/25)	L = 1, M = 1	M = 1, H = 1

	High	Medium	Low	Low	Medium	High
Absolute difference (His, Lys)	≥ 20%	5% to 20%	≤ 5%	≥ -5%	-5% to -20%	≤ -20%
Ratio difference (Ser, Thr, Tyr)	≥ 70%	40% to 70%	≤ 40%	≥ -40%	-40% to -70%	≤ -70%

Increased labeling   Decreased labeling

Figure 3.8: Sites on rituximab that undergo significant labeling changes after heat stress at (a) 45 °C for 4 h and (b) 55 °C for 4 h, as compared to non-stressed rituximab. Spheres represent residues that undergo significant changes in label levels ($p < 0.05$). Red represents labeling increases while blue represents decreases. Note that, for clarity, only one asymmetric unit of rituximab structure is labeled in this figure. As no full-length structures of rituximab are available in a protein data bank (PDB), we used existing F_{ab} and F_c crystal structures of rituximab to generate a molecular model for the entire rituximab molecule. A full-length human IgG1 model, from which atomic coordinates were generated using PDBs 2IG2 (F_c) and 1FC2 (F_{ab}) with a hinge region and other details theoretically modeled[75, 76], was used as a template. F_c (PDB 4W4N) and F_{ab} (PDB 4KAQ) structures of rituximab were then aligned to the template, using the molecular visualization system PyMOL.

Activity assays were conducted to determine if the DEPC labeling changes reflect significant enough structural changes to affect rituximab activity and thus validate the DEPC labeling results. Three types of activity assays were performed: (a) rituximab bridging ELISA, (b) Alamar blue assay, and (c) Raji cell pull-down assay (**Figure 3.9**). The bridging ELISA assay was used to evaluate binding of the F_c region, thereby reporting on the structural integrity of F_c region. In addition, complement-dependent cytotoxicity (CDC) activity of the F_c region was assessed using an Alamar blue assay.[70, 77] The structural integrity of F_{ab} region was evaluated using CD20-positive Raji B cells as part of a pull-down assay. These activity assays were conducted at different temperatures spanning the range of temperatures studied by DEPC labeling. Results from the bridging ELISA show that there is no significant change in F_c binding activity of rituximab after preheating to 45 °C (**Figure 3.10a**). CDC activity of F_c region is also estimated to remain unchanged after heat stress at 45 °C (**Figure 3.10b**). Similarly, the F_{ab} binding activity of rituximab stressed at 45 °C is estimated not to differ from control samples (**Figure 3.10c**). Together, these activity assays are consistent with the idea that the protein does not undergo significant structural changes at 45 °C, as indicated by the DEPC labeling experiments. The number of residues undergoing labeling changes is relatively small in both the F_{ab} and F_c regions and consists primarily of Tyr, Ser, and Thr residues that are more sensitive to local microenvironment changes. Moreover, the modification sites are scattered throughout the protein (**Figure 3.8a**), suggesting little effect on the conformation of the F_{ab} and F_c regions, as confirmed by the activity assays.
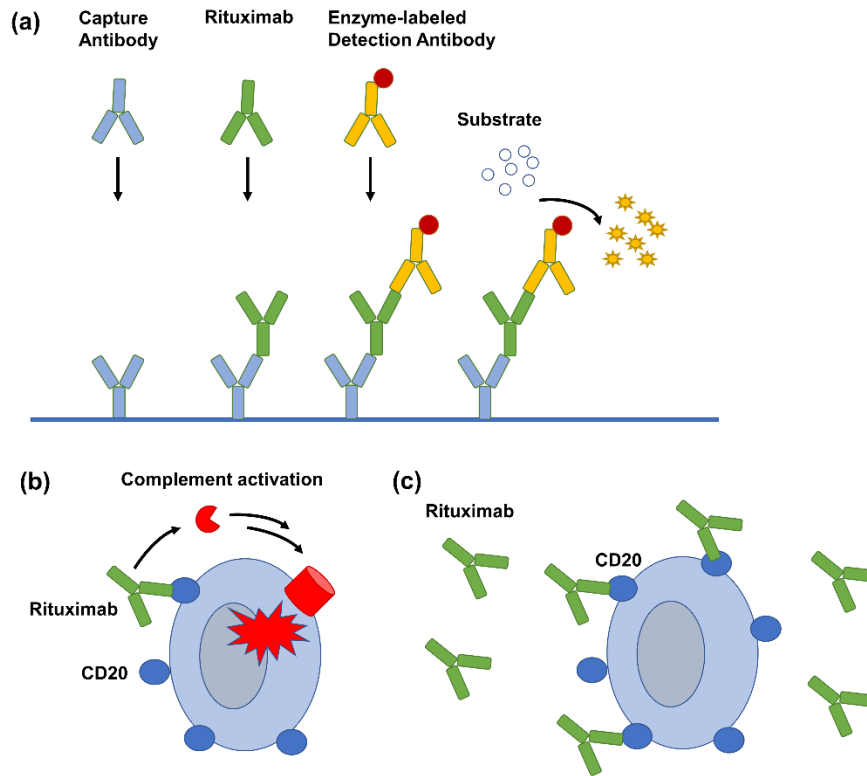
Figure 3.9: Schematics of three types of activity assays used in this study.
(a) The pharmacokinetic bridging format of the ELISA was used to evaluate binding of the F_c region to a capture antibody. The F_c of rituximab was first bound to an anti-human capture antibody, and the F_{ab} was then bound to an anti-rituximab detection antibody. (b) Complement-dependent cytotoxicity (CDC) activity of the F_c region was assessed using an Alamar blue assay. In this assay, Raji B cells were incubated with control or stressed rituximab. Rituximab binds to CD20 antigen expressed of cell surface, and cytotoxicity can be exerted by F_c after a normal human serum complement is added. CDC activity was then measured via a cell viability assay, allowing it to act as a functional indicator of the HOS of the F_c region. (c) The functional activity of F_{ab} binding to its antigen CD20 was measured using a Raji cell pull-down assay. A study has shown that rituximab is bound to the surface of CD20 cells *in vitro*.[71] In this experiment, CD20-positive Raji B cells were treated with control or stressed rituximab without human serum added, and the unbound rituximab was then quantified using ELISA This assay could measure binding activity of the F_{ab} region to its CD20 antigen on the cell surface.
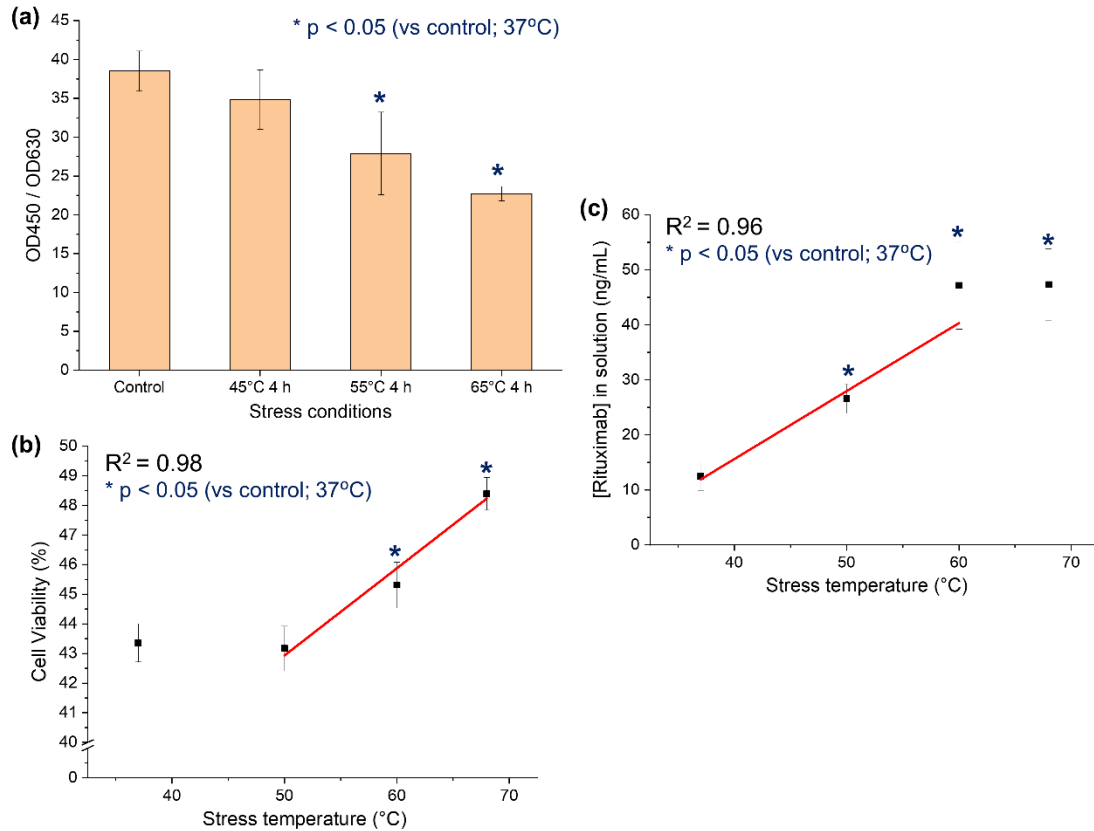
Figure 3.10: Structural changes revealed from DEPC CL-MS experiments are validated using rituximab activity assays.

The structural integrity of the F_c region is evaluated by (a) a rituximab bridging ELISA that measures F_c binding to a capture antibody and (b) an Alamar blue assay that measures complement dependent cytotoxicity (CDC). The structural integrity of the F_{ab} region is assessed by (c) a Raji cell pull-down assay that measures F_{ab} binding to CD-20 antigen on B cells. For each temperature condition, the means of two groups i.e. stressed vs. control (37 °C) were compared using student t-test at 95% confidence level. Each assay was performed in triplicate (n = 3).

More residues undergo significant changes in modification extent after heating at 55 °C for 4 h, and these changes can be found in all rituximab domains, particularly in the V_H and V_L domains of the F_{ab} region (**Table 3.2**). Most of the changes in each domain are increases in labeling, with an exception of the V_H and C_L domains where there is a similar number of labeling decreases. Almost twice as many labeling changes occur at His and Lys residues upon heating to 55 °C compared to heating at 45 °C, although most of the changes are characterized as low or medium changes. Most of the labeling changes at Tyr, Ser, and Thr are characterized as medium or high changes. Overall, these results indicate likely changes in protein topology, although the moderate labeling changes at His and Lys residues suggest somewhat modest structural changes. In contrast to the heat stress at 45 °C, after heating at 55 °C, the sites whose label levels undergo significant changes are found to cluster in the F_{ab} region of the protein, especially in V_H, V_L, and C_L domains (**Figure 3.8b**). Such clustering suggests localized structural changes in the F_{ab} region of rituximab. Similar clustering of residues undergoing labeling changes are not observed in the F_c region (**Figure 3.8b**).

Further insight into the meaning of the labeling data can be obtained via comparison to the activity assays. A small, but statistically significant, change in F_c binding activity from the bridging ELISA is measured at 55 °C (**Figure 3.10a**), but the CDC activity of the F_c region does not change significantly after heating at 55 °C (**Figure 3.10b**). The lack of consistency in the F_c-related assays might correlate with the lack of clustering of labeled residues in the F_c as seen in the DEPC labeling experiments (**Figure 3.10c**). One possible explanation is that the conformational changes in the F_c region are significant enough to influence the ELISA but too subtle to affect the Alamar blue assay. A significant change

in F_{ab} binding activity is observed from the Raji cell pull-down assay for rituximab preheated at 55 °C (**Figure 3.10c**). This result is very consistent with the observed labeling changes in the F_{ab} region that show significant clustering of residues that undergo labeling changes (**Figure 3.8b**).

Table 3.2: Changes in DEPC modification extents after the heat stress at 55 °C for 4 h.
The pie charts indicate the fraction of modified residues that undergo statistically significant labeling changes within each domain of rituximab. Red represents labeling increases while blue represents decreases. The bar charts for each domain indicate the number of residues whose extent of covalent labeling (CL) change falls within low (L), medium (M), and high (H) bins.

Domain	# Labeled residues with CL change / # Total modifiable residues		Extent of CL change	
(55 °C 4 h)	His, Lys	Ser, Thr, Tyr	His, Lys	Ser, Thr, Tyr
V_L	(2/6)	(6/33)	2 at L	3 at H; 1,2 at M
C_L	(1/10)	(3/28)	1 at M	1 at L; 2 at M
V_H	(4/9)	(4/36)	1 at L; 2,1 at M	3 at L; 1 at M
C_H1	(0/10)	(1/32)		1 at H
C_H2	(1/17)	(2/20)	1 at H	1 at M; 1 at H
C_H3	(1/10)	(6/25)	1 at H	1 at L; 5 at H

	High	Medium	Low	Low	Medium	High
Absolute difference (His, Lys)	≥ 20%	5% to 20%	≤ 5%	≥ -5%	-5% to -20%	≤ -20%
Ratio difference (Ser, Thr, Tyr)	≥ 70%	40% to 70%	≤ 40%	≥ -40%	-40% to -70%	≤ -70%

Increased labeling   **Decreased labeling**

3.3.4 Investigation of Conformation Change upon Higher Heat Stress

As indicated earlier, heating at 65 °C leads to greater changes in the fluorescence, CD, and DLS data. In fact, heating at this temperature for 4 h results in a cloudy sample, and DLS data suggest the formation of small protein aggregates, as indicated by a small percentage of species at 100 nm (**Figure 3.2c**). Additional support for this conclusion is also found from size-exclusion chromatography (SEC) measurements (**Figure 3.11**), which reveal the presence of high molecular weight species upon heating at 65 °C and not upon heating at 45 °C or 55 °C. Upon applying our DEPC CL-MS method to rituximab thermally-stressed at 65 °C, we find significant labeling changes in all rituximab domains (**Table 3.3 and Table B.3**). Most notably, many more decreases in labeling are observed than increases in labeling, which is likely explained by protein aggregation. Significant decreases in labeling are found in both the F_{ab} and F_c regions, while only a few residues in the F_{ab} region undergo increases in labeling. Around 30% of the labeling changes at 65 °C are found at His and Lys residues, and most of these can be characterized as medium or high changes in labeling extent. For Tyr, Ser, and Thr residues, most sites undergo medium or high extents of labeling changes. Overall, these results suggest more profound changes in the solvent accessibility and the local microenvironment for many sites in the protein upon heating at 65 °C. When the residues that undergo changes in labeling are mapped on the rituximab structure, clusters of residues are found in the C_H3 domain of F_c region and the variable domains of the F_{ab} region (**Figure 3.12)**.

The labeling changes are consistent with activity assays that indicate changes in the F_{ab} and F_c regions. Results from the bridging ELISA show that there is a significant change in F_c binding after heating to 65 °C (**Figure 3.10a**). CDC activity of the F_c region (**Figure

3.10b) and F_{ab} binding activity (**Figure 3.10c**) are also significantly different from control samples. The activity assay results are consistent with the DEPC labeling results, indicating significant structural changes to rituximab in both F_{ab} and F_c regions.
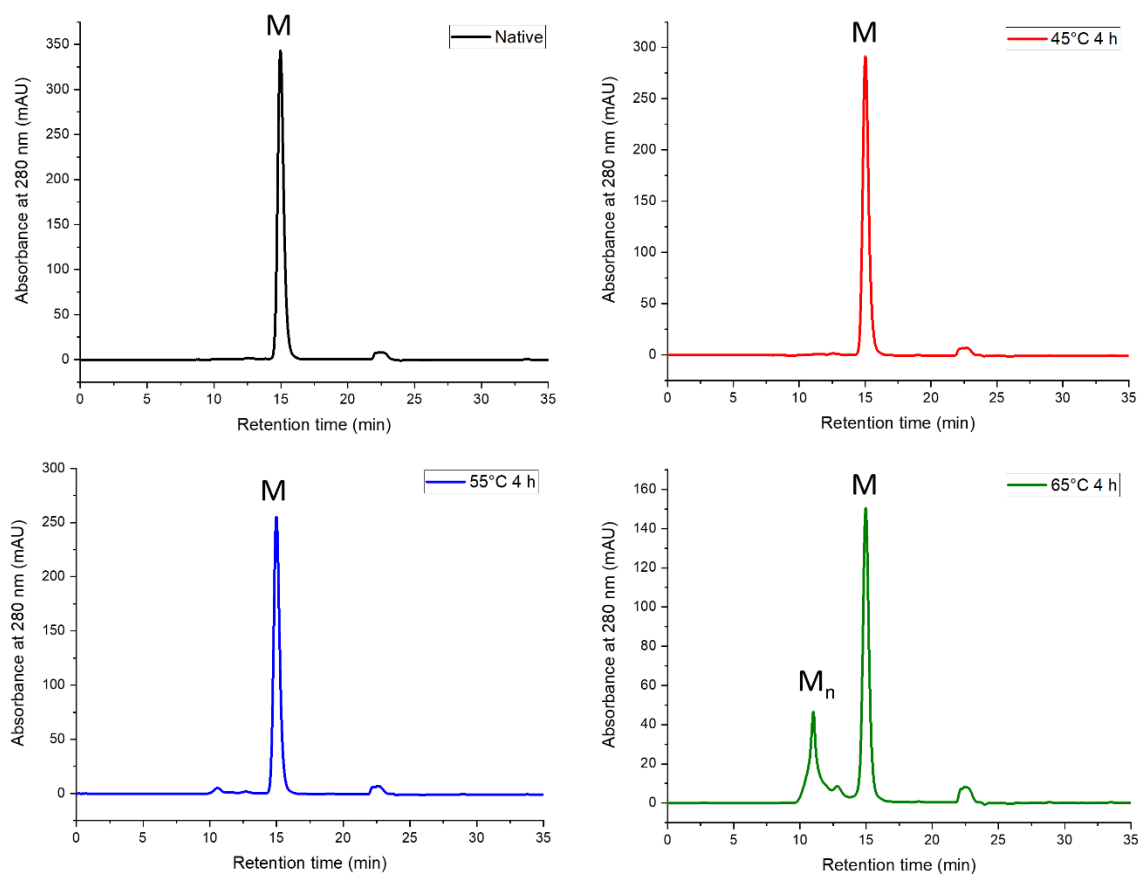
Figure 3.11: Size-exclusion chromatography (SEC) of rituximab at 37 °C (native, black) and thermally-stressed conditions, after incubation of the rituximab formulation at 45 °C for 4 h (red), 55 °C for 4 h (blue), and 65 °C for 4 h (green). After heating at 65 °C, the chromatogram of the stressed rituximab indicates the significant formation of high molecular weight species (i.e. M_n), whereas the other temperatures show almost exclusively monomeric rituximab (i.e. M).

Table 3.3: Changes in DEPC modification extents after the heat stress at 65 °C for 4 h.

The pie charts indicate the fraction of modified residues that undergo statistically significant labeling changes within each domain of rituximab. Red represents labeling increases while blue represents decreases. The bar charts for each domain indicate the number of residues whose extent of covalent labeling (CL) change falls within low (L), medium (M), and high (H) bins.

Domain	# Labeled residues with CL change / # Total modifiable residues		Extent of CL change	
(65 ºC 4 h)	His, Lys	Ser, Thr, Tyr	His, Lys	Ser, Thr, Tyr
V_L	(0/6)	(3/33)	—	M=2, H=1
C_L	(1/10)	(0/28)	L=1	—
V_H	(3/9)	(3/36)	M=1, H=1, H=1	L=2, H=1
C_H1	(1/10)	(4/32)	L=1	L=1, H=3
C_H2	(0/17)	(2/20)	—	H=2
C_H3	(2/10)	(3/25)	L=1, H=1	M=3

	High	Medium	Low	Low	Medium	High
Absolute difference (His, Lys)	≥ 20%	5% to 20%	≤ 5%	≥ -5%	-5% to -20%	≤ -20%
Ratio difference (Ser, Thr, Tyr)	≥ 70%	40% to 70%	≤ 40%	≥ -40%	-40% to -70%	≤ -70%

Increased labeling — Decreased labeling

Figure 3.12: Sites on rituximab that undergo significant (a) increases and (b) decreases in DEPC modification after heat stress at 65 °C for 4 h, as compared to non-stressed rituximab.

Spheres represent residues that undergo significant changes in label levels ($p < 0.05$). Red represents labeling increases while blue represents decreases. Note that, for clarity, only one asymmetric unit of rituximab structure is labeled in this figure. (See **Figure 3.8** for more details about rituximab's molecular model.)

3.4 Discussion

In this study, we have shown the applicability of DEPC CL-MS as an analytical tool to provide site-specific information about changes in the HOS of protein therapeutics. Sites of the structural changes are revealed from differences in DEPC modification levels compared to a control sample. The full meaning of the quantitative changes in labeling is not fully understood at this stage, but labeling increases often suggest greater unfolding in a given region and labeling decreases imply burial of side chains usually from aggregation. Changes in the solvent-accessible surface area (SASA) of a given residue is a primary factor that governs side chain reactivity,[36-38] but changes in reactivity can also be caused by changes in local tertiary structure around a given residue (e.g., microenvironment).[38]

Previous studies have identified the melting temperatures (T_m) for different domains of rituximab to be 71 °C, 74 °C, and 83 °C for the C_H2, F_{ab}, and C_H3 domains, respectively.[78] We stressed the rituximab samples at temperatures far below these T_m values as we were interested in studying the protein under mild to moderate thermal stress to see if DEPC CL-MS could report on any subtle structural changes. These lower temperatures more realistically mimic stresses that protein therapeutics may undergo.[79] Results from DEPC CL-MS after heating at 45 °C suggest that no significant structural changes occur at this temperature, which is consistent with fluorescence, CD, and DLS measurements as well as the activity assays. Even though a few residues undergo statistically significant labeling changes, the number of labeling changes that occur at His and Lys residues, which are the better reporters of SASA changes, is relatively small. Most of the labeling changes occur at Tyr, Ser, and Thr residues (**Table 3.1**), which are sites that are more sensitive to local microenvironment changes such as changes in H-bonding or

nearby electrostatic interactions. While the extent of DEPC modification of both sets of residues are influenced by SASA, the labeling changes that occur at the weakly nucleophilic Tyr, Ser, and Thr residues are much less correlated with the changes in SASA.[38] Similar observations are made in HRF experiments, where the primary structure and HOS more significantly influence residues with poor intrinsic reactivity.[36] The predominance of labeling changes at Tyr, Ser, and Thr residues suggests that any HOS changes at 45 °C are changes in the local microenvironment instead of significant unfolding that would dramatically change the SASA of residues. Moreover, the labeling changes are scattered throughout the protein structure rather than clustered at certain regions of the protein (**Figure 3.8a**), again implying very local effects. The data from the activity assays (**Figure 3.10**) are in agreement with this idea.

A greater number of labeling changes occur upon heating rituximab to 55 °C than at 45 °C, and these labeling changes are clustered in the F_{ab} domain of the protein. This extent and localized nature of labeling changes indicate that CL-MS is revealing structural perturbations, changes that are not measurable by fluorescence, CD, and DLS. Around 30% of the labeling changes at 55 °C occur at His and Lys residues, and these changes are mostly moderate or high increases in labeling (**Table 3.2**), indicating that the protein is unfolding in these regions. As alluded to earlier, the reactivity of the more nucleophilic His and Lys residues is better correlated with SASA, and thus the labeling of these residues are more sensitive to local unfolding events that would expose them to a greater extent.[35] It is important to note that most of the His and Lys labeling changes at 55 °C are classified as moderate, meaning there is less than a 20% increase in labeling. While we have not established a quantitative relationship between labeling extents and SASA, the fact that the

biophysical techniques do not detect any structural changes suggests that any structural changes that do occur after heating to 55 °C are subtle. A larger number of Tyr, Ser, and Thr residues also undergo labeling changes after stressing at 55 °C than at 45 °C, and most are moderate or high increases (**Table 3.2**). Again, these residues are less correlated with SASA changes, but the significant number of labeling changes to these residues also suggests a structural change.

The locations of the structural changes after heating to 55 °C are worth noting. The most significant and clustered labeling changes are found in the F_{ab} region, which is reported to be the most sensitive region to heat stress in IgG1 molecules.[80] Our results would seem to contrast with DSC data for rituximab obtained by Andersen *et al.* in which the C_H2 domain is found to have the lowest T_m of 71 °C;[78] however, this difference is almost certainly due to differences in the nature of our labeling experiments and the DSC experiments. In our experiments, the protein reaches a thermal equilibrium after 4 h of heating at 55 °C before being cooled and analyzed, whereas the DSC experiment involves a temperature ramp where the protein is exposed to a given temperature for only a minute without associated cooling.[78] Because different protein domains can change structures at different rates and re-fold to different extents upon cooling, one might not expect there to be a strong correlation between our heat-stress experiments and DSC experiments. Very telling is the fact that several residues that undergo significant labeling changes cluster in the variable domains, V_H and V_L, where the complementarity-determining regions (CDRs) for CD-20 are located (**Figure 3.8b and Table 3.2**).[81] Among the residues undergoing significant labeling changes at 55 °C are Thr91 on the LC, located in the middle of CDR L3, His33 on the LC and His35 and Lys74 on the HC, which are located close to CDRs

L1, H1, and H2, respectively **(Figure 3.13b)**. Structural changes near these residues, as indicated by the labeling changes, are nicely consistent with the decreased F_{ab} binding that was found in the activity assay (**Figure 3.10c**). To the best of our knowledge, our CL-MS technique is the first to report a subtle HOS change of a therapeutic mAb at the temperature as low as 55 °C.
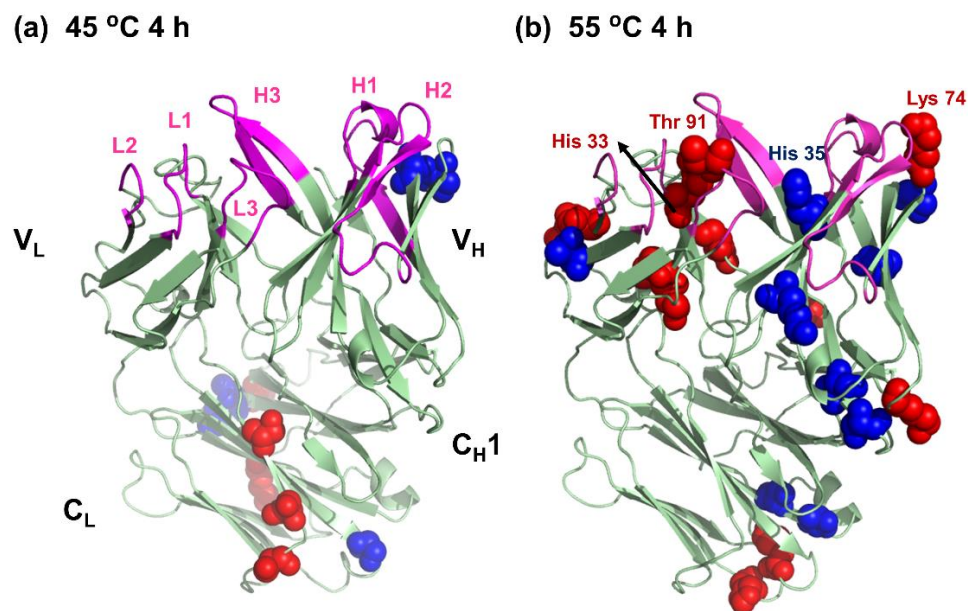
Figure 3.13: Illustration of rituximab's F_{ab} domain highlighting structural changes that may affect CDRs after heat stress at (a) 45 °C for 4 h and (b) 55 °C for 4 h. Red represents the increase in label level while blue represents the decrease in label level. Six CDRs of rituximab (H1-H3 in heavy chain and L1-L3 in light chain) have been previously reported in Du *et al.*,[81] and they are highlighted in magenta. Thr91 (light chain), His33 (light chain), His35 and Lys74 (heavy chain) which sit in or nearby CDRs are found to undergo significant changes in labeling after heat stress at 55 °C. (PDB accession code 4KAQ)

Heating at 65 °C more significantly alters the conformation of rituximab, as revealed by biophysical measurements, activity assays, and CL-MS. Indeed, some fraction of the protein aggregates upon heating for 4 h at this temperature, which is consistent with the propensity of rituximab to transiently unfold at temperatures below its T_m.[60] Aggregation is also indicated by the fact that most of the measured changes in CL-MS are

moderate or high labeling decreases (**Table 3.3**). It may seem surprising that the 65 °C data does not have a more even mix of labeling increases and decreases because the protein probably unfolds to some degree in addition to aggregating. It should be noted, though, that our technique is inherently much more sensitive to aggregation events than unfolding events. The amino acids that are labeled by DEPC are primarily polar residues that are more likely to reside on the protein surface anyway. Consequently, protein unfolding events only mildly increase their SASA and DEPC reactivity, while aggregation events completely bury these normally exposed residues, dramatically changing their SASA and DEPC reactivity. If a given residue is more exposed in some protein molecules but buried in other protein molecules, the decrease in labeling due to residue burial will likely counterbalance any labeling increase due to increased solvent exposure.

This greater sensitivity to aggregation events gives us excellent insight into regions of the protein that self-associate, as we have shown previously for other proteins.[52, 53, 56] Clusters of residues that undergo decreases in labeling are found in both F_{ab} and F_c regions (**Figure 3.12**). Structural changes in the F_{ab} region are consistent with a reduction in F_{ab} binding to CD20 (**Figure 3.10c**), and many labeling changes are found near the CDRs. Among the labeling changes are Tyr70 on the LC and His35, Lys67, and Lys74 on the HC, which sit close to CDRs L2, H1, H2 and H2, respectively (**Figure 3.14a**). Labeling changes in the F_c region, particularly the C_H3 domain are also confirmed by a reduction in F_c binding and CDC activity (**Figure 3.10**). Clusters of residues that undergo decreases in DEPC labeling in the C_H3 domain of F_c (Ser258, Ser428, His439, and Tyr440) and the variable domains of F_{ab} (Ser5, Tyr70, and Tyr101 on the LC and Thr118 and Ser120 on the HC) suggest these regions as the likely aggregation interfaces (**Figure 3.14b**).
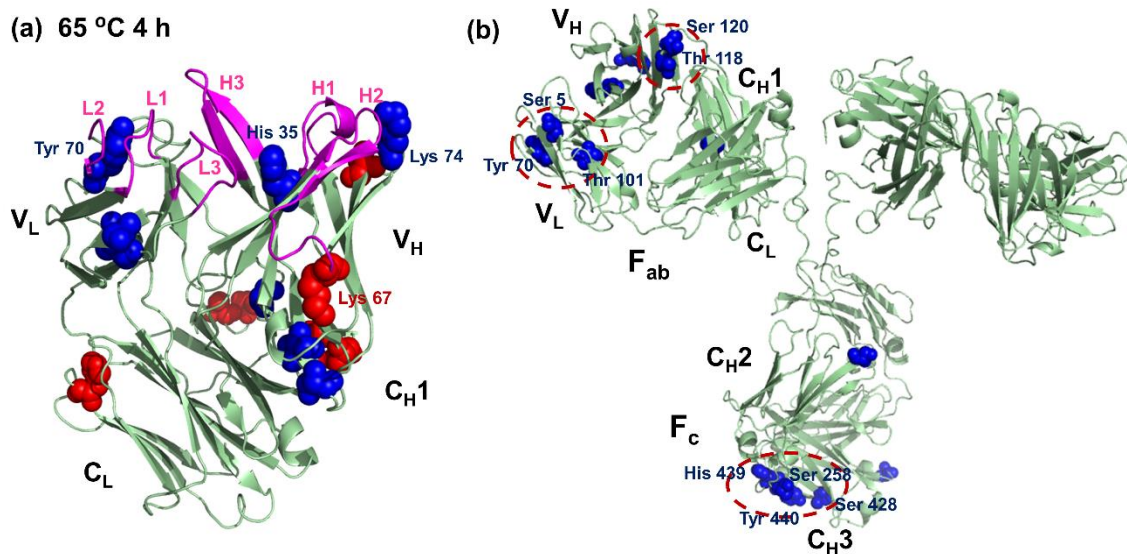
Figure 3.14: Illustration of rituximab structure highlighting structural changes that may affect CDRs and the possible aggregation sites after heat stress at 65 °C for 4 h. (a) Cartoon representation of the F_{ab} domain after the heat stress. Red represents the increase in label level while blue represents the decrease in label level. Six CDRs of rituximab (H1-3 in heavy chain and L1-3 in light chain) have been previously reported in Du *et al.*,[81] and they are highlighted in magenta. Tyr70 (light chain), His35, Lys67, and Lys74 (heavy chain), which sit nearby CDRs, are found to undergo significant changes in labeling after heat stress at 65 °C. (PDB accession code 4KAQ) (b) Cartoon representation of possible aggregation sites in the F_{ab} and F_c regions. Only decreases in labeling are shown in this figure. Those residues cluster to each other with a distance less than 12 Å. Note that only one asymmetric unit of rituximab structure is labeled in this figure. (See **Figure 3.8** for more details about rituximab's molecular model.)

Our DEPC CL-MS technique compares favorably to other MS-based structural techniques (e.g., HDX,[28-30] HRF,[48] and dimethyl labeling[82]) that have been used recently to study HOS changes or identify aggregation interfaces of mAbs. Most of the previous studies used harsh stress conditions (freeze-thaw, UV light, or temperature around the T_m) to induce conformation changes to the mAbs under study. Consistent with the results reported in our work here, most aggregation sites in mAbs are found to be in the F_{ab} region, particularly in the variable domains,[28, 30, 48, 56, 82] although there is at least one example of aggregation involving the F_c region.[29] To our knowledge, there are no published

experimental studies that report the aggregation interface for rituximab upon heating. Nevertheless, computational tools have predicted aggregation prone regions[83-85] in F_{ab} and F_c regions, of which many are located close to the sites found to decrease in labeling, such as Ser5 (LC), Thr118 and Ser120 (HC, F_{ab}) and Ser258 (HC, F_c).

From a methodological perspective, CL-MS using DEPC has some advantages over other MS-based techniques, especially for revealing aggregation interfaces. In HDX-MS, reduced deuterium exchange along a protein backbone is used to identify aggregation sites, even though protein-protein interactions are primarily mediated by side chain interactions. DEPC labeling reports on the SASA of side chains, thereby providing a more direct indication of aggregation sites. CL-MS also does not require specialized equipment like HDX-MS, thereby simplifying the overall approach. HRF is another side-chain labeling technique that has been used for studying mAb interactions and structure, but HRF techniques also require specialized lasers[41, 45, 46, 86, 87] or synchrotron[40, 48] sources to generate the radicals. In addition, radicals generate multiple reaction products, as opposed to a single one with DEPC, which can decrease the sensitivity of the technique.

3.5 Conclusion

In conclusion, we demonstrate here the ability of DEPC CL-MS to investigate HOS of antibody therapeutics. DEPC directly probes solvent accessibility and the microenvironment of side chains, and any changes in DEPC reactivity are indicative of changes in protein conformation. With a broad spectrum of modifiable residues coupled with the sensitivity of MS, the structural resolution and sensitivity of our DEPC labeling technique is high enough for probing subtle protein conformational changes. We have

shown that DEPC CL-MS can reveal subtle HOS changes in rituximab at the temperatures below the T_m (e.g., 55 °C), temperatures at which classical biophysical techniques are not sensitive enough to detect any changes. The results from DEPC labeling are validated via activity assays. At higher heat stress (e.g., 65 °C), HOS changes and aggregation are apparent from clusters of residues that undergo decreases in labeling in both F_{ab} and F_c regions. These results are confirmed by complementary biophysical and activity measurements. With its structural resolution, sensitivity, and simplicity, DEPC CL-MS should be amenable to the structural investigations of other antibody therapeutics. Future work will demonstrate its applicability to HOS characterization of other therapeutic antibodies and antibody-antigen complexes.

3.6 References

1.   Ecker, D. M.; Jones, S. D.; Levine, H. L., The therapeutic monoclonal antibody market. *mAbs* **2015,** *7* (1), 9-14.
2.   Lindsley, C. W., New 2017 Data and Statistics for Pharmaceutical Products. *ACS Chemical Neuroscience* **2018,** *9* (7), 1518-1519.
3.   Kaplon, H.; Reichert, J. M., Antibodies to watch in 2019. *mAbs* **2018**, In Press, DOI: 10.1080/19420862.2018.1556465.
4.   EvaluatePharma® *World Preview 2017, Outlook to 2022*; Evaluate Ltd: London, England, United Kingdom, 2017; p 8.
5.   Frokjaer, S.; Otzen, D. E., Protein drug stability: a formulation challenge. *Nat Rev Drug Discov* **2005,** *4* (4), 298-306.
6.   Berkowitz, S. A.; Engen, J. R.; Mazzeo, J. R.; Jones, G. B., Analytical tools for characterizing biopharmaceuticals and the implications for biosimilars. *Nat Rev Drug Discov* **2012,** *11* (7), 527-540.
7.   Weiss, W. F.; Gabrielson, J. P.; Al-Azzam, W.; Chen, G.; Davis, D. L.; Das, T. K.; Hayes, D. B.; Houde, D.; Singh, S. K., Technical Decision Making With Higher Order Structure Data: Perspectives on Higher Order Structure Characterization From the Biopharmaceutical Industry. *Journal of Pharmaceutical Sciences* **2016,** *105* (12), 3465-3470.
8.   Kaltashov, I. A.; Bobst, C. E.; Abzalimov, R. R.; Wang, G.; Baykal, B.; Wang, S., Advances and challenges in analytical characterization of biotechnology products: Mass spectrometry-based approaches to study properties and behavior of protein therapeutics. *Biotechnology Advances* **2012,** *30* (1), 210-222.

9.   Kaltashov, I. A.; Bobst, C. E.; Abzalimov, R. R.; Berkowitz, S. A.; Houde, D., Conformation and Dynamics of Biopharmaceuticals: Transition of Mass Spectrometry-Based Tools from Academe to Industry. *Journal of the American Society for Mass Spectrometry* **2010,** *21* (3), 323-337.

10.   Beck, A.; Diemer, H.; Ayoub, D.; Debaene, F.; Wagner-Rousset, E.; Carapito, C.; Van Dorsselaer, A.; Sanglier-Cianférani, S., Analytical characterization of biosimilar antibodies and Fc-fusion proteins. *TrAC Trends in Analytical Chemistry* **2013,** *48* (Supplement C), 81-95.

11.   Jung, S. K.; Lee, K. H.; Jeon, J. W.; Lee, J. W.; Kwon, B. O.; Kim, Y. J.; Bae, J. S.; Kim, D.-I.; Lee, S. Y.; Chang, S. J., Physicochemical characterization of Remsima®. *mAbs* **2014,** *6* (5), 1163-1177.

12.   Magnenat, L.; Palmese, A.; Fremaux, C.; D'Amici, F.; Terlizzese, M.; Rossi, M.; Chevalet, L., Demonstration of physicochemical and functional similarity between the proposed biosimilar adalimumab MSB11022 and Humira®. *mAbs* **2017,** *9* (1), 127-139.

13.   Seo, N.; Polozova, A.; Zhang, M.; Yates, Z.; Cao, S.; Li, H.; Kuhns, S.; Maher, G.; McBride, H. J.; Liu, J., Analytical and functional similarity of Amgen biosimilar ABP 215 to bevacizumab. *mAbs* **2018,** *10* (4), 678-691.

14.   Lee, K. H.; Lee, J.; Bae, J. S.; Kim, Y. J.; Kang, H. A.; Kim, S. H.; Lee, S. J.; Lim, K. J.; Lee, J. W.; Jung, S. K.; Chang, S. J., Analytical similarity assessment of rituximab biosimilar CT-P10 to reference medicinal product. *mAbs* **2018,** *10* (3), 380-396.

15.   Nupur, N.; Chhabra, N.; Dash, R.; Rathore, A. S., Assessment of structural and functional similarity of biosimilar products: Rituximab as a case study. *mAbs* **2018,** *10* (1), 143-158.

16.   Ambrogelly, A.; Gozo, S.; Katiyar, A.; Dellatore, S.; Kune, Y.; Bhat, R.; Sun, J.; Li, N.; Wang, D.; Nowak, C.; Neill, A.; Ponniah, G.; King, C.; Mason, B.; Beck, A.; Liu, H., Analytical comparability study of recombinant monoclonal antibody therapeutics. *mAbs* **2018,** *10* (4), 513-538.

17.   Flores-Ortiz, L. F.; Campos-García, V. R.; Perdomo-Abúndez, F. C.; Pérez, N. O.; Medina-Rivero, E., Physicochemical Properties of Rituximab. *Journal of Liquid Chromatography & Related Technologies* **2014,** *37* (10), 1438-1452.

18.   Hickey, J. M.; Toprani, V. M.; Kaur, K.; Mishra, R. P. N.; Goel, A.; Oganesyan, N.; Lees, A.; Sitrin, R.; Joshi, S. B.; Volkin, D. B., Analytical Comparability Assessments of 5 Recombinant CRM197 Proteins From Different Manufacturers and Expression Systems. *Journal of Pharmaceutical Sciences* **2018,** *107* (7), 1806-1819.

19.   Thiagarajan, G.; Semple, A.; James, J. K.; Cheung, J. K.; Shameem, M., A comparison of biophysical characterization techniques in predicting monoclonal antibody stability. *mAbs* **2016,** *8* (6), 1088-1097.

20.   ICH Harmonised Tripartite Guideline. Q5E Comparability of Biotechnological/Biological Products Subject to Changes in their Manufacturing Process. 2004 Nov [accessed 2016 Mar 12]. http://www.ich.org/products/guidelines/quality/quality-single/article/comparability-of-biotechnologicalbiological-products-subject-to-changes-in-their-manufacturing-proc.html.

21.   ICH Harmonised Tripartite Guideline. Q6B Specifications: Test Procedures and Acceptance Criteria for Biotechnological/Biological Products.1999 Mar [accessed 2016 Mar 12]. http://www.ich.org/products/guidelines/quality/quality-single/article/specifications-test-procedures-and-acceptance-criteria-for-biotechnologicalbiological-products.html.

22.   U.S. Food and Drug Administration. Quality Considerations in Demonstrating Biosimilarity of a Therapeutic Protein Product to a Reference Product to a Reference Product; 2015 Apr [accessed 2017 Sep 25]. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291134.pdf.

23.   U.S. Food and Drug Administration. Points to Consider in the Manufacture and Testing of Monoclonal Antibody Products for Human Use. 1997 Feb 28 [accessed 2017 Sep 23]. https://www.fda.gov/downloads/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/OtherRecommendationsforManufacturers/UCM153182.pdf.

24.   U.S. Food and Drug Administration. Guidance concerning demonstration of comparability of human biological products, including therapeutic biotechnologyderived products. 1996 Apr [accessed 2018 Jul 21]. https://www.fda.gov/drugs/guidancecomplianceregulatoryinformation/guidances/ucm122879.htm.

25.   European Medicines Agency. Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: quality issues (revision 1). 2014 Jun 3 [accessed 2017 Sep 25]. http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000886.jsp&mid=WC0b01ac058002956b.

26.   European Medicines Agency. Guideline on development, production, characterisation and specification for monoclonal antibodies and related products. 2016 Aug 4 [accessed 2017 Sep 25]. http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000867.jsp&mid=WC0b01ac058002956b.

27.   Rogstad, S.; Faustino, A.; Ruth, A.; Keire, D.; Boyne, M.; Park, J., A Retrospective Evaluation of the Use of Mass Spectrometry in FDA Biologics License Applications. *Journal of The American Society for Mass Spectrometry* **2017,** *28* (5), 786-794.

28.   Zhang, A.; Singh, S. K.; Shirts, M. R.; Kumar, S.; Fernandez, E. J., Distinct Aggregation Mechanisms of Monoclonal Antibody Under Thermal and Freeze-Thaw Stresses Revealed by Hydrogen Exchange. *Pharmaceutical Research* **2012,** *29* (1), 236-250.

29.   Iacob, R. E.; Bou-Assaf, G. M.; Makowski, L.; Engen, J. R.; Berkowitz, S. A.; Houde, D., Investigating Monoclonal Antibody Aggregation Using a Combination of H/DX-MS and Other Biophysical Measurements. *Journal of Pharmaceutical Sciences* **2013,** *102* (12), 4315-4329.

30.   Bommana, R.; Chai, Q.; Schöneich, C.; Weiss, W. F.; Majumdar, R., Understanding the Increased Aggregation Propensity of a Light-Exposed IgG1 Monoclonal Antibody Using Hydrogen Exchange Mass Spectrometry, Biophysical

Characterization, and Structural Analysis. *Journal of Pharmaceutical Sciences* **2018,** *107* (6), 1498-1511.

31.   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.

32.   Majumdar, R.; Middaugh, C. R.; Weis, D. D.; Volkin, D. B., Hydrogen-Deuterium Exchange Mass Spectrometry as an Emerging Analytical Tool for Stabilization and Formulation Development of Therapeutic Monoclonal Antibodies. *Journal of Pharmaceutical Sciences* **2015,** *104* (2), 327-345.

33.   Bonnington, L.; Lindner, I.; Gilles, U.; Kailich, T.; Reusch, D.; Bulau, P., Application of Hydrogen/Deuterium Exchange-Mass Spectrometry to Biopharmaceutical Development Requirements: Improved Sensitivity to Detection of Conformational Changes. *Analytical Chemistry* **2017,** *89* (16), 8233-8237.

34.   Mendoza, V. L.; Vachet, R. W., Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spectrom Rev* **2009,** *28* (5), 785-815.

35.   Mendoza, V. L.; Vachet, R. W., Protein Surface Mapping Using Diethylpyrocarbonate with Mass Spectrometric Detection. *Analytical Chemistry* **2008,** *80* (8), 2895-2904.

36.   Xie, B.; Sood, A.; Woods, R. J.; Sharp, J. S., Quantitative Protein Topography Measurements by High Resolution Hydroxyl Radical Protein Footprinting Enable Accurate Molecular Model Selection. *Scientific Reports* **2017,** *7* (1), 4552.

37.   Huang, W.; Ravikumar, Krishnakumar M.; Chance, Mark R.; Yang, S., Quantitative Mapping of Protein Structure by Hydroxyl Radical Footprinting-Mediated Structural Mass Spectrometry: A Protection Factor Analysis. *Biophysical Journal* **2015,** *108* (1), 107-115.

38.   Limpikirati, P.; Liu, T.; Vachet, R. W., Covalent labeling-mass spectrometry with non-specific reagents for studying protein structure and interactions. *Methods* **2018,** *144*, 79-93.

39.   Kaur, P.; Tomechko, S. E.; Kiselar, J.; Shi, W.; Deperalta, G.; Wecksler, A. T.; Gokulrangan, G.; Ling, V.; Chance, M. R., Characterizing monoclonal antibody structure by carboxyl group footprinting. *mAbs* **2015,** *7* (3), 540-552.

40.   Kaur, P.; Kiselar, J.; Shi, W.; Yang, S.; Chance, M. R., Covalent Labeling Techniques for Characterizing Higher Order Structure of Monoclonal Antibodies. In *State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 3. Defining the Next Generation of Analytical and Biophysical Techniques*, American Chemical Society: 2015; Vol. 1202, pp 45-73.

41.   Li, K. S.; Chen, G.; Mo, J.; Huang, R. Y. C.; Deyanova, E. G.; Beno, B. R.; O'Neil, S. R.; Tymiak, A. A.; Gross, M. L., Orthogonal Mass Spectrometry-Based Footprinting for Epitope Mapping and Structural Characterization: The IL-6 Receptor upon Binding of Protein Therapeutics. *Analytical Chemistry* **2017,** *89* (14), 7742-7749.

42.   Wecksler, A. T.; Kalo, M. S.; Deperalta, G., Mapping of Fab-1:VEGF Interface Using Carboxyl Group Footprinting Mass Spectrometry. *Journal of The American Society for Mass Spectrometry* **2015,** *26* (12), 2077-2080.

43.   Pan, L. Y.; Salas-Solano, O.; Valliere-Douglass, J. F., Localized conformational interrogation of antibody and antibody-drug conjugates by site-specific carboxyl group footprinting. *mAbs* **2017,** *9* (2), 307-318.

44.   Xu, G.; Chance, M. R., Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics. *Chem. Rev.* **2007,** *107* (8), 3514-3543.

45.   Li, J.; Wei, H.; Krystek, S. R.; Bond, D.; Brender, T. M.; Cohen, D.; Feiner, J.; Hamacher, N.; Harshman, J.; Huang, R. Y. C.; Julien, S. H.; Lin, Z.; Moore, K.; Mueller, L.; Noriega, C.; Sejwal, P.; Sheppard, P.; Stevens, B.; Chen, G.; Tymiak, A. A.; Gross, M. L.; Schneeweis, L. A., Mapping the Energetic Epitope of an Antibody/Interleukin-23 Interaction with Hydrogen/Deuterium Exchange, Fast Photochemical Oxidation of Proteins Mass Spectrometry, and Alanine Shave Mutagenesis. *Analytical Chemistry* **2017,** *89* (4), 2250-2258.

46.   Zhang, Y.; Wecksler, A. T.; Molina, P.; Deperalta, G.; Gross, M. L., Mapping the Binding Interface of VEGF and a Monoclonal Antibody Fab-1 Fragment with Fast Photochemical Oxidation of Proteins (FPOP) and Mass Spectrometry. *Journal of The American Society for Mass Spectrometry* **2017,** *28* (5), 850-858.

47.   Opuni Kwabena, F. M.; Al-Majdoub, M.; Yefremova, Y.; El-Kased Reham, F.; Koy, C.; Glocker Michael, O., Mass spectrometric epitope mapping. *Mass Spectrometry Reviews* **2016,** *37* (2), 229-241.

48.   Deperalta, G.; Alvarez, M.; Bechtel, C.; Dong, K.; McDonald, R.; Ling, V., Structural analysis of a therapeutic monoclonal antibody dimer by hydroxyl radical footprinting. *mAbs* **2013,** *5* (1), 86-101.

49.   Trinquier, G.; Sanejouand, Y. H., Which effective property of amino acids is best preserved by the genetic code? *Protein Engineering, Design and Selection* **1998,** *11* (3), 153-169.

50.   Zhou, Y.; Vachet, R. W., Increased Protein Structural Resolution from Diethylpyrocarbonate-based Covalent Labeling and Mass Spectrometric Detection. *Journal of The American Society for Mass Spectrometry* **2012,** *23* (4), 708-717.

51.   Zhou, Y.; Vachet, R. W., Diethylpyrocarbonate Labeling for the Structural Analysis of Proteins: Label Scrambling in Solution and How to Avoid It. *Journal of The American Society for Mass Spectrometry* **2012,** *23* (5), 899-907.

52.   Mendoza, V. L.; Barón-Rodríguez, M. A.; Blanco, C.; Vachet, R. W., Structural Insights into the Pre-Amyloid Tetramer of β-2-Microglobulin from Covalent Labeling and Mass Spectrometry. *Biochemistry* **2011,** *50* (31), 6711-6722.

53.   Mendoza, V. L.; Antwi, K.; Barón-Rodríguez, M. A.; Blanco, C.; Vachet, R. W., Structure of the Preamyloid Dimer of β-2-Microglobulin from Covalent Labeling and Mass Spectrometry. *Biochemistry* **2010,** *49* (7), 1522-1532.

54.   Srikanth, R.; Mendoza, V. L.; Bridgewater, J. D.; Zhang, G.; Vachet, R. W., Copper Binding to β-2-Microglobulin and Its Pre-Amyloid Oligomers. *Biochemistry* **2009,** *48* (41), 9871-9881.

55.   Liu, T.; Marcinko, T. M.; Kiefer, P. A.; Vachet, R. W., Using Covalent Labeling and Mass Spectrometry To Study Protein Binding Sites of Amyloid Inhibiting Molecules. *Analytical Chemistry* **2017,** *89* (21), 11583-11591.

56.   Borotto, N. B.; Zhou, Y.; Hollingsworth, S. R.; Hale, J. E.; Graban, E. M.; Vaughan, R. C.; Vachet, R. W., Investigating Therapeutic Protein Structure with

Diethylpyrocarbonate Labeling and Mass Spectrometry. *Analytical Chemistry* **2015,** *87* (20), 10627-10634.

57. Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T. A.; Brusniak, M.-Y.; Paulse, C.; Creasy, D.; Flashner, L.; Kani, K.; Moulding, C.; Seymour, S. L.; Nuwaysir, L. M.; Lefebvre, B.; Kuhlmann, F.; Roark, J.; Rainer, P.; Detlev, S.; Hemenway, T.; Huhmer, A.; Langridge, J.; Connolly, B.; Chadick, T.; Holly, K.; Eckels, J.; Deutsch, E. W.; Moritz, R. L.; Katz, J. E.; Agus, D. B.; MacCoss, M.; Tabb, D. L.; Mallick, P., A cross-platform toolkit for mass spectrometry and proteomics. *Nat Biotech* **2012,** *30* (10), 918-920.

58. Holman, J. D.; Tabb, D. L.; Mallick, P., Employing ProteoWizard to Convert Raw Mass Spectrometry Data. In *Current Protocols in Bioinformatics*, John Wiley & Sons, Inc.: 2002.

59. Vaudel, M.; Barsnes, H.; Berven, F. S.; Sickmann, A.; Martens, L., SearchGUI: An open-source graphical user interface for simultaneous OMSSA and X!Tandem searches. *Proteomics* **2011,** *11* (5), 996-999.

60. The Global Proteome Machine Organization. cRAP protein sequences. 2012 Jan 1 [accessed 2017 Oct 18]. http://www.thegpm.org/crap/index.html.

61. Craig, R.; Beavis, R. C., A method for reducing the time required to match protein sequences with tandem mass spectra. *Rapid Communications in Mass Spectrometry* **2003,** *17* (20), 2310-2316.

62. Craig, R.; Beavis, R. C., TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* **2004,** *20* (9), 1466-1467.

63. Kim, S.; Pevzner, P. A., MS-GF+ makes progress towards a universal database search tool for proteomics. **2014,** *5*, 5277.

64. Geer, L. Y.; Markey, S. P.; Kowalak, J. A.; Wagner, L.; Xu, M.; Maynard, D. M.; Yang, X.; Shi, W.; Bryant, S. H., Open Mass Spectrometry Search Algorithm. *Journal of Proteome Research* **2004,** *3* (5), 958-964.

65. Tabb, D. L.; Fernando, C. G.; Chambers, M. C., MyriMatch: Highly Accurate Tandem Mass Spectral Peptide Identification by Multivariate Hypergeometric Analysis. *Journal of Proteome Research* **2007,** *6* (2), 654-661.

66. Vaudel, M.; Burkhart, J. M.; Zahedi, R. P.; Oveland, E.; Berven, F. S.; Sickmann, A.; Martens, L.; Barsnes, H., PeptideShaker enables reanalysis of MS-derived proteomics data sets. *Nat Biotech* **2015,** *33* (1), 22-24.

67. Katajamaa, M.; Miettinen, J.; Orešič, M., MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data. *Bioinformatics* **2006,** *22* (5), 634-636.

68. Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M., MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **2010,** *11* (1), 395.

69. Myers, J. K.; Pace, C. N.; Scholtz, J. M., Helix Propensities Are Identical in Proteins and Peptides. *Biochemistry* **1997,** *36* (36), 10923-10929.

70. Zhang, H.; Song, L.; Ye, H.; Hu, L.; Liang, W.; Liu, D., Characterization of a Novel Humanized Anti-CD20 Antibody with Potent Anti-Tumor Activity against Non-Hodgkin's Lymphoma. *Cellular Physiology and Biochemistry* **2013,** *32* (3), 645-654.

71.   Cragg, M. S.; Bayne, M. B.; Tutt, A. L.; French, R. R.; Beers, S.; Glennie, M. J.; Illidge, T. M., A new anti-idiotype antibody capable of binding rituximab on the surface of lymphoma cells. *Blood* **2004,** *104* (8), 2540.

72.   Hampson, G.; Ward, T. H.; Cummings, J.; Bayne, M.; Tutt, A. L.; Cragg, M. S.; Dive, C.; Illidge, T. M., Validation of an ELISA for the determination of rituximab pharmacokinetics in clinical trials subjects. *Journal of Immunological Methods* **2010,** *360* (1), 30-38.

73.   Miller, J. N.; Miller, J. C., *Statistics and chemometrics for analytical chemistry*. 6th ed.; Pearson/Prentice Hall: Harlow, England, United Kingdom, 2005.

74.   Shrivastava, A.; Gupta, V., Methods for the determination of limit of detection and limit of quantitation of the analytical methods. *Chronicles of Young Scientists* **2011,** *2* (1), 21-25.

75.   Padlan, E. A., Anatomy of the antibody molecule. *Molecular Immunology* **1994,** *31* (3), 169-217.

76.   Martz, E. Atomic Coordinate Files for Whole IgG1; 1996 Feb 7 [accessed 2018 Aug 4]. https://www.umass.edu/microbio/rasmol/padlan.htm.

77.   Gazzano-Santoro, H.; Ralph, P.; Ryskamp, T. C.; Chen, A. B.; Mukku, V. R., A non-radioactive complement-dependent cytotoxicity assay for anti-CD20 monoclonal antibody. *Journal of Immunological Methods* **1997,** *202* (2), 163-171.

78.   Andersen Christian, B.; Manno, M.; Rischel, C.; Thórólfsson, M.; Martorana, V., Aggregation of a multidomain protein: A coagulation mechanism governs aggregation of a model IgG1 antibody under weak thermal stress. *Protein Science* **2009,** *19* (2), 279-290.

79.   Nejadnik, M. R.; Randolph, T. W.; Volkin, D. B.; Schöneich, C.; Carpenter, J. F.; Crommelin, D. J. A.; Jiskoot, W., Postproduction Handling and Administration of Protein Pharmaceuticals and Potential Instability Issues. *Journal of Pharmaceutical Sciences* **2018,** *107* (8), 2013-2019.

80.   Vermeer, A. W.; Norde, W., The thermal stability of immunoglobulin: unfolding and aggregation of a multi-domain protein. *Biophysical Journal* **2000,** *78* (1), 394-404.

81.   Du, J.; Wang, H.; Zhong, C.; Peng, B.; Zhang, M.; Li, B.; Huo, S.; Guo, Y.; Ding, J., Structural Basis for Recognition of CD20 by Therapeutic Antibody Rituximab. *Journal of Biological Chemistry* **2007,** *282* (20), 15073-15080.

82.   Jhan, S.-Y.; Huang, L.-J.; Wang, T.-F.; Chou, H.-H.; Chen, S.-H., Dimethyl Labeling Coupled with Mass Spectrometry for Topographical Characterization of Primary Amines on Monoclonal Antibodies. *Analytical Chemistry* **2017,** *89* (7), 4255-4263.

83.   Courtois, F.; Schneider, C. P.; Agrawal, N. J.; Trout, B. L., Rational Design of Biobetters with Enhanced Stability. *Journal of Pharmaceutical Sciences* **2015,** *104* (8), 2433-2440.

84.   Wang, X.; Das, T. K.; Singh, S. K.; Kumar, S., Potential aggregation prone regions in biotherapeutics: A survey of commercial monoclonal antibodies. *mAbs* **2009,** *1* (3), 254-267.

85.   van der Kant, R.; Karow-Zwick, A. R.; Van Durme, J.; Blech, M.; Gallardo, R.; Seeliger, D.; Aßfalg, K.; Baatsen, P.; Compernolle, G.; Gils, A.; Studts, J. M.; Schulz, P.; Garidel, P.; Schymkowitz, J.; Rousseau, F., Prediction and Reduction of the

Aggregation of Monoclonal Antibodies. *Journal of Molecular Biology* **2017,** *429* (8), 1244-1261.

86.   Watson, C.; Sharp, J. S., Conformational analysis of therapeutic proteins by hydroxyl radical protein footprinting. *AAPS J.* **2012,** *14* (2), 206-217.

87.   Jones, L. M.; B. Sperry, J.; A. Carroll, J.; Gross, M. L., Fast Photochemical Oxidation of Proteins for Epitope Mapping. *Analytical Chemistry* **2011,** *83* (20), 7657-7661.

CHAPTER 4

COVALENT LABELING/MASS SPECTROMETRY OF MONOCLONAL ANTIBODIES WITH DIETHYLPYROCARBONATE: REACTION KINETICS FOR ENSURING STRUCTURAL INTEGRITY

4.1 Introduction

Mass spectrometry (MS) has emerged as one of the leading techniques used for the routine structural analysis of antibody therapeutics to ensure the quality, efficacy and safety throughout a product life cycle.[1-4] Higher-order structure (HOS) is one of the quality attributes of therapeutic monoclonal antibody (mAb) that has to be characterized as changes in HOS can cause decrease in therapeutic efficacy, reduced stability, or possible immunogenicity.[5-7] In addition, identifying specific residues involved in antigen-antibody interactions, i.e. epitope/paratope mapping, is particularly useful to evaluate epitope novelty, predict immunogenicity, and assess and optimize binding characteristics.[8, 9] The wider use of MS-based methods to probe protein's HOS results from (a) an ability to obtain residue-level structural information that optical spectroscopy cannot provide, and (b) a more rapid, sensitive, and sample-efficient analysis compared to X-ray crystallography and NMR spectroscopy.

The two most commonly used MS-based tools for studying mAb HOS are hydrogen-deuterium exchange (HDX) and covalent labeling (CL). HDX has been extensively used to investigate HOS changes and aggregation of mAbs during drug development.[10-14] Epitope mapping and comparability studies are also examples of how HDX-MS has been used.[14-18] CL of amino acid side chains is another technique that can be used to encode a protein's structural properties into the mass of protein, allowing

information about a protein's structure in solution to be read-out through MS-based bottom-up approaches.[19-22] CL-MS has more recently been applied to study mAb higher order structure. Compared to hydrogen-deuterium exchange (HDX), CL benefits from limited label loss and scrambling due to the irreversible nature of the labeling reaction. Non-specific reagents (e.g., hydroxyl radicals, carbenes, and diethylpyrocarbonate) can simultaneously probe a range of different side chains, allowing excellent structural resolution to be obtained. Hydroxyl radical footprinting (HRF) has been used to characterize the conformations and dimer interfaces of mAb therapeutics and to perform epitope mapping for various antigen-antibody interactions.[23-27] Carboxyl group footprinting has also been used to characterize the HOS of mAbs and was found to provide complementary information to HRF.[28]

Our group has developed and used diethylpyrocarbonate (DEPC) as a labeling reagent because it is simple to use, commercially available, and labels a range of nucleophilic residues and the N-terminus of proteins once added into a solution. DEPC CL-MS has been used to reveal aggregation sites and HOS changes for heat-stressed murine IgG1 and rituximab, providing structural information for up to 30% of the residues in these mAbs.[29, 30] An important principle of CL-MS experimental design is that a protein's structure must not be perturbed during the labeling reaction, so that the probe can correctly report accurate structural information.[19, 31] There are at least three general ways to assess a protein's structural integrity upon CL. One way is to use a complementary measurement, such as CD spectroscopy, fluorescence spectroscopy, or an activity assay, to monitor any HOS change after labeling.[30-32] These techniques are typically not sensitive enough to detect local structural perturbations. A second way is to ensure that the labeling

138

reaction occurs faster than any label-induced HOS change can happen, guaranteeing a protein's HOS is not affected during CL. This approach only works for fast labeling chemistries, e.g. HRF and carbene labeling, that require laser or synchrotron sources.[33-36]

A third method, which is more reliable and effective, is to monitor labeling reaction kinetics. By plotting the unmodified fraction of a protein (or better, each of its peptide fragments) as function of reagent concentration, any labeling-induced structural perturbations can be revealed by deviations from the proper reaction order kinetics, e.g. pseudo first-order kinetics for HRF and second-order kinetics for DEPC CL.[31, 37-39] Such dose-response plots of an intact protein are less useful than plots of peptide fragments as intact protein plots are ensemble averages, much like CD measurements, and do not provide site-specific information. Dose-response plots for proteolytic fragments obtained from bottom-up analysis of a labeled protein have been shown to be much more sensitive and reliable probes of possible protein structural perturbations as they provide information about any local structural perturbations caused by either nearby or distant modifications.[31] Deviations in the kinetics allow the researcher to identify the highest reagent dose to use while still ensuring the structural integrity of the protein. Using the highest dose, while preventing structural perturbations, maximizes the labeling extent and thus the structural resolution of the method. Kinetic measurements via dose-response plots have been effective for ensuring the reliability of the structural information for proteins ranging from cytochrome c to lysozyme to gelsolin.[37-39] These measurements, however, have only been used for, single-domain proteins, and so an interesting question is whether they are beneficial for multi-domain proteins like mAbs. Moreover, it would be valuable to know

if CL on a residue in one domain of a protein (e.g. variable light chain in a mAb) can affect the structure and labeling of a distant domain (e.g. constant heavy chain in a mAb).

In this work, we have acquired DEPC labeling dose-response plots for proteolytic fragments of different mAbs to: (a) test the applicability of this method for ensuring the structural integrity of multi-domain mAbs during CL-MS-based structural analyses, (b) identify if a common DEPC molar excess might be suitable for DEPC-based CL-MS analyses of mAbs, and (c) evaluate the extent to which a multi-domain protein's structure is sensitive to multiple DEPC modifications. Results suggest that multi-domain proteins can withstand many more than 1 DEPC label without being perturbed structurally, which allows a more sensitive and higher resolution measure of structure because more modifications can be detected. In addition, based on the limited set of mAbs studied here, there seems to be a common DEPC molar ratio that can be used without worrying about structural perturbations.

4.2 Experimental Section

4.2.1 Materials

Rituximab (Rituxan® 100 mg/10 mL vial, lot#3209283, Genentech) and the NIST Monoclonal Antibody Reference Material 8671 (NISTmAb 10 mg/mL vial, lot#14HB-D-002) were ordered from Myoderm (Norristown, PA) and the National Institute of Standards and Technology (NIST, Gaithersburg, MD), respectively. Diethylpyrocarbonate (DEPC) (#D5758), imidazole (#I5513), iodoacetamide (#I6125), tris(2-carboxyethyl)phosphine (TCEP) (#C4706), and trypsin (#T1426) were all purchased from Sigma-Aldrich (St. Louis, MO). Sodium phosphate monobasic monohydrate (#S0710) was ordered from EM

Science (Darmstadt, Hesse, Germany). Sodium phosphate dibasic anhydrous (#S374), LC/MS-grade formic acid (#A117), acetonitrile (#A998), and water (#W7) were obtained from Fisher Scientific (Fair Lawn, NJ). All reagents used in this study have no known potential hazards.

4.2.2 Sample Preparation and DEPC Labeling Reactions

Rituximab was used as is in its formulation, which contains 10 mg/mL rituximab, 0.7 mg/mL polysorbate 80, 7.35 mg/mL sodium citrate dihydrate and 9 mg/mL sodium chloride in water at pH 6.5. For the NISTmAb formulation, L-histidine was removed from the formulation using PD SpinTrap G-25 spin columns (#28918004, GE Healthcare Life Sciences, Marlborough, MA) to prevent this amino acid from interfering with the DEPC reaction. The NISTmAb was then buffer exchanged in 50 mM phosphate buffer at pH 6.0, which is the pH of the formulation before L-histidine removal. Aliquots of rituximab and the NISTmAb (10 mg/mL, 70 μM) with only minor dilution (to 60 μM) were reacted with DEPC. Stock solutions of DEPC (69 mM) were freshly prepared in acetonitrile, and the final solution of DEPC was then prepared in water. Labeling of each mAb was performed for 5 min at 37 °C and was initiated by adding various molar excesses of DEPC, from 2-fold up to 30-fold). The reaction was quenched by the addition of imidazole at a 1:50 DEPC to imidazole molar ratio. For experiments at each DEPC concentration, at least three replicates were performed on the rituximab and NISTmAb samples.

4.2.3 Proteolytic Digestion

After quenching, the labeled mAb samples were diluted in 50 mM phosphate buffer at pH 7.4 and added into a urea-containing tube. The resulting mixture contained 8 M urea

for protein denaturation at room temperature. TCEP was added at the final concentration of 25 mM to reduce the disulfide bonds, and iodoacetamide was simultaneously added at the same final concentration to alkylate the reduced Cys residues. The samples were then kept in the dark at room temperature for 20 min. For rituximab samples, DetergentOUT™ Tween® Micro spin columns (#786–214, G-Biosciences, St. Louis, MO) were used to remove Tween® 80 from the samples. Subsequently, overnight digestion at 37 °C was performed with trypsin at a 1:10 (w/w) enzyme to substrate ratio. Following digestion, trypsin was removed from the resulting rituximab or NISTmAb peptides through ultrafiltration using an Amicon® centrifugal filter with a 10 kDa molecular weight cutoff (#UFC501096, Millipore, Burlington, MA). The flow-through was collected, flash-frozen in liquid nitrogen, and stored at −20 °C until LC-MS/MS analysis.

4.2.4 HPLC Separation

Online LC-MS/MS analyses were performed on all rituximab and NISTmAb digests. A sample containing approximately 5 μg of the digested protein was loaded on a Thermo Scientific Dionex Ultimate 3000 HPLC system (Waltham, MA). The separation of peptides was performed on a Thermo Scientific Acclaim™ PepMap™ RSLC C18 column (15 cm x 300 μm, 2 μm particle size, 100 Å pore size) with a flow rate of 4 μL/min. LC/MS-grade water (solvent A) and acetonitrile (solvent B), each containing 0.1% formic acid, were used as mobile phases. Desalting was performed at 5% B during the first 5 min after sample injection. A linear gradient of solvent B was increased from 5% B to 35% B over 65 min. The gradient was finally elevated to and held at 95% B to flush the column.

4.2.5 Mass Spectrometry

Mass spectra were acquired on a Thermo Scientific Orbitrap Fusion mass spectrometer (Waltham, MA). The electrospray needle voltage was kept at ~4 kV (positive mode), and the ion transfer tube temperature was set to 275 °C. Tandem mass spectrometry (MS/MS) was conducted on a set of selected mAb peptides. The precursor ions (unmodified or modified peptide ions) were selected using a quadrupole mass filter at an isolation width of 2.0. The AGC target and maximum injection time were set to 5×10^4 ions and 50 msec, respectively. Collision induced dissociation (CID) was performed in a linear quadrupole ion trap with a normalized collision energy of 35%. Mass spectra of product ions were acquired on an Orbitrap analyzer with a resolution of 30,000.

4.2.6 Dose-Response Plots

DEPC can modify His, Lys, Ser, Thr, Tyr, and N-termini, and the labeling results in a single type of modification product with a mass addition of 72.02 Da (**Figure 4.1a**). Under the conditions used in this study, the reaction of DEPC with a specific site in the protein follows second order kinetics.[31]

Nucleophilic side chain (P) + DEPC (X) → Carbethoxylated P + CO_2 + CH_3CH_2OH
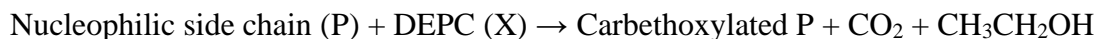
If the reaction follows second order kinetics, the rate is defined by:

$$Rate \ = \ -\frac{d[P]}{dt} = k[P][X] \tag{4.1}$$

where [P] is the concentration of unmodified mAb at time t, [X] is the DEPC concentration at time t, and k is the second-order rate coefficient.

Let $\quad a = reaction\ progress = [P]_0 - [P] = [X]_0 - [X]$ (4.2)

where $[P]_0$ is the initial concentration of unmodified mAb, $[X]_0$ is the initial concentration

of DEPC

$$-\frac{d[P]}{dt} = \frac{da}{dt} = k([P]_0 - a)([X]_0 - a)$$ (4.3)

$$\frac{da}{([P]_0 - a)([X]_0 - a)} = k\ dt$$ (4.4)

$$\int_0^a \frac{da}{([P]_0 - a)([X]_0 - a)} = k\int_0^t dt$$ (4.5)

$$\frac{1}{[X]_0 - [P]_0}\left[\ln\frac{[P]_0}{[P]_0 - a} - \ln\frac{[X]_0}{[X]_0 - a}\right] = kt$$ (4.6)

$$\frac{1}{[X]_0 - [P]_0}\left[\ln\frac{[P]_0[X]}{[P][X]_0}\right] = kt$$ (4.7)

$$\ln\left(\frac{[X]_0[P]}{[X][P]_0}\right) = -kt[X]_0 + kt[P]_0$$ (4.8)

For each specific labeling site, a plot between $\ln\left(\frac{[X]_0[P]}{[X][P]_0}\right)$ and $[X]_0$ was produced for

a given peptide from LC-MS/MS data of that peptide. $[P]/[P]_0$ is the ratio of the peak area

for the unmodified peptide to the sum of the peak areas for the modified and unmodified

peptide, and $[X]$ is determined by the difference between the $[P]$ and $[P]_0$ values. A

hypothetical dose-response plot for a second order reaction is shown in **Figure 4.1b**.

A custom software pipeline specifically designed for protein CL-MS studies with

DEPC[29, 30] was used for the initial identification of the peptides that were chosen for

MS/MS. Assignments of b and y ions from CID tandem mass spectra were performed with

a mass tolerance of 0.5 Da. From the LC-MS/MS analyses on a set of selected mAb

peptides, Thermo Scientific Xcalibur™ software was used to reconstruct product ion

chromatograms of unmodified and modified peptides. Peptide identification and peak quantification were performed using tandem spectra and mass spectral peak areas, respectively.
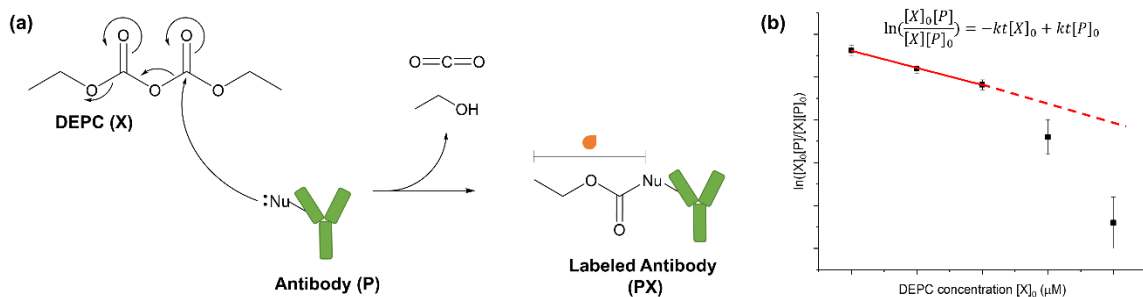
Figure 4.1: The reaction of DEPC with a specific site in protein
(a) DEPC labeling of nucleophilic side chains (His, Lys, Tyr, Ser, and Thr) in the antibodies. (b) A hypothetical dose-response plot for a given peptide fragment can be used to identify any labeling-induced structural perturbation to the protein in the region represented by the measured peptide. In this example, a break in linearity is indicative of a structural perturbation at a high DEPC concentration.

4.2.7 LC-MS Analysis of DEPC-Labeled Light and Heavy Chains of the mAbs

Aliquots of rituximab and the NISTmAb (10 mg/mL, 70 μM) were reacted with DEPC. Information about sample preparation and DEPC labeling reaction conditions can be found in **Section 4.2.2**. After labeling and quenching, TCEP was added to the DEPC-labeled samples at a 500:1 TCEP:mAb molar ratio to reduce the disulfide bonds of the mAb and yield heavy and light chains. The resulting mixtures were diluted in 50 mM phosphate buffer at pH 7.4 to a final concentration of 0.15 mg/mL.

Online LC-MS analyses were performed on the TCEP-reduced samples. A sample containing approximately 2 μg protein was loaded on a Thermo Scientific Ultimate 3000 HPLC system (Waltham, MA). The separation was performed on a Waters Acquity UPLC Protein BEH C4 column (50 mm x 2.1 mm, 1.7 μm particle size, 300 Å pore size; Milford, MA) with a flow rate of 200 μL/min. LC/MS-grade water (solvent A) and acetonitrile

145

(solvent B), each containing 0.1% formic acid, were used as mobile phases. Desalting was performed at 5% B during the first 4 min after sample injection. A linear gradient of solvent B was increased from 5% B to 100% B over 12 min. The gradient was held at 95% B for additional 3 min to flush a column.

Mass spectra were acquired on a Thermo Scientific Orbitrap Fusion mass spectrometer (Waltham, MA). The electrospray needle voltage was kept at ~4 kV (positive mode), and the ion transfer tube temperature was set to 330 °C. In-source activation energy was applied at 35 V to help remove water and other adducts from protein ions. Mass spectra were acquired on an Orbitrap analyzer, with a resolution of 15,000. The AGC target and maximum injection time were set to 1×10^6 ions and 100 msec, respectively. Measurements were conducted in the high mass range mode with 3 microscans per spectrum.

4.3 Results and Discussion

4.3.1 Dose-Response Plots as Indicators of Antibody Structural Changes upon DEPC Labeling

Reliable information about the HOS of mAbs can be obtained from CL-MS experiments only when the protein's structural integrity is preserved during the labeling. Under the conditions that DEPC is normally used for labeling, its reaction with a specific site in a protein follows second order kinetics.[31] Deviations from these kinetics at high reagent concentrations can indicate a structural perturbation, as the labeling kinetics are sensitive to any structural changes that affect the reactivity of a given residue. We have shown in previous work with DEPC that breaks in linearity in dose-response plots are

146

indicators of labeling-induced structural perturbations in small proteins, such as cytochrome c and myoglobin.[31] In addition, other investigators have used dose-response plots to measure structural changes during HRF and carboxyl group footprinting.[28, 38] Because DEPC can hydrolyze in water over time, the DEPC reactions are conducted at a constant reaction time t, and a rate constant can be determined from a dose-response plot. Changes in local structural features like solvent accessibility and microenvironment[19, 31, 40] that can occur at higher DEPC concentrations lead to breaks in the linearity of the plot, indicating these structural perturbations (**Figure 4.1b**). Dose-response plots for individual proteolytic fragments of the labeled protein provide a sensitive measure of structural perturbations as they report on local changes. The experimental workflow to generate dose-response plots for proteolytic fragments is shown in **Figure 4.2**.
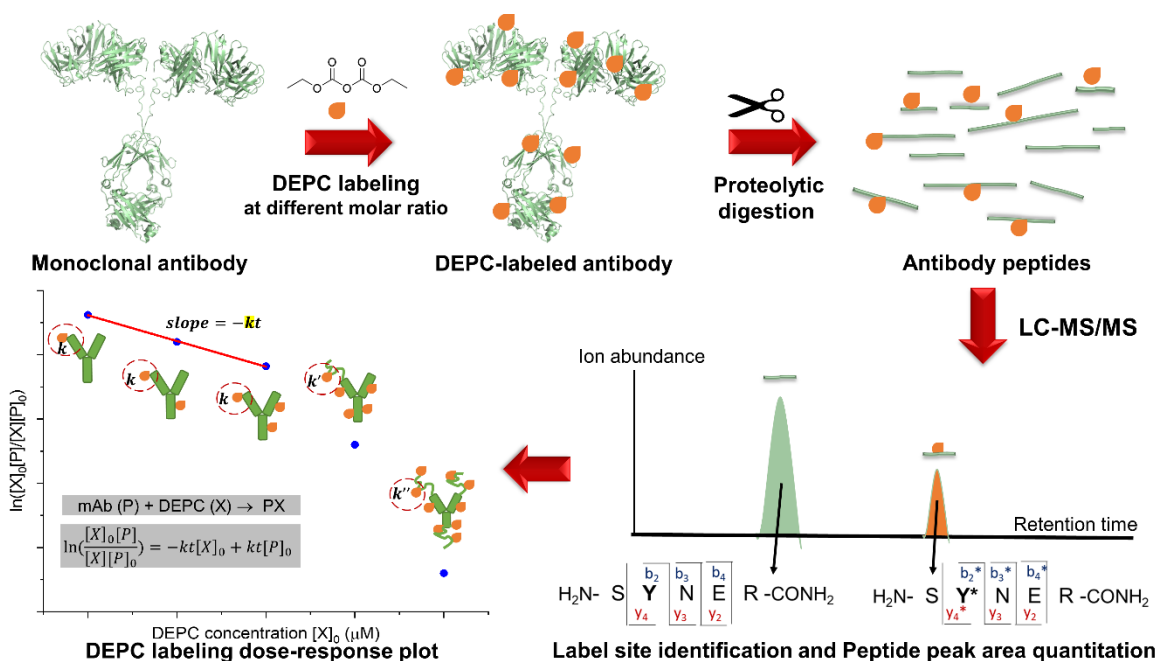
Figure 4.2: The experimental workflow used to generate dose-response plots for selected proteolytic fragments of the mAbs studied here.
The mAb reacts with DEPC at different reagent concentrations. The labeled mAb is subjected to proteolytic digestion, and the labeled peptides are analyzed using LC–MS/MS. Peak areas of unmodified and modified species are used to generate a dose-response plot.

147

The antibodies rituximab and NISTmAb were reacted with a range of DEPC concentrations, and dose-response plots were generated for a set of representative peptides with labeled sites spread throughout the antibody structure (**Figure 4.3**). One to two peptides per domain were selected, and these peptides (a) had no post-translational modifications, and (b) were generated with no missed cleavages during the tryptic digestion. Several of the peptides that are reported here include residues in the complementarity-determining regions (CDRs), which makes them critically important for the antibody's structure and function. Examples of tandem mass spectra and label site identifications of these representative peptides can be found in **Figures 4.4 and 4.5**.
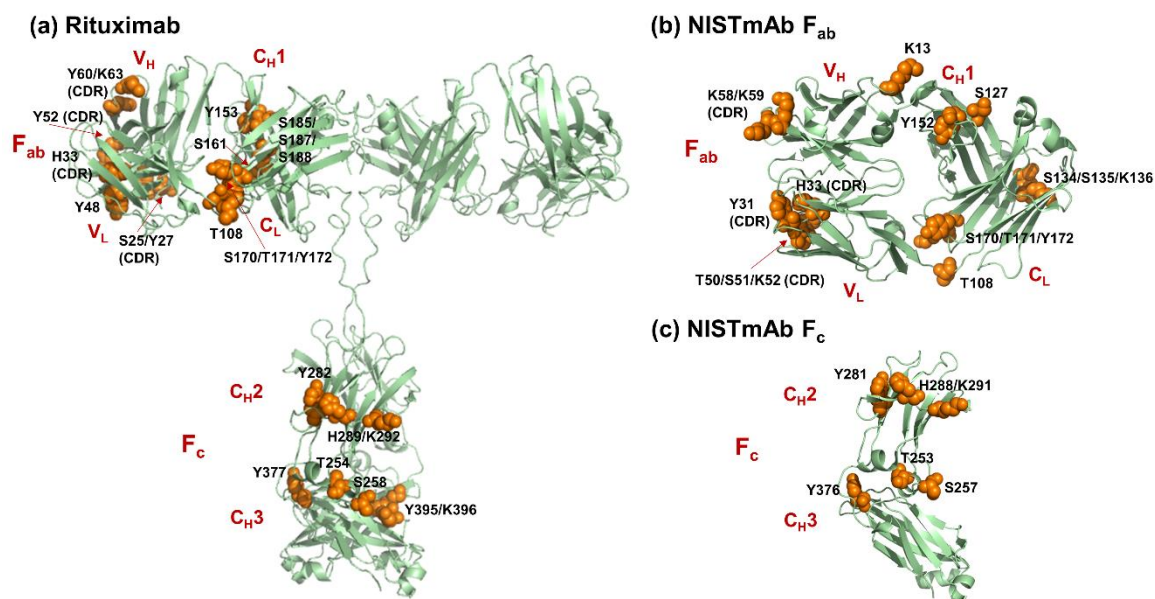
Figure 4.3: Labeling sites (orange sphere) from a set of representative peptides that were selected from different domains throughout the antibody structures.
(a) Rituximab - F_{ab} (PDB 4KAQ) and F_c (PDB 4W4N) structures of rituximab are aligned to the human IgG1 model, using the molecular visualization system PyMOL.[29,41] Note that, for clarity, only one asymmetric unit of rituximab structure is labeled in this figure. (b) NISTmAb F_{ab} (PDB 5K8A) and (c) NISTmAb F_c (PDB 4W4N)

(a) HC_Y52
CDR-H2 (44-63, V$_H$)
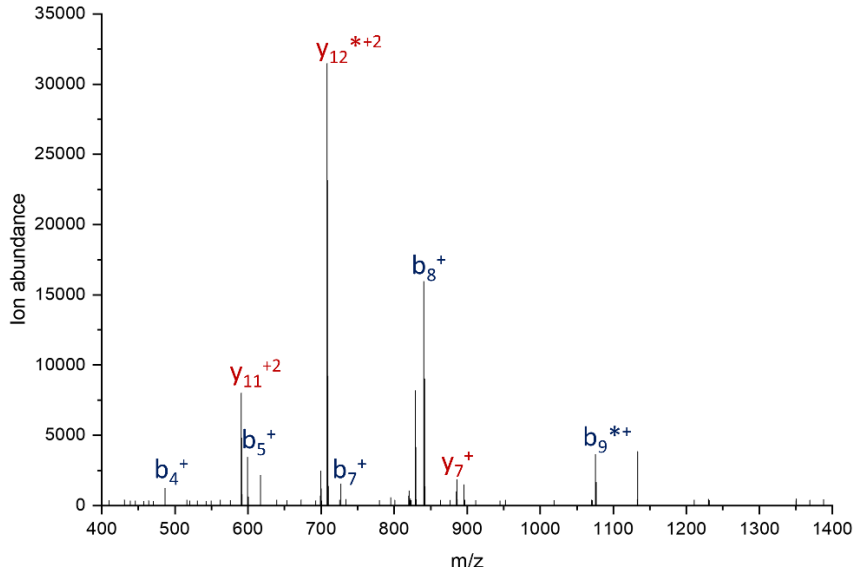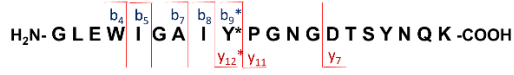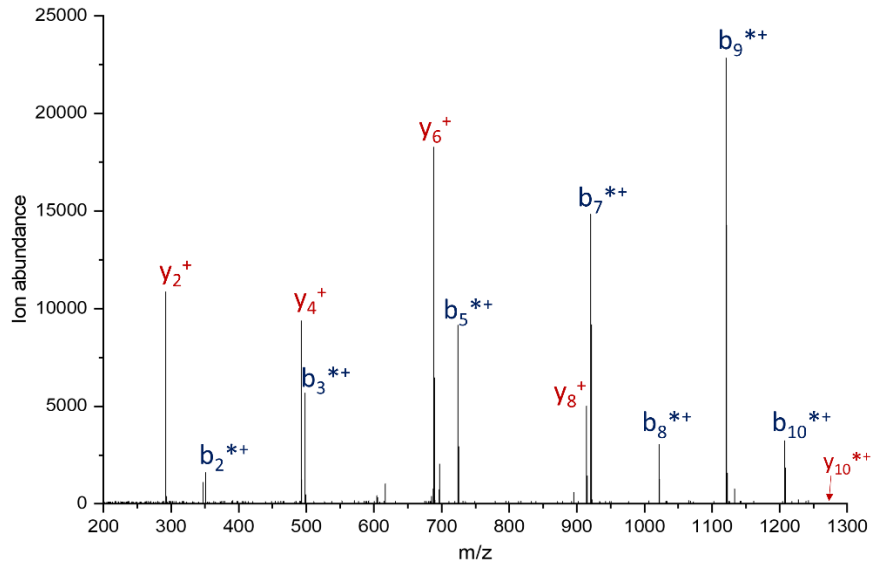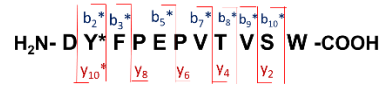
H$_2$N- G L E W I G A I Y* P G N G D T S Y N Q K -COOH

(b) HC_Y153
(152-162, C$_H$1)

H$_2$N- D Y* F P E P V T V S W -COOH

**(c) HC_T254
(253-259, C$_H$2)**

H$_2$N- D T* L M I S R -COOH

**(d) HC_Y377
(375-396, C$_H$3)**

H$_2$N- G F Y* P S D I A V E W E S N G Q P E N N Y K-COOH

150

(e) LC_H33
CDR-L1 (25-34, V$_L$)

H$_2$N- A S S S V S Y I H* W -COOH

(f) LC_T108
(108-115, C$_L$)

H$_2$N- T* V A A P S V F-COOH
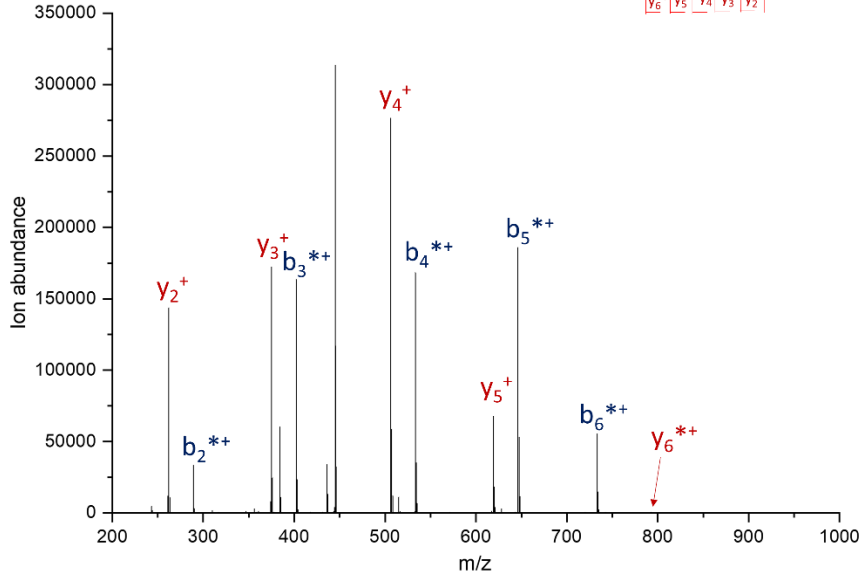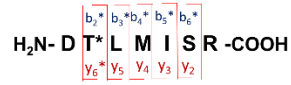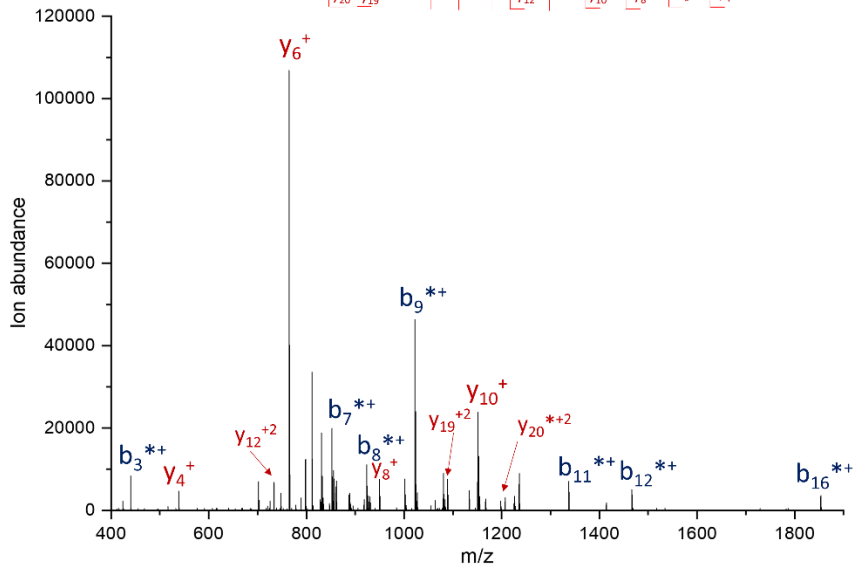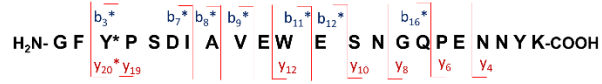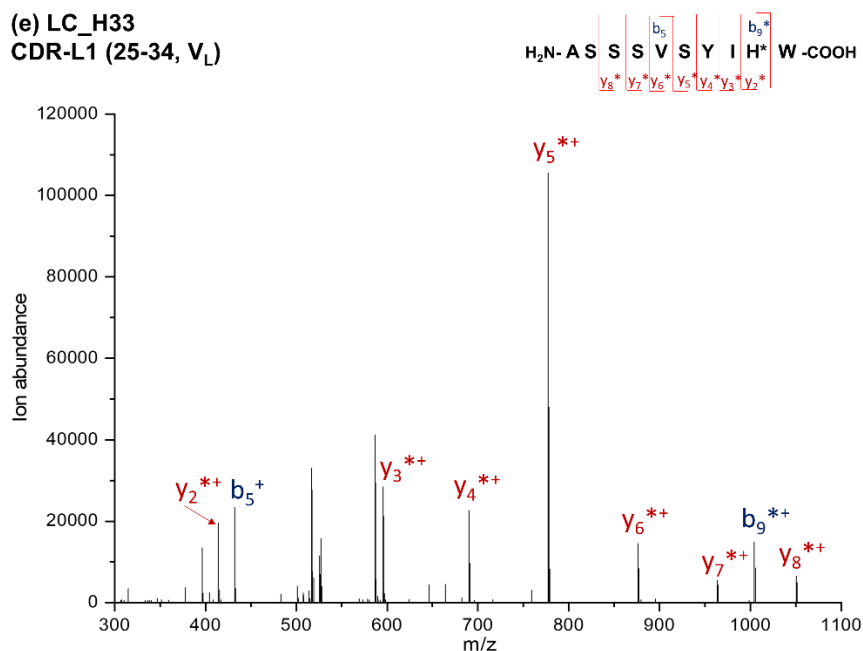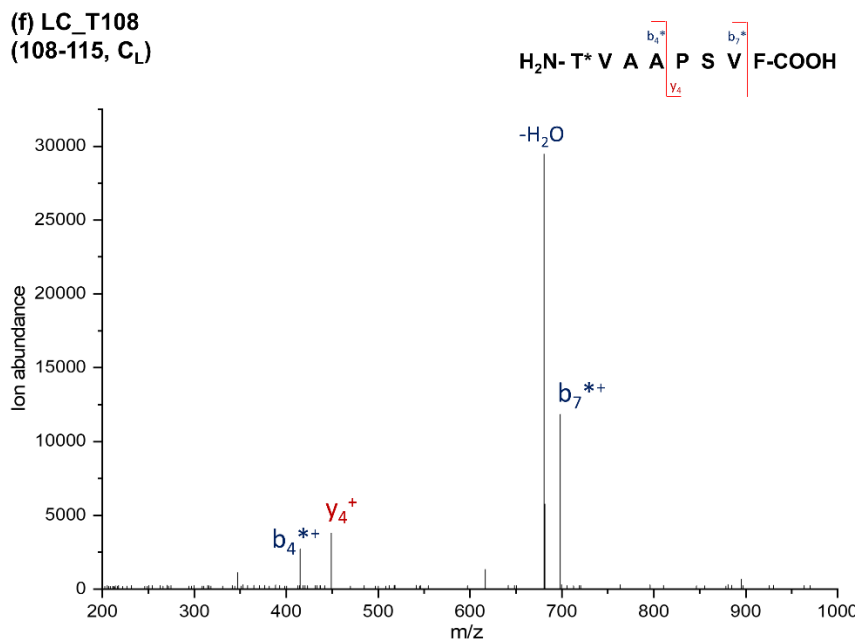
Figure 4.4: MS/MS assignments for DEPC-labeled peptides from each of the six antibody domains in heavy chain (HC) and light chain (LC) of rituximab, (a) V$_H$, (b) C$_H$1, (c) C$_H$2, (d) C$_H$3, (e) V$_L$, and (f) C$_L$

(a) HC_K13 (6-18, V_H)

H_2N- E S G P A L V K* P T Q T L -COOH

(b) HC_Y152 (151-161, C_H1)

H_2N- D Y* F P E P V T V S W -COOH

(c) HC_T253
(252-258, C_H2)

H_2N- D T* L M I S R -COOH
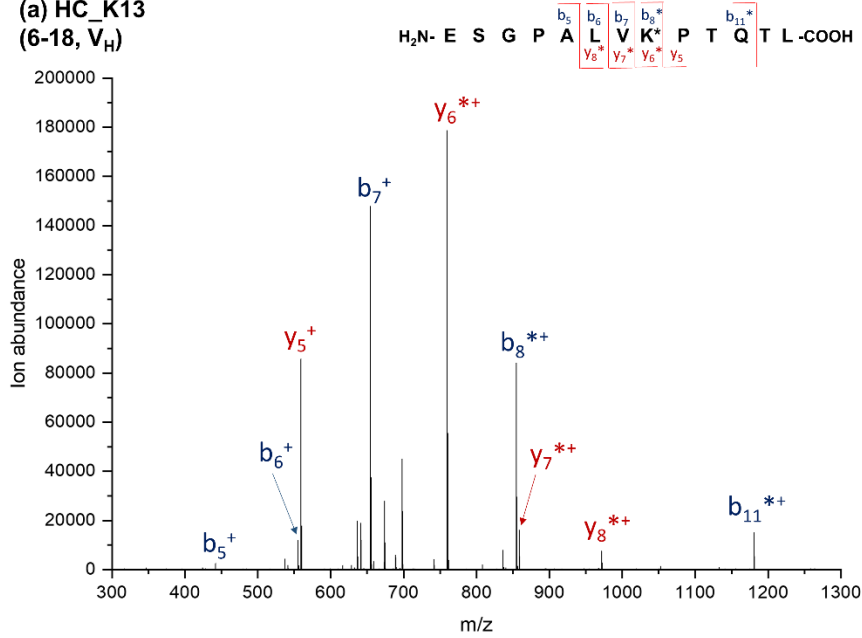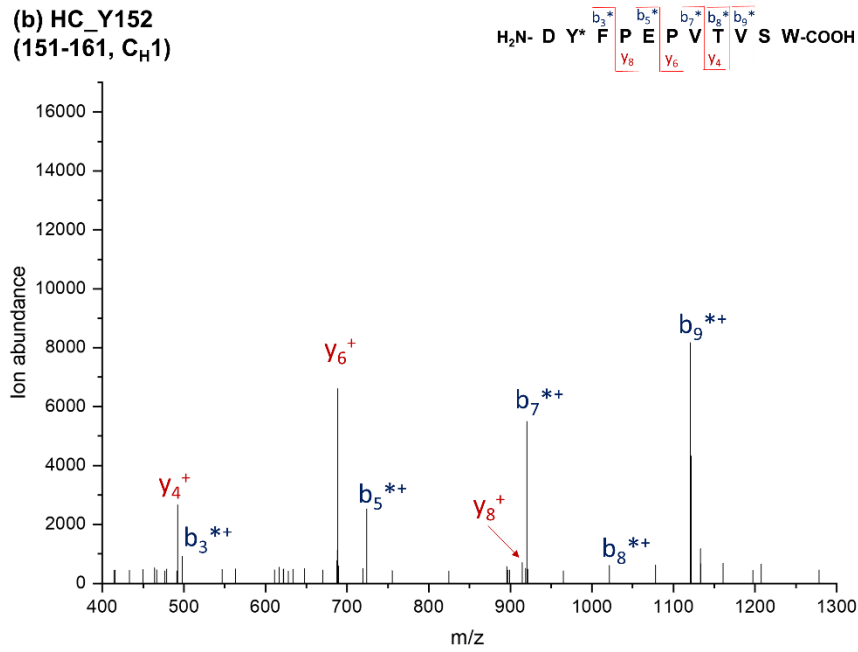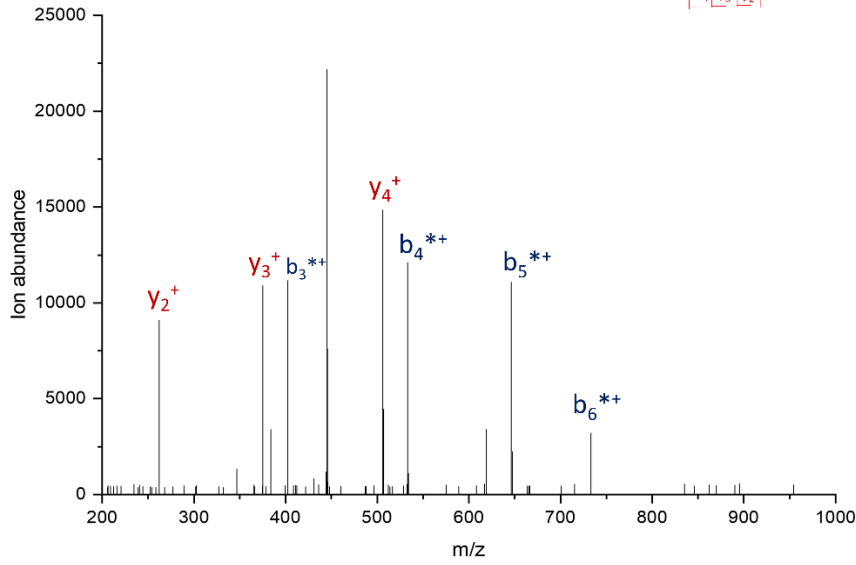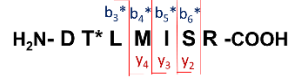
(d) HC_Y376
(376-395, C_H3)

H_2N- G F Y* P S D I A V E W E S N G Q P E N N Y K-COOH

153
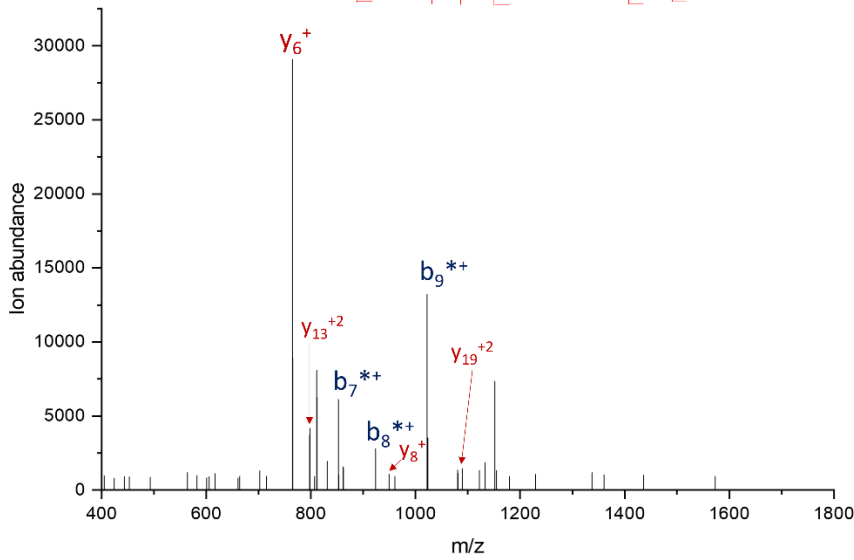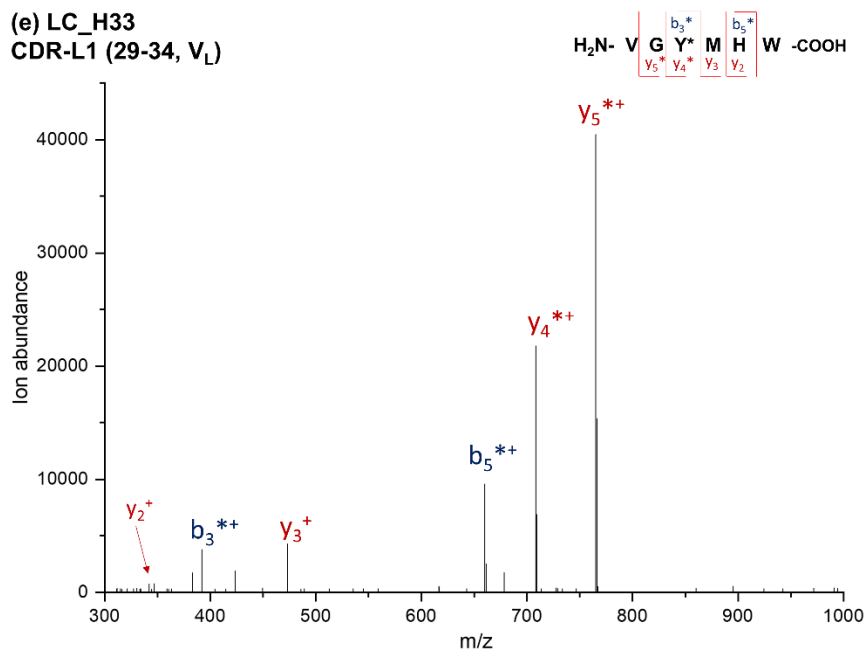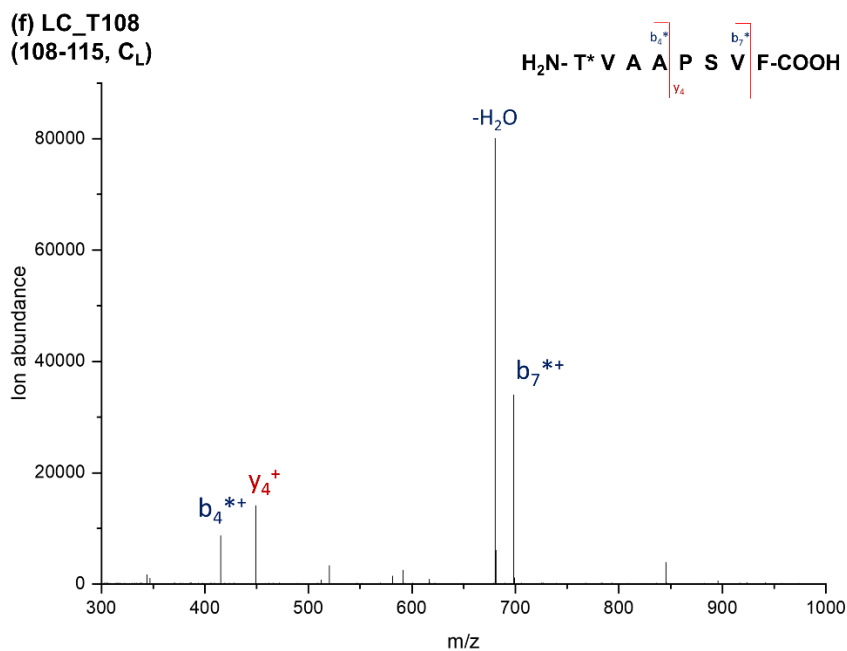
Figure 4.5: MS/MS assignments for DEPC-labeled peptides from each of the six antibody domains in heavy chain (HC) and light chain (LC) of NISTmAb, (a) V_H, (b) C_H1, (c) C_H2, (d) C_H3, (e) V_L, and (f) C_L

4.3.2 Dose-response plots can be used to reveal labeling-induced structural perturbations for antibodies, which are multi-domain proteins

Dose-response plots for selected proteolytic fragments of rituximab are shown in **Figure 4.6**. For all of the peptides, linear relationships are observed at low DEPC concentrations, suggesting that the protein's HOS in different domains is maintained during the DEPC labeling at these concentrations. Breaks in linearity of the dose-response plots were determined from R^2 values and standard errors of regression ($s_{y/x}$).[42, 43] Specifically, R^2 values above 0.95 and $s_{y/x}$ below 30% were used as cutoffs for assessing linearity. The 30% value for $s_{y/x}$ was chosen because we found that the average %RSD for all the data for rituximab and NISTmAb is around 30%. One conclusion from the rituximab results is that the linearity of the dose-response plots is maintained for at least up to 6-fold DEPC in all cases, and for some residues it is maintained up to 10- or 15-fold DEPC [**Table 4.1** and **Figure 4.6 (j) to (p)**]. This observation means that rituximab's HOS can be reliably maintained at DEPC:protein molar ratios of 6-fold. An almost identical conclusion is made from the dose-response plots for proteolytic fragments of NISTmAb (**Figure 4.7**). The dose-response plots maintain linearity up to 6-fold for all peptides, while some peptides retain linearity at DEPC:protein ratios as high as 15-fold [**Table 4.1** and **Figure 4.7 (f) to (o)**]. The fact that deviations in linearity occur at higher than 6-fold DEPC for both antibodies suggests that this might be a DEPC:protein ratio that could be commonly used for mAbs without worry of structural perturbations. Indeed, because therapeutic mAbs have the same overall fold and tend to have identical F_c domains, the results obtained here are probably transferrable to all mAbs; however, more studies would likely be needed to more thoroughly evaluate this observation.

Table 4.1: A summary of the dose-response data for residues in rituximab and the NISTmAb.

Antibody Residue[a]	Secondary structure	Linearity up to	Rate coefficient k[b] (M^{-1} s^{-1})
Rituximab			
LC S170/T171/Y172[c]	β-sheet & random coil	6X	$7.4 \pm 0.8 \times 10^{-1}$
HC H289/K292[c]	Random coil	10X	$3.1 \pm 0.2 \times 10^{-1}$
HC Y282	β-sheet	6X	$1.6 \pm 0.1 \times 10^{-1}$
LC H33	β-sheet	6X	$1.0 \pm 0.1 \times 10^{-1}$
HC S25/Y27[c]	Random coil	10X	$7.2 \pm 0.6 \times 10^{-2}$
LC Y48	Random coil	15X	$4.9 \pm 0.3 \times 10^{-2}$
HC Y395/K396[c]	β-sheet	10X	$2.2 \pm 0.1 \times 10^{-2}$
HC T254	β-sheet	6X	$2.1 \pm 0.4 \times 10^{-2}$
HC Y377	Random coil	6X	$1.34 \pm 0.02 \times 10^{-2}$
HC S258	Random coil	6X	$1.31 \pm 0.06 \times 10^{-2}$
HC Y153	Random coil	6X	$8.5 \pm 0.1 \times 10^{-3}$
HC Y52	β-sheet	6X	$3.6 \pm 0.2 \times 10^{-3}$
HC Y60/K63[c]	Random coil	6X	$3.5 \pm 0.2 \times 10^{-3}$
HC S185/S187/S188[c]	β-sheet	15X	$2.5 \pm 0.1 \times 10^{-3}$
HC S161	β-sheet	10X	$2.5 \pm 0.2 \times 10^{-3}$
LC T108	Random coil	10X	$2.2 \pm 0.1 \times 10^{-3}$
NISTmAb			
HC K58/K59[c]	β-sheet & random coil	6X	$8.0 \pm 0.2 \times 10^{-1}$
HC Y281	β-sheet	10X	$3.6 \pm 0.1 \times 10^{-1}$
HC H288/K291[c]	Random coil	10X	$3.5 \pm 0.1 \times 10^{-1}$
HC K13	Random coil	8X	$3.2 \pm 0.2 \times 10^{-1}$
LC Y31	Random coil	15X	$4.3 \pm 0.5 \times 10^{-2}$
LC H33	β-sheet	15X	$2.3 \pm 0.2 \times 10^{-2}$
LC T50/S51/K52[c]	β-sheet & random coil	6X	$1.22 \pm 0.04 \times 10^{-2}$
LC S170/T171/Y172[c]	β-sheet & random coil	10X	$1.22 \pm 0.09 \times 10^{-2}$
HC Y376	Random coil	10X	$1.1 \pm 0.1 \times 10^{-2}$
HC T253	β-sheet	6X	$4.7 \pm 0.5 \times 10^{-3}$
HC Y152	β-sheet	10X	$4.5 \pm 0.3 \times 10^{-3}$
HC S127	β-sheet	6X	$3.6 \pm 0.1 \times 10^{-3}$
HC S257	Random coil	6X	$3.1 \pm 0.2 \times 10^{-3}$
HC S134/S135/K136[c]	α-helix	10X	$2.9 \pm 0.2 \times 10^{-3}$
LC T108	Random coil	8X	$2.2 \pm 0.1 \times 10^{-3}$

[a] LC refers to light chain, and HC refers to heavy chain.

[b] The k values are obtained by dividing the measured slope by the reaction time (see **Equation 4.8**). Error bars are calculated from the standard error of the slope.

[c] Tandem mass spectrometry does not enable the precise modification site to be definitively identified.

❑ **Linearity up to 6X**

(a) HC_Y52   CDR-H2 (44-63, V_H)
H$_2$N- GLEWIGAI**Y**PGNGDTSYNQK-COOH
$k = 3.6 \pm 0.2 \times 10^{-3}$

(b) HC_Y60/K63   CDR-H2 (44-63, V_H)
H$_2$N- GLEWIGAIYPGNGDTS**Y**NQ**K**-COOH
$k = 3.5 \pm 0.2 \times 10^{-3}$

(c) HC_Y153   (152-162, C_H1)
H$_2$N- D**Y**FPEPVTVSW-COOH
$k = 8.5 \pm 0.1 \times 10^{-3}$

(d) HC_T254   (253-259, C_H2)
H$_2$N- D**T**LMISR-COOH
$k = 2.1 \pm 0.4 \times 10^{-2}$

(e) HC_S258   (253-259, C_H2)
H$_2$N- DTLMI**S**R-COOH
$k = 1.31 \pm 0.06 \times 10^{-2}$

(f) HC_Y282   (279-292, C_H2)
H$_2$N- FNW**Y**VDGVEVHNAK-COOH
$k = 1.6 \pm 0.1 \times 10^{-1}$

(g) HC_Y377   (375-396, C_H3)
H$_2$N- GF**Y**PSDIAVEWESNGQPENNYK-COOH
$k = 1.34 \pm 0.02 \times 10^{-2}$

(h) LC_H33   CDR-L1 (25-34, V_L)
H$_2$N- ASSSVSYI**H**W-COOH
$k = 1.0 \pm 0.1 \times 10^{-1}$

(i) LC_S170/T171/Y172   (169-180, C_L)
H$_2$N- D**STY**SLSSTL-COOH
$k = 7.4 \pm 0.8 \times 10^{-1}$

157

Figure 4.6: Dose-response plots for selected proteolytic fragments of rituximab after labeling with DEPC at different concentrations varying from 2- to 30-fold DEPC to protein molar ratio (2X to 30X).

The plots of reactive residues from different antibody domains (V_H, C_H1, C_H2, C_H3, V_L, and C_L) in heavy chain (HC) and light chain (LC) are shown here. From all of the representative peptides, linear relationships are observed between the unmodified fraction $\ln(\frac{[X]_0[P]}{[X][P]_0})$ and the DEPC concentrations at low reagent concentrations up to 6X [plots (a) to (i)] or beyond 6X [plots (j) to (p)]. The rate coefficient (k) values in M^{-1} s^{-1} are obtained by dividing the measured slopes by the reaction time (see **Equation 4.8**), and the error bars are calculated from standard error of slope.

❑ **Linearity up to 6X**

(a) HC_K58/K59   CDR-H2 (46-59, V$_H$)
H$_2$N- ALEWLADIWWDD**KK**-COOH

$k = 8.0 \pm 0.2 \times 10^{-1}$

2X
4X
6X 8X
10X
15X 20X 25X 30X

y-axis: ln([X]$_0$[P]/[X][P]$_0$)
x-axis: [X]$_0$ (μM)

(b) HC_S127   (125-136, C$_H$1)
H$_2$N- GP**S**VFPLAPSSK-COOH

$k = 3.6 \pm 0.1 \times 10^{-3}$

y-axis: ln([X]$_0$[P]/[X][P]$_0$)
x-axis: [X]$_0$ (μM)

(c) HC_T253   (252-258, C$_H$2)
H$_2$N- D**T**LMISR-COOH

$k = 4.7 \pm 0.5 \times 10^{-3}$

y-axis: ln([X]$_0$[P]/[X][P]$_0$)
x-axis: [X]$_0$ (μM)

(d) HC_S257   (252-258, C$_H$2)
H$_2$N- DTLMI**S**R-COOH

$k = 3.1 \pm 0.2 \times 10^{-3}$

y-axis: ln([X]$_0$[P]/[X][P]$_0$)
x-axis: [X]$_0$ (μM)

(e) LC_T50/S51/K52   CDR-L2 (45-52, V$_L$)
H$_2$N- LLIYD**TSK**-COOH

$k = 1.22 \pm 0.04 \times 10^{-2}$

y-axis: ln([X]$_0$[P]/[X][P]$_0$)
x-axis: [X]$_0$ (μM)

❑ **Linearity beyond 6X**

(f) HC_K13   (6-18, V$_H$)
H$_2$N-ESGPALV**K**PTQTL-COOH

$k = 3.2 \pm 0.2 \times 10^{-1}$

2X
4X
6X
8X
10X
15X
20X
25X 30X

y-axis: ln([X]$_0$[P]/[X][P]$_0$)
x-axis: [X]$_0$ (μM)

(g) HC_S134/S135/K136   (125-136, C$_H$1)
H$_2$N- GPSVFPLAP**SSK**-COOH

$k = 2.9 \pm 0.2 \times 10^{-3}$

y-axis: ln([X]$_0$[P]/[X][P]$_0$)
x-axis: [X]$_0$ (μM)

(h) HC_Y152   (151-161, C$_H$1)
H$_2$N- D**Y**FPEPVTVSW-COOH

$k = 4.5 \pm 0.3 \times 10^{-3}$

y-axis: ln([X]$_0$[P]/[X][P]$_0$)
x-axis: [X]$_0$ (μM)

(i) HC_Y281   (279-292, C$_H$2)
H$_2$N- FNW**Y**VDGVEVHNAK-COOH

$k = 3.6 \pm 0.1 \times 10^{-1}$

y-axis: ln([X]$_0$[P]/[X][P]$_0$)
x-axis: [X]$_0$ (μM)

(j) HC_H288/K291   (279-292, C$_H$2)
H$_2$N- FNWYVDGVEV**H**NA**K**-COOH

$k = 3.5 \pm 0.1 \times 10^{-1}$

y-axis: ln([X]$_0$[P]/[X][P]$_0$)
x-axis: [X]$_0$ (μM)

(k) HC_Y376   (374-395, C$_H$3)
H$_2$N- GF**Y**PSDIAVEWESNGQPENNYK-COOH

$k = 1.1 \pm 0.1 \times 10^{-2}$

y-axis: ln([X]$_0$[P]/[X][P]$_0$)
x-axis: [X]$_0$ (μM)

159

Figure 4.7: Dose-response plots for selected proteolytic fragments of NISTmAb after labeling with DEPC at different concentrations varying from 2- to 30-fold DEPC to protein molar ratio (2X to 30X).

The plots of reactive residues from different antibody domains (V_H, C_H1, C_H2, C_H3, V_L, and C_L) in heavy chain (HC) and light chain (LC) are shown here. From all of the representative peptides, linear relationships are observed between the unmodified fraction $\ln(\frac{[X]_0[P]}{[X][P]_0})$ and the DEPC concentrations at low reagent concentrations up to 6X [plots (a) to (e)] or beyond 6X [plots (f) to (o)]. The rate coefficient (k) values in M^{-1} s^{-1} are obtained by dividing the measured slopes by the reaction time (see **Equation 4.8**), and the error bars are calculated from standard error of slope.

A closer look at the labeling data reveals that some antibody domains are more susceptible to labeling-induced structural changes than others (**Table 4.1** and **Figure 4.8**). Most notably, the F_{ab} has a greater percentage of sites (60%) whose linearity goes beyond 6-fold DEPC. In contrast, the C_H2 domain has a smaller percentage (38%) of sites that maintain linearity past 6-fold DEPC. These observations are consistent with the known stability differences of the F_{ab} and C_H2 domains. Previous studies have identified the melting temperatures (Tm) to be 71°C and 74°C for the C_H2 and F_{ab} domains of rituximab, respectively, and 69°C and 83°C for those of the NISTmAb.[44, 45] The C_H2 domains of both antibodies have lower T_m values than that of F_{ab} domains, which may explain why the C_H2 domain is more sensitive to labeling-induced structural perturbations. Upon comparing the F_{ab} sites in each antibody, the NISTmAb has higher number of sites whose linearity goes beyond 6-fold DEPC (70%) than rituximab (50%). Perhaps the higher overall stability of NISTmAb's F_{ab} explains why it has higher number of sites whose linearity goes beyond 6-fold DEPC.
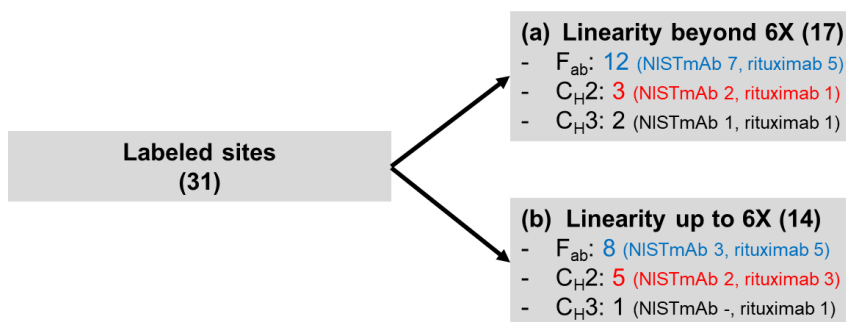
Figure 4.8: Trends of structural perturbations with regard to different antibody domains.

The F_{ab} has a greater percentage of sites whose linearity goes beyond 6X (60%, 12 out of 20 sites) whereas C_H2 has a smaller percentage of sites whose linearity goes beyond 6X (38%, 3 out of 8 sites). Considering the number of F_{ab} sites in each antibody, NISTmAb has higher number of sites whose linearity is beyond 6X (70%, 7 out of 10 sites) while rituximab has lower number (50%, 5 out of 10 sites).

4.3.3 Labeling rate coefficients can be determined from dose-response plots

DEPC modification rate coefficients (k) range from 2.2 x 10^{-3} to 8.0 x 10^{-1} M^{-1} s^{-1} (**Table 4.1**). These are 'apparent' rate coefficients because differences in ionization efficiencies between the labeled and unlabeled peptides make it difficult to obtain real rate coefficients from the experiments. The reaction rate coefficients tend to be higher for His and Lys residues than for Ser, Thr, and Tyr residues. Around 85% of His/Lys residues have k values greater than 1 x 10^{-2} M^{-1} s^{-1}, while about 50% of the Ser/Thr/Tyr residues have k values greater than 1 x 10^{-2} M^{-1} s^{-1}. These results are consistent with our previous studies that His and Lys side chains have higher intrinsic reactivities towards DEPC labeling than Ser, Thr, and Tyr side chains.[19, 31, 40]

Because the two mAbs have the same constant domain sequences, and presumably similar/ structures in these regions, we also can compare DEPC labeling rate coefficients for eight common residues in these constant regions (C_L, C_H1, and F_c) to understand how the formulation affects the labeling rates. Upon comparing the common residues, we find that H289/K292 (H288/K291) and Y377 (Y376) in the heavy chain and T108 in the light chain have essentially the same rate coefficients for the two mAbs. Note that the NISTmAb's residue numbers are shown in parentheses as its heavy chain has one less residue at the N-terminus than rituximab. In contrast, the rate coefficients for Y153 (Y152), Y282 (Y281), T254 (T253), and S258 (S257) in the heavy chain change by a factor of 2 to 4, and the rate coefficient for S170/T171/Y172 in the light chain changes by a factor of almost 60. The difference in labeling rate coefficients for these residues can be ascribed to (a) differences in the rituximab and NISTmAb formulations and/or (b) changes in the

microenvironment around these residues due to the formulation differences or non-identical local structures in the constant regions of these mAbs.

The key differences of the rituximab and NISTmAb formulations are pH (6.5 and 6.0, respectively) and the presence of polysorbate 80 in the former. The residues Y153, Y282, T254 have nearby acidic residues (within 4 Å distance). The difference in the pH of the two mAb formulations could affect the protonation states of these adjacent acidic residues, thereby influencing the nucleophilicity and thus the DEPC reactivity of these Tyr, Ser, and Thr residues.[19] In contrast, there are no acidic residues adjacent to H289/K292, Y377, and T108, perhaps explaining why the rate coefficients for these residues are almost the same for the two mAbs. Interestingly, S170 has a nearby acidic residue in the NISTmAb, but in rituximab an Arg residue is close, perhaps explaining why S170/T171/Y172 is 60-fold more reactive in rituximab. Moreover, the microenvironment around these residues in each mAb is quite different in terms of nearby hydrophobic residues and other polar residues.

4.3.4 Multi-domain antibodies can be labeled with more than one DEPC molecules before being structurally perturbed

The 'single hit rule' in CL indicates that the addition of subsequent covalent labels after the first will necessarily occur to a chemically perturbed protein and should thus be avoided. We were interested, though, to see how many DEPC labels a multi-domain protein could accommodate before structural changes actually occurred, as evidenced by changes in the reaction kinetics. Because the dose-response plots for both antibodies were still linear up to a 6-fold excess of DEPC, we separately reacted each antibody with this level of DEPC and then measured the extent of labeling on the intact light and heavy chains using LC-MS

We find that the weighted average numbers of labels on the light chains of rituximab and NISTmAb are 0.6 and 0.7, respectively, whereas the weighted average number of labels on the two heavy chains of rituximab and NISTmAb are 2.3 and 2.7, respectively (**Figure 4.9**). Note that the heavy chain's spectra are complicated by N-glycan heterogeneity. Enzymatic removal of the glycoform to reduce spectral complexity is possible, but we did not perform this because we wanted to limit any possible hydrolysis of the DEPC labeled sites during the relatively long enzymatic reaction time. Because each protein has two light chains, the total extent of labeling is 3.5 for rituximab and 4.1 for NISTmAb.

The fact that both antibodies can accommodate 3 or 4 labels before any structural perturbation is observed suggests that the 'single hit rule' is not relevant to such large multi-domain proteins such as antibodies. Previous work in our group on relatively small (< 30 kDa) single-domain proteins showed that modification conditions that lead to 0.8 to 1.2 labels per protein on average can maintain the structural integrity of these proteins during the labeling.[31, 46, 47] Being able to add more than one label per mAb is beneficial for CL-MS studies, as it allows a more sensitive measure of structure because more modifications are detected. It is perhaps not surprising that multi-domain proteins such as mAbs are capable of accommodating more than one label, as the multi-domain structure of these proteins likely make them less readily perturbed when modifications are occurring at distant sites (e.g. different domains). This explanation is supported by the success of therapeutic antibody-drug conjugates (ADCs) that have hydrophobic drugs attached. Conjugation of these drugs, often at multiple sites, do not significantly perturb the structure and function of the multi-domain antibody.[48-51] While up to 4 DEPC modifications can be accommodated by the mAbs studied here, these proteins do have 12 domains, so

modification-induced structural changes still do occur across domains. Overall, our findings indicate that multi-domain antibodies can be labeled with more than one DEPC molecule before being structurally perturbed, thus improving the CL-MS resolution substantially compared to the single-hit rule. In addition, these results could be helpful for understanding the extent to which ADCs could be conjugated with drug molecules before structural perturbations are observed.
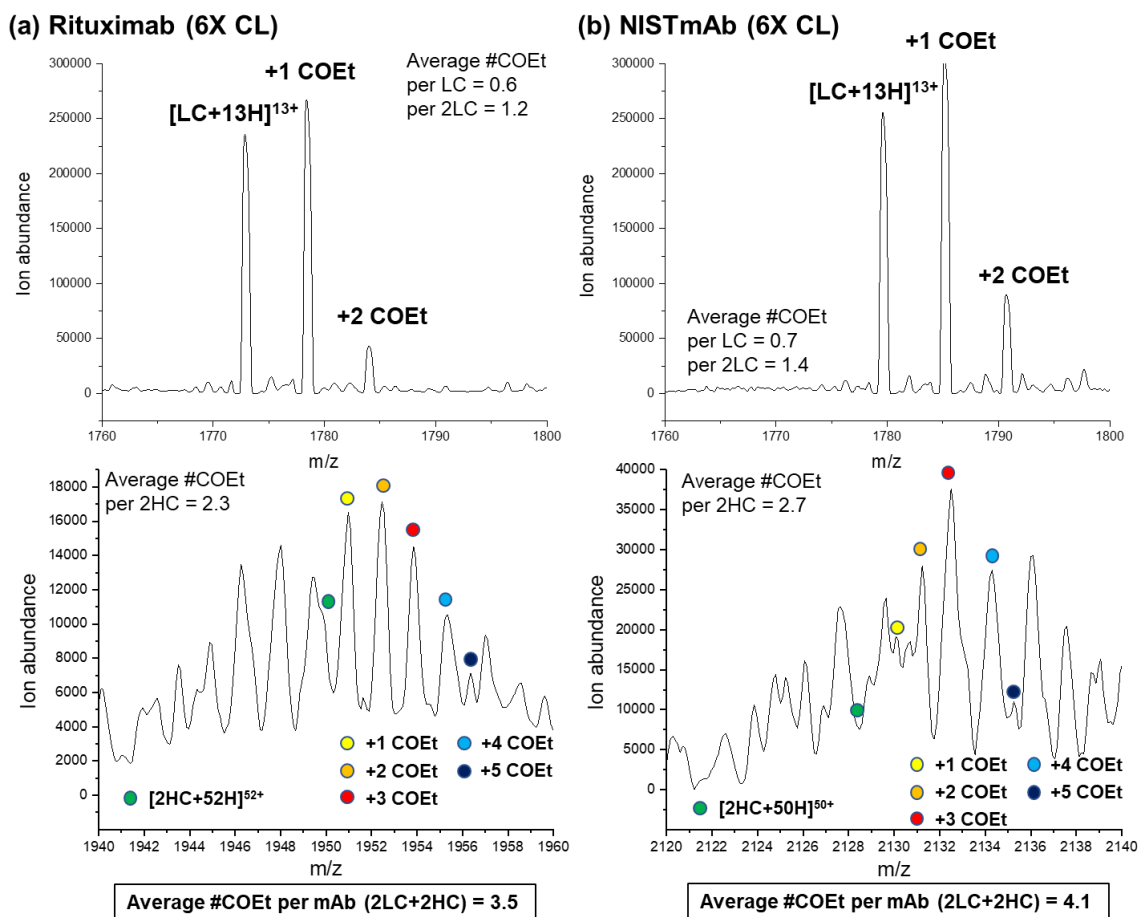
Figure 4.9: Mass spectra obtained from LC-MS analyses of DEPC-labeled mAbs, showing the extent of modification for the light and heavy chains of (a) rituximab and (b) NISTmAb after labeling at 6-fold DEPC:protein molar ratio.
COEt refers to a carbethoxy group that is the product upon modification with DEPC. Because of glycan heterogeneity on the heavy chain (HC) and signal overlap with the light chain (LC), only the G1F/G1F glycoform of a non-reduced heavy chain (2HC) is shown.

4.4 Conclusion

Measuring the kinetics of DEPC CL reactions can reveal the DEPC concentrations at which labeling-induced structural perturbations occur for antibodies. For two separate antibodies that have different formulations, we find that a 6-fold molar excess of DEPC can be used without perturbing the protein's structure, suggesting that this is a reliable DEPC concentration for CL studies of antibodies. This practical finding should avoid the need for time-consuming optimization of DEPC CL conditions for antibodies. Moreover, at these DEPC:mAb labeling ratios, up to four DEPC modifications can be added, indicating that these multi-domain proteins can accommodate more than one label without being structurally perturbed. Interestingly, most of the domains in the NISTmAb maintain their structural integrity at higher DEPC concentrations than rituximab, which is consistent with the greater stability of this mAb. This observation may suggest that more thermally stable proteins could accommodate even higher levels of labeling, allowing a more sensitive measure of HOS by DEPC CL. Overall, these studies improve our ability to use DEPC CL to study the HOS of protein therapeutics.

4.5 References

1.   Berkowitz, S. A.; Engen, J. R.; Mazzeo, J. R.; Jones, G. B., Analytical tools for characterizing biopharmaceuticals and the implications for biosimilars. *Nature Reviews Drug Discovery* **2012,** *11* (7), 527-540.
2.   Zhang, H.; Cui, W.; Gross, M. L., Mass spectrometry for the biophysical characterization of therapeutic monoclonal antibodies. *FEBS Letters* **2014,** *588* (2), 308-317.
3.   Rogstad, S.; Faustino, A.; Ruth, A.; Keire, D.; Boyne, M.; Park, J., A Retrospective Evaluation of the Use of Mass Spectrometry in FDA Biologics License Applications. *Journal of The American Society for Mass Spectrometry* **2017,** *28* (5), 786-794.
4.   Rathore, D.; Faustino, A.; Schiel, J.; Pang, E.; Boyne, M.; Rogstad, S., The role of mass spectrometry in the characterization of biologic protein products. *Expert Review of Proteomics* **2018,** *15* (5), 431-449.

5.   Wang, W.; Singh, S.; Zeng, D. L.; King, K.; Nema, S., Antibody Structure, Instability, and Formulation. *Journal of Pharmaceutical Sciences* **2007,** *96* (1), 1-26.

6.   Dingman, R.; Balu-Iyer, S. V., Immunogenicity of Protein Pharmaceuticals. *Journal of Pharmaceutical Sciences* **2019,** *108* (5), 1637-1654.

7.   Frokjaer, S.; Otzen, D. E., Protein drug stability: a formulation challenge. *Nature Reviews Drug Discovery* **2005,** *4* (4), 298-306.

8.   Abbott, W. M.; Damschroder, M. M.; Lowe, D. C., Current approaches to fine mapping of antigen-antibody interactions. *Immunology* **2014,** *142* (4), 526-535.

9.   Opuni, K. F. M.; Al-Majdoub, M.; Yefremova, Y.; El-Kased, R. F.; Koy, C.; Glocker, M. O., Mass spectrometric epitope mapping. *Mass Spectrom Rev* **2018,** *37* (2), 229-241.

10.   Majumdar, R.; Middaugh, C. R.; Weis, D. D.; Volkin, D. B., Hydrogen-Deuterium Exchange Mass Spectrometry as an Emerging Analytical Tool for Stabilization and Formulation Development of Therapeutic Monoclonal Antibodies. *Journal of Pharmaceutical Sciences* **2015,** *104* (2), 327-345.

11.   Iacob, R. E.; Bou-Assaf, G. M.; Makowski, L.; Engen, J. R.; Berkowitz, S. A.; Houde, D., Investigating Monoclonal Antibody Aggregation Using a Combination of H/DX-MS and Other Biophysical Measurements. *Journal of Pharmaceutical Sciences* **2013,** *102* (12), 4315-4329.

12.   Bommana, R.; Chai, Q.; Schöneich, C.; Weiss, W. F.; Majumdar, R., Understanding the Increased Aggregation Propensity of a Light-Exposed IgG1 Monoclonal Antibody Using Hydrogen Exchange Mass Spectrometry, Biophysical Characterization, and Structural Analysis. *Journal of Pharmaceutical Sciences* **2018,** *107* (6), 1498-1511.

13.   Zhang, A.; Singh, S. K.; Shirts, M. R.; Kumar, S.; Fernandez, E. J., Distinct Aggregation Mechanisms of Monoclonal Antibody Under Thermal and Freeze-Thaw Stresses Revealed by Hydrogen Exchange. *Pharmaceutical Research* **2012,** *29* (1), 236-250.

14.   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.

15.   Houde, D.; Berkowitz, S. A.; Engen, J. R., The Utility of Hydrogen/Deuterium Exchange Mass Spectrometry in Biopharmaceutical Comparability Studies. *Journal of Pharmaceutical Sciences* **2011,** *100* (6), 2071-2086.

16.   Coales, S. J.; Tuske, S. J.; Tomasso, J. C.; Hamuro, Y., Epitope mapping by amide hydrogen/deuterium exchange coupled with immobilization of antibody, on-line proteolysis, liquid chromatography and mass spectrometry. *Rapid Communications in Mass Spectrometry* **2009,** *23* (5), 639-647.

17.   Malito, E.; Faleri, A.; Lo Surdo, P.; Veggi, D.; Maruggi, G.; Grassi, E.; Cartocci, E.; Bertoldi, I.; Genovese, A.; Santini, L.; Romagnoli, G.; Borgogni, E.; Brier, S.; Lo Passo, C.; Domina, M.; Castellino, F.; Felici, F.; van der Veen, S.; Johnson, S.; Lea, S. M.; Tang, C. M.; Pizza, M.; Savino, S.; Norais, N.; Rappuoli, R.; Bottomley, M. J.; Masignani, V., Defining a protective epitope on factor H binding protein, a key meningococcal virulence factor and vaccine antigen. *Proceedings of the National Academy of Sciences* **2013,** *110* (9), 3304.

18.	Zhang, Q.; Willison, L. N.; Tripathi, P.; Sathe, S. K.; Roux, K. H.; Emmett, M. R.; Blakney, G. T.; Zhang, H.-M.; Marshall, A. G., Epitope Mapping of a 95 kDa Antigen in Complex with Antibody by Solution-Phase Amide Backbone Hydrogen/Deuterium Exchange Monitored by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Analytical Chemistry* **2011,** *83* (18), 7129-7136.

19.	Limpikirati, P.; Liu, T.; Vachet, R. W., Covalent labeling-mass spectrometry with non-specific reagents for studying protein structure and interactions. *Methods* **2018,** *144*, 79-93.

20.	Mendoza, V. L.; Vachet, R. W., Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spectrom Rev* **2009,** *28* (5), 785-815.

21.	Xu, G.; Chance, M. R., Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics. *Chem. Rev.* **2007,** *107* (8), 3514-3543.

22.	Zhang, B.; Cheng, M.; Rempel, D.; Gross, M. L., Implementing fast photochemical oxidation of proteins (FPOP) as a footprinting approach to solve diverse problems in structural biology. *Methods* **2018,** *144*, 94-103.

23.	Deperalta, G.; Alvarez, M.; Bechtel, C.; Dong, K.; McDonald, R.; Ling, V., Structural analysis of a therapeutic monoclonal antibody dimer by hydroxyl radical footprinting. *mAbs* **2013,** *5* (1), 86-101.

24.	Li, J.; Wei, H.; Krystek, S. R.; Bond, D.; Brender, T. M.; Cohen, D.; Feiner, J.; Hamacher, N.; Harshman, J.; Huang, R. Y. C.; Julien, S. H.; Lin, Z.; Moore, K.; Mueller, L.; Noriega, C.; Sejwal, P.; Sheppard, P.; Stevens, B.; Chen, G.; Tymiak, A. A.; Gross, M. L.; Schneeweis, L. A., Mapping the Energetic Epitope of an Antibody/Interleukin-23 Interaction with Hydrogen/Deuterium Exchange, Fast Photochemical Oxidation of Proteins Mass Spectrometry, and Alanine Shave Mutagenesis. *Analytical Chemistry* **2017,** *89* (4), 2250-2258.

25.	Zhang, Y.; Wecksler, A. T.; Molina, P.; Deperalta, G.; Gross, M. L., Mapping the Binding Interface of VEGF and a Monoclonal Antibody Fab-1 Fragment with Fast Photochemical Oxidation of Proteins (FPOP) and Mass Spectrometry. *Journal of The American Society for Mass Spectrometry* **2017,** *28* (5), 850-858.

26.	Jones, L. M.; B. Sperry, J.; A. Carroll, J.; Gross, M. L., Fast Photochemical Oxidation of Proteins for Epitope Mapping. *Analytical Chemistry* **2011,** *83* (20), 7657-7661.

27.	Watson, C.; Sharp, J. S., Conformational analysis of therapeutic proteins by hydroxyl radical protein footprinting. *AAPS J.* **2012,** *14* (2), 206-217.

28.	Kaur, P.; Tomechko, S. E.; Kiselar, J.; Shi, W.; Deperalta, G.; Wecksler, A. T.; Gokulrangan, G.; Ling, V.; Chance, M. R., Characterizing monoclonal antibody structure by carboxyl group footprinting. *mAbs* **2015,** *7* (3), 540-552.

29.	Limpikirati, P.; Hale, J. E.; Hazelbaker, M.; Huang, Y.; Jia, Z.; Yazdani, M.; Graban, E. M.; Vaughan, R. C.; Vachet, R. W., Covalent labeling and mass spectrometry reveal subtle higher order structural changes for antibody therapeutics. *mAbs* **2019,** *11* (3), 463-476.

30.	Borotto, N. B.; Zhou, Y.; Hollingsworth, S. R.; Hale, J. E.; Graban, E. M.; Vaughan, R. C.; Vachet, R. W., Investigating Therapeutic Protein Structure with Diethylpyrocarbonate Labeling and Mass Spectrometry. *Analytical Chemistry* **2015,** *87* (20), 10627-10634.

31.   Mendoza, V. L.; Vachet, R. W., Protein Surface Mapping Using Diethylpyrocarbonate with Mass Spectrometric Detection. *Analytical Chemistry* **2008,** *80* (8), 2895-2904.

32.   Manzi, L.; Barrow, A. S.; Scott, D.; Layfield, R.; Wright, T. G.; Moses, J. E.; Oldham, N. J., Carbene footprinting accurately maps binding sites in protein–ligand and protein–protein interactions. *Nature Communications* **2016,** *7* (1), 13288.

33.   Gau, B. C.; Sharp, J. S.; Rempel, D. L.; Gross, M. L., Fast Photochemical Oxidation of Protein Footprints Faster than Protein Unfolding. *Analytical Chemistry* **2009,** *81* (16), 6563-6571.

34.   Jumper, C. C.; Schriemer, D. C., Mass Spectrometry of Laser-Initiated Carbene Reactions for Protein Topographic Analysis. *Analytical Chemistry* **2011,** *83* (8), 2913-2920.

35.   Hambly, D. M.; Gross, M. L., Laser Flash Photolysis of Hydrogen Peroxide to Oxidize Protein Solvent-Accessible Residues on the Microsecond Timescale. *Journal of the American Society for Mass Spectrometry* **2005,** *16* (12), 2057-2063.

36.   Jumper, C. C.; Bomgarden, R.; Rogers, J.; Etienne, C.; Schriemer, D. C., High-Resolution Mapping of Carbene-Based Protein Footprints. *Analytical Chemistry* **2012,** *84* (10), 4411-4418.

37.   Gupta, S.; Celestre, R.; Petzold, C. J.; Chance, M. R.; Ralston, C., Development of a microsecond X-ray protein footprinting facility at the Advanced Light Source. *Journal of Synchrotron Radiation* **2014,** *21* (4), 690-699.

38.   Kiselar, J. G.; Janmey, P. A.; Almo, S. C.; Chance, M. R., Structural Analysis of Gelsolin Using Synchrotron Protein Footprinting. *Molecular & Cellular Proteomics* **2003,** *2* (10), 1120.

39.   Kiselar, J. G.; Maleknia, S. D.; Sullivan, M.; Downard, K. M.; Chance, M. R., Hydroxyl radical probe of protein surfaces using synchrotron X-ray radiolysis and mass spectrometry. *International Journal of Radiation Biology* **2002,** *78* (2), 101-114.

40.   Limpikirati, P.; Pan, X.; Vachet, R. W., Covalent Labeling with Diethylpyrocarbonate: Sensitive to the Residue Microenvironment, Providing Improved Analysis of Protein Higher Order Structure by Mass Spectrometry. *Analytical Chemistry* **2019,** *91* (13), 8516-8523.

41.   Schrodinger, LLC *The PyMOL Molecular Graphics System, Version 1.8*, 2015.

42.   Asuero, A. G.; Sayago, A.; González, A. G., The Correlation Coefficient: An Overview. *Critical Reviews in Analytical Chemistry* **2006,** *36* (1), 41-59.

43.   Miller, J. N.; Miller, J. C., *Statistics and chemometrics for analytical chemistry*. 6th ed.; Pearson/Prentice Hall: Harlow, England, United Kingdom, 2005.

44.   Andersen, C. B.; Manno, M.; Rischel, C.; Thórólfsson, M.; Martorana, V., Aggregation of a multidomain protein: A coagulation mechanism governs aggregation of a model IgG1 antibody under weak thermal stress. *Protein Science* **2010,** *19* (2), 279-290.

45.   Borotto, N. B.; Degraan-Weber, N.; Zhou, Y.; Vachet, R. W., Label Scrambling During CID of Covalently Labeled Peptide Ions. *J. Am. Soc. Mass Spectrom.* **2014,** *25* (10), 1739-1746.

46.   Liu, T.; Marcinko, T. M.; Kiefer, P. A.; Vachet, R. W., Using Covalent Labeling and Mass Spectrometry To Study Protein Binding Sites of Amyloid Inhibiting Molecules. *Analytical Chemistry* **2017,** *89* (21), 11583-11591.

47.   Liu, T.; Limpikirati, P.; Vachet, R. W., Synergistic Structural Information from Covalent Labeling and Hydrogen–Deuterium Exchange Mass Spectrometry for Protein–Ligand Interactions. *Analytical Chemistry* **2019**.

48.   Pan, L. Y.; Salas-Solano, O.; Valliere-Douglass, J. F., Antibody Structural Integrity of Site-Specific Antibody-Drug Conjugates Investigated by Hydrogen/Deuterium Exchange Mass Spectrometry. *Analytical Chemistry* **2015,** *87* (11), 5669-5676.

49.   Huang, R. Y. C.; O'Neil, S. R.; Lipovšek, D.; Chen, G., Conformational Assessment of Adnectin and Adnectin-Drug Conjugate by Hydrogen/Deuterium Exchange Mass Spectrometry. *Journal of The American Society for Mass Spectrometry* **2018,** *29* (7), 1524-1531.

50.   Pan, L. Y.; Salas-Solano, O.; Valliere-Douglass, J. F., Localized conformational interrogation of antibody and antibody-drug conjugates by site-specific carboxyl group footprinting. *mAbs* **2017,** *9* (2), 307-318.

51.   Pan, L. Y.; Salas-Solano, O.; Valliere-Douglass, J. F., Conformation and Dynamics of Interchain Cysteine-Linked Antibody-Drug Conjugates as Revealed by Hydrogen/Deuterium Exchange Mass Spectrometry. *Analytical Chemistry* **2014,** *86* (5), 2657-2664.

CHAPTER 5

SUMMARY AND FUTURE DIRECTIONS

5.1 Summary

This dissertation has described the use of a diethylpyrocarbonate (DEPC-based covalent labeling (CL) method combined with mass spectrometry (MS) to characterize the higher-order structure (HOS) of proteins and protein therapeutics. DEPC is a simple to use, commercially-available CL reagent. From reactions of amino acid side chains with DEPC, a protein's structural properties can be encoded into the mass of protein, and the site-specific identification of modified residues can be performed by MS. This dissertation aims to address the need in pharmaceutical industry, specifically the protein therapeutics field. Characterizing HOS of monoclonal antibodies (mAbs) is challenging given their size and the multidomain nature. As one of the non-specific reagents, DEPC can readily react with a range of residues in proteins, allowing for higher structural resolution to obtain insight into conformational changes of therapeutic proteins.

In **Chapter 2**, we studied the effect of residue microenvironment on DEPC labeling reactivity of weakly nucleophilic residues (Ser, Thr, and Tyr). We find that in intact proteins Ser, Thr, and Tyr can be modified by DEPC in addition to other residues such as His and Lys, which allows for higher structural resolution to obtain insight into conformational changes of therapeutic proteins. Our findings indicate that a hydrophobic microenvironment influences Ser, Thr, and Tyr reactivity toward DEPC labeling in intact proteins more than solvent accessibility. Meanwhile, His and Lys are less likely to be affected by a hydrophobic microenvironment because they are inherently much more reactive than Ser, Thr, and Tyr, so they do not need high local concentrations of DEPC to

be modified. Additional studies are needed to further verify the hydrophobic microenvironment effect, but the reactivity of specific Ser, Thr, and Tyr residues could be used to provide insight into the local microenvironment around these residues and could be useful in CL-MS-based protein structural prediction. For example, the reactivity of specific Ser, Thr, and Tyr residues in proteins could indicate the presence of nearby hydrophobic residues, and this information could be used to improve confidence in certain structural models. In protein-ligand complexes, we anticipate that if a ligand is hydrophobic enough, the labeling reactivity of close-by Ser, Thr, and Tyr residues should reveal the presence of ligand binding sites.

In **Chapter 3**, we demonstrated the use of DEPC-based CL-MS to detect conformational changes caused by heat stress of therapeutic mAbs. Our DEPC CL-MS technique compares favorably to other MS-based structural techniques (e.g., hydrogen-deuterium exchange [HDX],[1-3] hydroxyl radical footprinting [HRF],[4] and dimethyl labeling[5]) that have been used recently to study HOS changes or identify aggregation interfaces of mAbs. To the best of our knowledge, our CL-MS technique is the first to report a subtle HOS change of a therapeutic mAb at the temperature far below the antibody melting temperature, changes that are not detected by common characterization techniques. At higher stress temperature, conformational changes and aggregation sites are also identified.

CL-MS using DEPC offers some advantages over other MS-based techniques, especially for revealing protein aggregation interfaces and protein-ligand binding sites. In HDX-MS, reduced deuterium exchange along a protein backbone is used to identify binding interfaces, even though protein-protein and protein-ligand interactions are

primarily mediated by side chain interactions. DEPC labeling reports on side chains, thereby providing a more direct indication of binding interfaces. Decreased labeling at specific residues can be used to determine the binding interfaces because protein aggregation and ligand binding decrease the solvent accessibility of the side chains involved in a binding event. CL-MS also does not require specialized equipment like HDX-MS, thereby simplifying the overall approach. HRF is another side-chain labeling technique, but HRF techniques also require specialized lasers or synchrotron sources to generate the radicals. In addition, radicals generate multiple reaction products which can complicate the analyses, as opposed to a single one with DEPC. However, DEPC CL has the slower labeling time scale than HDX and HRF, which should make it only sensitive to changes in solvent accessibility and insensitive to changes in protein dynamics. Thus, changes in protein dynamics at allosteric sites upon protein-protein and protein-ligand interactions are likely not to be reported by DEPC CL-MS.

In **Chapter 4**, we evaluated applicability of DEPC reaction kinetics to ensure site-specific or local structural integrity of multi-domain mAbs during labeling as reliable information about HOS of mAbs can be obtained only when the protein's structural integrity is preserved during the labeling. We find that multi-domain mAbs can withstand many more than one label, which contrasts to previously studied single-domain proteins. This more extensive labeling provides a more sensitive measure of structure, making DEPC-based CL-MS suitable for the HOS analyses of mAbs. However, we do not expect that all big proteins can withstand more than one label per protein because different proteins have different folding and thermal stability. Structure of big proteins with limited stability may be perturbed even after having only one label per protein. In addition, even

though big proteins are likely less perturbed when modifications occur at distant sites, an allosteric effect might be induced by a first label and results in a conformation change at a distant site.

For protein-ligand systems, measuring the kinetics of DEPC CL reactions should be able to reveal the DEPC concentrations at which labeling-induced structural perturbations occur for a protein, ensuring that there is no ligand dissociation upon labeling. If more than one labels on average per protein molecule is possible for the protein-ligand complex, the increased structural resolution can improve the identification of ligand binding sites and allosteric sites.

5.2 Future Directions

The following sections discuss possible future improvements to the covalent labeling method and the structural characterization of protein therapeutics.

5.2.1 Residue Microenvironment Effect on DEPC Covalent Labeling

5.2.1.1 Microenvironment Investigations of DEPC Covalent Labeling Data Obtained from Other Proteins

In **Chapter 2**, we find evidence that the presence of nearby hydrophobic residues increases the DEPC reactivity of Ser, Thr, and Tyr residues in intact proteins.[6] However, the examination of the structural features was performed for only three different proteins, β-2-microglobulin (β2m), ubiquitin, and human growth hormone (hGH) (**Tables A.1, A.2, and A.3 in Appendix A**). In order to get more evidence to support that the nearby hydrophobic residues are essential for the enhanced reactivity of certain Ser, Thr, and Tyr residues, we should expand the microenvironment investigations to DEPC labeling results

174

of Ser, Thr, and Tyr residues in other proteins. In this dissertation, we have used DEPC CL-MS data[7] recently obtained from bovine carbonic anhydrase (BCA), equine holo-myoglobin (MYG), and hen egg-white lysozyme (LYZ) intact proteins. Structural features and DEPC modification percentages for nucleophilic residues in these proteins are shown in **Tables A.4, A.5, and A.6**. Adding data from these three proteins to the data from three proteins in **Chapter 2**, altogether, the total number of Ser, Thr, and Tyr residues in six different proteins is 139 of which 55 residues are found to be labeled by DEPC in these proteins. Interestingly, almost 90% of the labeled Ser, Thr, and Tyr residues that sits close to at least one hydrophobic side chain (**Figure 5.1**). In contrast, only half of the unlabeled Ser, Thr, and Tyr residues have nearby hydrophobic residues (**Figure 5.1**). This additional evidence can emphasize the importance of nearby hydrophobic groups for enhancing Ser, Thr, or Tyr reactivity.
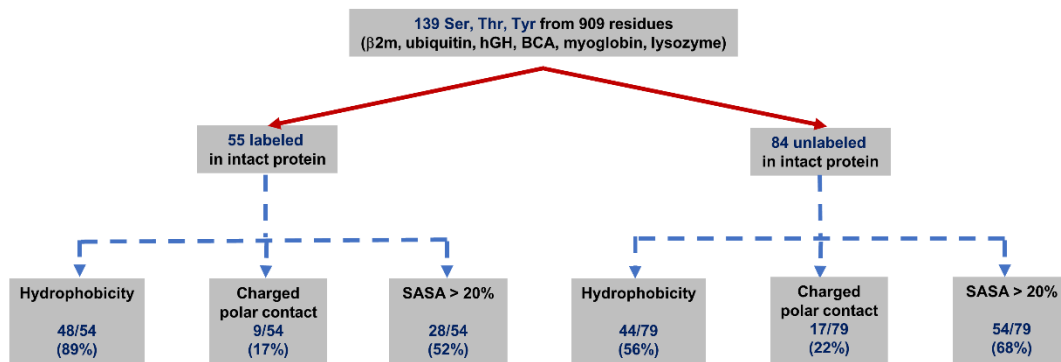
Figure 5.1: Flow chart summarizing the covalent labeling results and the structural features of weakly nucleophilic residues in six different proteins.

However, we do not see evidence that the presence of nearby hydrophobic residues can enhance the DEPC reactivity of Ser, Thr, and Tyr residues in rituximab. Around 70% of both the labeled and the unlabeled Ser, Thr, and Tyr residues have nearby hydrophobic residues. Because a full crystal structure for rituximab is not available, the comparison may

175

have flaws. Another explanation is based on the rituximab formulation (10 mg/mL rituximab) which contains 0.7 mg/mL polysorbate 80 (Tween® 80). As a molar ratio of polysorbate 80 to rituximab is more than 7,000-fold, and the concentration of polysorbate 80 is above its critical micellar concentration (13 to 15 mg/L),[8] we hypothesize that polysorbate 80 may form a micelle around rituximab which may affect the microenvironment effect from hydrophobic residues. To investigate the microenvironment effect regarding the presence of nearby hydrophobic residues without a micelle, labeling experiments should be performed on (a) rituximab with polysorbate 80 removal or (b) NISTmAb whose formulation does not contain any surfactant.

Additionally, because the microenvironment investigations described herein were performed on PDB atomic coordinate of each protein in a manual manner, it is very time consuming to expand the investigations into the larger proteins such as mAbs. A computer script should be written and programmed to identify and determine the number of hydrophobic neighboring residues around Ser, Thr, and Tyr residues from atomic coordinates of the 3D structures of proteins. The computer script will allow further investigations of the microenvironment effect on DEPC covalent labeling of proteins. Ultimately, we expect that the reactivity of certain Ser, Thr, and Tyr residues could be used to indicate the presence of nearby hydrophobic groups and thus could be used as constraints in protein structure prediction and to improve confidence in certain structural models.

5.2.1.2 Quantitative correlation between DEPC Labeling Extent and the Distance of the Close-by Hydrophobic Residues

Investigations of the microenvironment around Ser, Thr, and Tyr residues in intact proteins indicate that hydrophobicity is important for labeling rather than solvent

accessibility or polar interactions. However, we still have no evidence of a quantitative correlation between DEPC labeling extent and the distance of the close-by hydrophobic residues. This lack of a quantitative correlation could be for many reasons. One reason could be the structures we use for determining the distance of the hydrophobic residues. Some of the structures come from X-ray crystallography, and therefore do not take into account of protein dynamics. One might expect that the distances might vary to some extent as the protein undergoes local motion. In order to address this issue, simulation of the molecular dynamics of proteins should be performed to investigate the distance between the hydrophobic residues and Ser, Thr, and Tyr residues. We expect to obtain a quantitative correlation between the distance of the nearby hydrophobic residues and the DEPC modification percentage of these residues.

Another reason for the lack of quantitative correlation could be the fact that we do not measure an absolute level of labeling extent, but instead we measure a relative extent of labeling. The modification levels here do not reflect the absolute quantitation of modified species as the addition of a carbethoxyl group to the modified peptide and the different LC solvent conditions during gradient elution of peptides result in different ionization efficiencies for the unmodified and modified peptides. If possible, we might be able to correct for the ionization efficiency differences between the labeled and unlabeled peptides and thus get a more accurate measurement of the modification percentage.

5.2.1.3 Structure-Reactivity Relationship of Reagent's Functional Group and the Labeling Reactivity at Weakly Nucleophilic Side Chains

Despite the successful use of DEPC, their chemical tunability and/or amino acid coverage could still be improved upon. As our study suggests that hydrophobicity may

contribute to higher labeling reactivity of weak nucleophiles due to the increase in local concentration of a CL reagent at protein's surface,[6] hydrophobicity can be one of the reagent properties to be considered for reagent development for the CL-based structural analysis of proteins and protein therapeutics. Recently, we have developed a new class of CL reagents having an α,β-unsaturated carbonyl (ABUC) scaffold (**Figure 5.2**) for studying protein HOS.[9] ABUC reagents with hydrophobic functional group variations can be synthesized. Effect of hydrophobic moiety (e.g. R_2 = H, Me, Et, n-Bu, t-Bu) on the labeling reactivity at weakly nucleophilic side chains should be evaluated using specific model peptides and proteins.
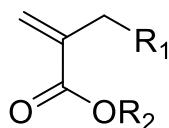
Figure 5.2: Structure of the α,β-unsaturated carbonyl (ABUC) reagent
(R_1 = leaving group, R_2 = hydrophobic functional group)

We anticipate that the hydrophobic moiety of these CL reagents, if not limited by solubility, can contribute to the increase in CL reactivity of these residues. If tuning hydrophobicity does not significantly improve probing resolution of CL reagents, an alternate plan will focus on tuning stability of a leaving group (e.g. with a better leaving group as R_1) and/or electrophilicity of a reaction site (e.g. with a more electron withdrawing group as R_2) to accelerate the labeling reaction. Tuning functional groups of ABUC-scaffold CL reagents will allow the development of novel CL reagents for HOS characterization of protein therapeutics with expanded structural coverage and good sensitivity to detect subtle structural changes.

5.2.2 Structural Characterization of mAb Therapeutics

5.2.2.1 Synergistic Structural Information about Stressed Therapeutic Antibodies from Hydrogen Deuterium Exchange (HDX) and CL-MS

In **Chapter 3**, we demonstrate the ability of DEPC CL-MS to investigate HOS of antibody therapeutics. DEPC can probe solvent accessibility and the microenvironment of side chains, and any changes in DEPC reactivity are indicative of changes in protein conformation. DEPC CL technique can provide structural information which is complementary to what can be obtained from HDX. HDX reports on the protein backbone structure and dynamics while DEPC CL reports on side chain accessibility. CL-MS experimental results can be compared with HDX-MS results to validate the consistency of structural characterization. In addition, because of the different labeling timescales of these two techniques, we predict that there might be synergy between HDX and CL data when studying HOS of mAbs. Using HDX to characterize structures of stressed rituximab after heating at different temperatures may offer further insights into the HOS of antibody undergoing from subtle to massive conformation changes.

5.2.2.2 DEPC CL-MS for Epitope Mapping of Antigen-Antibody Interactions

Covalent labeling techniques are emerging as powerful tools for epitope mapping of mAbs. Previously, hydroxyl radical footprinting (HRF) has been used to perform epitope mapping for various antigen-antibody interactions.[10-12] To demonstrate the applicability of DEPC labeling for epitope mapping, DEPC CL-MS will be used to identify distinct three-dimensional epitopes for adalimumab, infliximab, and golimumab on tumor necrosis factor alpha (TNFα). We expect that the identified epitopes from CL-MS will be consistent with the epitopes obtained from the previously published crystal structures (PDB 3WD5, 5YOY,

and 4G3Y for TNFα in complex with adalimumab, infliximab, and golimumab, respectively). In addition, HDX-MS can be used to provide complementary information to CL-MS regarding the epitope mapping.

5.2.3 DEPC CL Reaction Kinetics for Ensuring Structural Integrity

5.2.3.1 Determination of DEPC Labeling Reaction Order Using Kinetic Isolation Method

The bimolecular reaction between DEPC and a specific site in protein should follow second order kinetics.[9] In **Chapter 4**, we demonstrated that dose-response plots based on reaction kinetics of DEPC CL can be used to reveal labeling-induced structural perturbation for antibodies with an assumption that deviation from the second order kinetics can indicate structural perturbation of mAbs. However, the rate law does not necessarily reflect that stoichiometry of reaction for once the products or intermediates that have been generated might participate in the reaction and affect the rate.[13] In addition, DEPC molecule is symmetric and has two reaction sites, hence we are concerned that the labeling reaction between DEPC and a nucleophilic site may not undergo the regular second order kinetics. Determination of DEPC labeling reaction order should be performed in order to get more understandings about the labeling kinetics. Kinetic isolation method can be used for the determination of the rate law.[13] To reduce complexity of the study, we will use the model peptide sequence of Fmoc-DGXGG-amide (X = nucleophilic residue) which has only single modifiable site for DEPC labeling. The kinetic isolation method will be applied to assess reaction orders with regard to the model peptide and DEPC.

In the kinetic isolation method, the reaction between DEPC and a nucleophilic site in the model peptide is assumed to be under higher-order kinetics with a rate equation shown below (**Equation 5.1**).

Peptide (P) + DEPC → Carbethoxylated P + CO_2 + CH_3CH_2OH

$$Rate = -\frac{d[P]}{dt} = k[DEPC]^a[P]^b \qquad (5.1)$$

where a and b are the reaction orders with regard to DEPC and the model peptide, respectively, [P] is the concentration of unmodified peptide at time t, [X] is the DEPC concentration at time t, and k is the labeling rate coefficient.

➤   Reaction with excess [DEPC]

Given $\qquad k_{DEPC} = k[DEPC]^a$ $\qquad\qquad\qquad\qquad\qquad\qquad$ (5.2)

The rate equation is reduced to $\qquad Rate = -\frac{d[P]}{dt} = k_{DEPC}[P]^b$ $\qquad\qquad$ (5.3)

$$\log \text{Rate} = \log k_{DEPC} + b \log [P] \qquad (5.4)$$

Initial rate method $\qquad\qquad \log \text{Rate}_0 = \log k_{DEPC} + b \log [P]_0$ $\qquad\qquad$ (5.5)

The initial rate (Rate_0) can be measured during an infinitesimal time ($\Delta t \rightarrow 0$), and we assume that 1-min labeling time is small enough.

$$\text{Rate}_0 = -\frac{d[P]}{dt} \approx -\frac{\Delta[P]}{\Delta t} \qquad (5.6)$$

Experimentally, the model peptide at different concentrations ($[P]_0$ = 25, 50, 75, 100, 150 µM) is labeled with DEPC at a fixed excess concentration ($[DEPC]_0$ = 150 µM) for 1 min, and the reaction is then quenched with imidazole. LC-MS is used to determine a DEPC modification percentage in order to obtain $\Delta[P]$ and calculate Rate_0. Reaction order with regard to the model peptide (b) can be obtained from a slope of a plot between

$\log \text{Rate}_0$ and $\log [P]_0$ (see **Equation 5.5**). Our study shows that in a reaction between DEPC and the model peptide Fmoc-DGYGG-amide, the reaction order with regard to the peptide is one. **(Figure 5.3)**

Fmoc-DGYGG-CONH$_2$
$\log \text{Rate}_0 = 0.96(\log [P]_0) - 0.13$
Slope = b = 0.96 ≈ 1
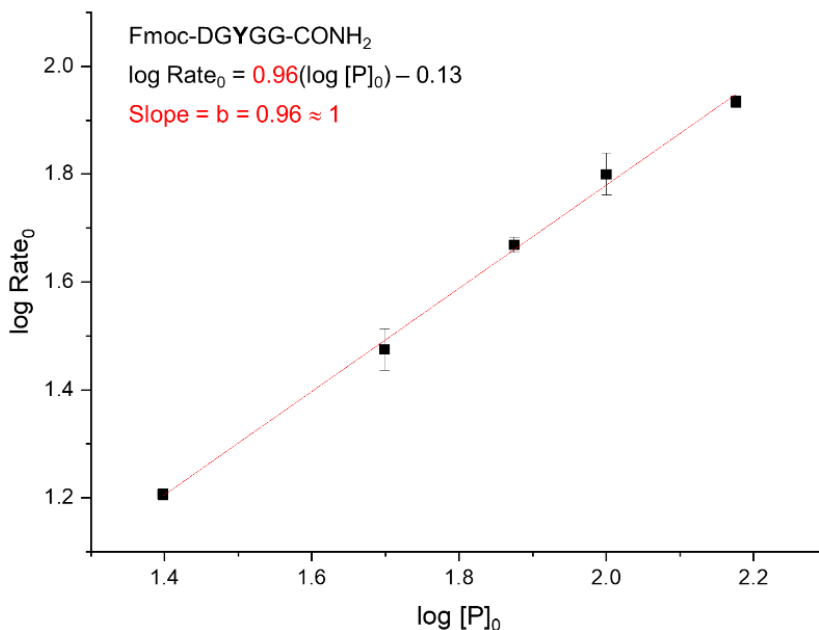
Figure 5.3: Determination of the reaction order with regard to the model peptide using kinetic isolation method

➢ Reaction with excess [P]

Given $\qquad k_P = k[P]^b$ (5.7)

The rate equation is reduced to $\qquad Rate = k_p[DEPC]^a$ (5.8)

$$\log \text{Rate} = \log k_P + a \log [DEPC] \qquad (5.9)$$

Initial rate method $\qquad \log \text{Rate}_0 = \log k_P + a \log [DEPC]_0 \qquad (5.10)$

Rate_0 can be measured during an infinitesimal time ($\Delta t \to 0$), and we assume that 1-min labeling time is small enough.

$$\text{Rate}_0 = -\frac{d[DEPC]}{dt} \approx -\frac{\Delta[DEPC]}{\Delta t} \qquad (5.11)$$

Our experimental plan is as follows. The model peptide at a fixed excess concentration ($[P]_0$) is labeled with DEPC at different concentrations ($[DEPC]_0$) for 1 min, the second peptide (e.g. Cys-containing peptide) is then added to react with the remaining DEPC, and the reaction is finally quenched with imidazole. LC-MS is used to determine a DEPC modification percentage of the second peptide, and a concentration of the remaining DEPC can be obtained from a plot between DEPC concentration and modification percentage of the second peptide. $\Delta[DEPC]$ and Rate_0 can be calculated from the concentration of the remaining DEPC. Reaction order with regard to DEPC (a) can be obtained from a slope of a kinetic plot between $\log \text{Rate}_0$ and $\log [DEPC]_0$ (see **Equation 5.10**).

5.2.3.2 Determination of DEPC Modification Rate Coefficients: Correction for the Ionization Efficiency Differences

In **Chapter 4**, DEPC modification rate coefficients (k) determined from dose-response plots are 'apparent' rate coefficients because differences in ionization efficiencies between the labeled and unlabeled peptides make it difficult to obtain real rate coefficients from the experiments. The ionization efficiency differences between the labeled and unlabeled peptides might be corrected in order to get a more accurate rate measurement. To accomplish this, the ion intensities of the labeled peptides should be plotted as a function of DEPC concentration up to the molar ratio at which deviation occurs. A plot for the unlabeled peptide should be separately prepared. The slopes of these plots would effectively be the sensitivities of the peptides, and the ratios of these sensitivities should be the ratio of the ionization efficiencies. We expect that this correction might work if the reproducibility of the data is good enough.

To see if there is any trend in the data with regard to rate coefficients, plotting the measured rates for a given residue type as a function of SASA and as a function of the number of nearby hydrophobic groups (or distance of hydrophobic groups or some combination of those factors) should be considered.

5.3 References

1.   Zhang, A.; Singh, S. K.; Shirts, M. R.; Kumar, S.; Fernandez, E. J., Distinct Aggregation Mechanisms of Monoclonal Antibody Under Thermal and Freeze-Thaw Stresses Revealed by Hydrogen Exchange. *Pharmaceutical Research* **2012,** *29* (1), 236-250.
2.   Iacob, R. E.; Bou-Assaf, G. M.; Makowski, L.; Engen, J. R.; Berkowitz, S. A.; Houde, D., Investigating Monoclonal Antibody Aggregation Using a Combination of H/DX-MS and Other Biophysical Measurements. *Journal of Pharmaceutical Sciences* **2013,** *102* (12), 4315-4329.
3.   Bommana, R.; Chai, Q.; Schöneich, C.; Weiss, W. F.; Majumdar, R., Understanding the Increased Aggregation Propensity of a Light-Exposed IgG1 Monoclonal Antibody Using Hydrogen Exchange Mass Spectrometry, Biophysical Characterization, and Structural Analysis. *Journal of Pharmaceutical Sciences* **2018,** *107* (6), 1498-1511.
4.   Deperalta, G.; Alvarez, M.; Bechtel, C.; Dong, K.; McDonald, R.; Ling, V., Structural analysis of a therapeutic monoclonal antibody dimer by hydroxyl radical footprinting. *mAbs* **2013,** *5* (1), 86-101.
5.   Jhan, S.-Y.; Huang, L.-J.; Wang, T.-F.; Chou, H.-H.; Chen, S.-H., Dimethyl Labeling Coupled with Mass Spectrometry for Topographical Characterization of Primary Amines on Monoclonal Antibodies. *Analytical Chemistry* **2017,** *89* (7), 4255-4263.
6.   Limpikirati, P.; Pan, X.; Vachet, R. W., Covalent Labeling with Diethylpyrocarbonate: Sensitive to the Residue Microenvironment, Providing Improved Analysis of Protein Higher Order Structure by Mass Spectrometry. *Analytical Chemistry* **2019,** *91* (13), 8516-8523.
7.   Liu, T.; Limpikirati, P.; Vachet, R. W., Synergistic Structural Information from Covalent Labeling and Hydrogen–Deuterium Exchange Mass Spectrometry for Protein–Ligand Interactions. *Analytical Chemistry* **2019**.
8.   Neugebauer, J. M., Detergents: An overview. In *Methods in Enzymology*, Deutscher, M. P., Ed. Academic Press: 1990; Vol. 182, pp 239-253.
9.   Mendoza, V. L.; Vachet, R. W., Protein Surface Mapping Using Diethylpyrocarbonate with Mass Spectrometric Detection. *Analytical Chemistry* **2008,** *80* (8), 2895-2904.
10.   Li, J.; Wei, H.; Krystek, S. R.; Bond, D.; Brender, T. M.; Cohen, D.; Feiner, J.; Hamacher, N.; Harshman, J.; Huang, R. Y. C.; Julien, S. H.; Lin, Z.; Moore, K.; Mueller, L.; Noriega, C.; Sejwal, P.; Sheppard, P.; Stevens, B.; Chen, G.; Tymiak, A.

A.; Gross, M. L.; Schneeweis, L. A., Mapping the Energetic Epitope of an Antibody/Interleukin-23 Interaction with Hydrogen/Deuterium Exchange, Fast Photochemical Oxidation of Proteins Mass Spectrometry, and Alanine Shave Mutagenesis. *Analytical Chemistry* **2017,** *89* (4), 2250-2258.

11.   Zhang, Y.; Wecksler, A. T.; Molina, P.; Deperalta, G.; Gross, M. L., Mapping the Binding Interface of VEGF and a Monoclonal Antibody Fab-1 Fragment with Fast Photochemical Oxidation of Proteins (FPOP) and Mass Spectrometry. *Journal of The American Society for Mass Spectrometry* **2017,** *28* (5), 850-858.

12.   Jones, L. M.; B. Sperry, J.; A. Carroll, J.; Gross, M. L., Fast Photochemical Oxidation of Proteins for Epitope Mapping. *Analytical Chemistry* **2011,** *83* (20), 7657-7661.

13.   Atkins, P.; de Paula, J., *Atkins' Physical Chemistry. 10th ed*. 6th ed.; Oxford University Press: Oxford, England, United Kingdom, 2014.

APPENDIX A

STRUCTURAL FEATURES AND DEPC MODIFICATION PERCENTAGES

FOR NUCLEOPHILIC RESIDUES IN DIFFERENT PROTEINS

Table A.1: DEPC modification percentages of nucleophilic residues in β2m intact protein and its proteolytic peptides.
Each experiment was performed in triplicate or quadruplicate (n = 3 or 4). Error bars shown in a table are standard deviations. Listed along are structural features of each residue from PDB 1JNJ. ✔ and ✖ represent presence or absence of each structural element, respectively. (+) and (-) represent positively charged and negatively charged polar contacts, respectively.

β2m residue	Secondary structure	pKa‡	%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)	Peptide (DEPC:peptide = 4:1)	Peptide (DEPC:peptide = 50:1)
			Solvent accessibility*		**Microenvironment**		**% DEPC CL**		
T4	Random coil	N/A[a]	81.5	✔	✖ **- Overall: Uncharged, nothing**	✔ - OH close to P5, L87	59 ± 5 (T4 & N-term)	N.D.[c]	0.02 ± 0.01
K6	β-sheet	10.42	79.6	✔	✖ - H-bonding with backbone CO of D8 **- Overall: Uncharged, polar contact**	✖	41 ± 5	N/A[b]	N/A
Y10	β-sheet	10.98	41.6	✔	✖ - H-bonding with side chain $CONH_2$ of N24 **- Overall: Uncharged, polar contact**	✔ - OH close to Y26, L65	15 ± 2	N.D.	3.9 ± 0.8
S11	β-sheet	N/A	10.3	✖	✖ - H-bonding with side chain aromatic amine of W95 - H-bonding with backbone CO of F22, R97, M99 - H-bonding with backbone NH of R12, D98, M99 **- Overall: Uncharged, polar contact**	✔ - OH close to A15, L23, F70, M99	3.0 ± 0.5	N.D.	0.02 ± 0.01
H13	Random coil	6.09	47.7	✔	✖ **- Overall: Uncharged, nothing**	✔ - imidazole amine close to F22	30 ± 3	N/A	N/A

[a] N/A = not applicable; [b] N-terminal blocking using SNHSA also results in side reaction (acetylation) at Lys and His residues in free peptides, so they are not discussed here.

[c] N.D. = not detected (less than DEPC labeling threshold of 0.01%[1])

* SASA calculated using GETAREA[2], as explained in detail in **Section 2.2.8**; ‡ pKa calculated using PROPKA[3]

β2m residue	Secondary structure	pKₐ‡	Solvent accessibility*		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)	Peptide (DEPC:peptide = 4:1)	Peptide (DEPC:peptide = 50:1)
K19	Random coil	10.42	77.7	✔	✔ - Salt bridge with side chain COOH of E16 - H-bonding with backbone CO of S20, N21 **- Overall: Charged, (-)**	✘	30 ± 2	N/A	N/A
S20	Random coil	N/A	61.6	✔	✔ - H-bonding with side chain OH of T71 - H-bonding with backbone CO of E69 - H-bonding with side chain COOH of E69 **- Overall: Charged, (-)**	✘	N.D.	2.6 ± 0.6	N.D.
Y26	β-sheet	10.20	26.4	✔	✘ **- Overall: Uncharged, nothing**	✔ - OH close to L65	N.D.	N.D.	N.D.
S28	β-sheet	N/A	18.2	✘	✘ - H-bonding with side chain CONH₂ of Q8 - H-bonding with side chain OH of Y26 - H-bonding with backbone NH of G29 **- Overall: Uncharged, polar contact**	✔ - OH close to Y26	N.D.	N.D.	N.D.
H31	Random coil	5.98	33.5	✔	✘ **- Overall: Uncharged, nothing**	✔ - imidazole amine close to I1	5 ± 3	N/A	N/A
S33	Random coil	N/A	70.1	✔	✘ **- Overall: Uncharged, nothing**	✔ - OH close to P32, F62	1.6 ± 0.4	N.D.	N.D.
K41	Random coil	12.73	12.4	✘	✔ - Salt bridge with side chain COOH of E44, D76 - H-bonding with side chain OH of Y78 - H-bonding with backbone CO of E44 **- Overall: Charged, (-)**	✔ - NH₂ close to I46	1.7 ± 0.8	N/A	N/A

β2m residue	Secondary structure	pK$_a$‡	Solvent accessibility*		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)	Peptide (DEPC:peptide = 4:1)	Peptide (DEPC:peptide = 50:1)
K48	Random coil	10.62	87.0	✔	✔ - Salt bridge with side chain COOH of E47 **- Overall: Charged, (-)**	✘	5 ± 2	N/A	N/A
H51	Random coil	7.15	57.5	✔	✔ - Salt bridge with side chain COOH of D53 - H-bonding with backbone NH of D53 **- Overall: Charged, (-)**	✔ - imidazole amine close to L64	4 ± 1	N/A	N/A
S52	Random coil	N/A	50.7	✔	✘ - H-bonding with side chain OH of Y67 - H-bonding with backbone CO of D53, L65 - H-bonding with backbone NH of D53, L65 **- Overall: Uncharged, polar contact**	✔ - OH close to L65, Y67	N.D.	N.D.	N.D.
S55	Random coil	N/A	55.3	✔	✘ - H-bonding with backbone CO of S57 **- Overall: Uncharged, polar contact**	✔ - OH close to L54	0.6 ± 0.3	N.D.	N.D.
S57	Random coil	N/A	34.9	✔	✔ - H-bonding with side chain COOH of D59 - H-bonding with side chain OH of S61 - H-bonding with backbone CO of F56, S61 - H-bonding with backbone NH of K58, D59 **- Overall: Charged, (-)**	✔ - OH close to F56	1.8 ± 0.9	N.D.	N.D.
K58	Random coil	10.29	88.5	✔	✘ **- Overall: Uncharged, nothing**	✔ - NH$_2$ close to F56	9 ± 2	N/A	N/A

β2m residue	Secondary structure	pK$_a$‡	Solvent accessibility*		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)	Peptide (DEPC:peptide = 4:1)	Peptide (DEPC:peptide = 50:1)
S61	Random coil	N/A	28.1	✔	✔ - H-bonding with side chain guanidinium of R3 - H-bonding with backbone NH of G29, S57, F62 - H-bonding with backbone CO of G29, F56, S57, D58 **- Overall: Charged, (+)**	✔ - OH close to F56, W60, F62, Y63	10 ± 1	N.D.	N.D.
Y63	Random coil	9.54	27.7	✔	✖ - H-bonding with side chain OH of Y26, S55 **- Overall: Uncharged, polar contact**	✔ - OH close to Y26, F56	0.3 ± 0.2	N.D.	N.D.
Y66	β-sheet	13.25	7.5	✖	✖ **- Overall: Uncharged, nothing**	✔ - OH close to V27, F30, I35, V37, L64	0.3 ± 0.2	N.D.	N.D.
Y67	β-sheet	10.46	22.8	✔	✔ - H-bonding with side chain OH of S52 - H-bonding with backbone CO of H51 - H-bonding with side chain COOH of E50 **- Overall: Charged, (-)**	✔ - OH close to L65	N.D.	N.D.	1.3 ± 0.6
T68	β-sheet	N/A	0.1	✖	✔ - H-bonding with backbone CO of K48, backbone NH of E69 - H-bonding with side chain COOH of E69 **- Overall: Charged, (-)**	✔ - OH close to L39, I46, V49, F70	N.D.	N.D.	N.D.
T71	Random coil	N/A	29.8	✔	✖ - H-bonding with side chain OH of S20 - H-bonding with backbone CO of E69 **- Overall: Uncharged, polar contact**	✖	N.D.	N.D.	N.D.

β2m residue	Secondary structure	pKa‡	Solvent accessibility*		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)	Peptide (DEPC:peptide = 4:1)	Peptide (DEPC:peptide = 50:1)
T73	Random coil	N/A	63.9	✔	✔ - H-bonding with side chain COOH of D76 - H-bonding with backbone NH of E74, K75, D76 **- Overall: Charged, (-)**	✘	N.D.	N.D.	N.D.
K75	Random coil	10.47	100	✔	✔ - Salt bridge with side chain COOH of E74, E77 **- Overall: Charged, (-)**	✘	2.3 ± 0.9	N/A	N/A
Y78	Random coil	10.34	5.1	✘	✔ - H-bonding with side chain NH$_2$ of K41 - H-bonding with backbone NH of T71 - H-bonding with backbone CO of T71 **- Overall: Charged, (+)**	✔ - OH close to I46, F70	N.D.	N.D.	N.D.
H84	Random coil	5.78	0.0	✘	✘ - H-bonding with backbone NH of T86 **- Overall: Uncharged, polar contact**	✔ - imidazole amine close to F30, V85	2.0 ± 0.3	N/A	N/A
T86	Random coil	NA	9.2	✘	✘ - H-bonding with backbone NH of L87 - H-bonding with backbone CO of V85 **- Overall: Uncharged, polar contact**	✔ - OH close to V85, L87	N.D.	N.D.	N.D.
S88	Random coil	NA	96.7	✔	✘ - H-bonding with backbone CO of V85 **- Overall: Uncharged, polar contact**	✔ - OH close to L87	N.D.	N.D.	N.D.

191

β2m residue	Secondary structure	pK$_a$‡	Solvent accessibility*		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)	Peptide (DEPC:peptide = 4:1)	Peptide (DEPC:peptide = 50:1)
K91	Random coil	10.36	71.3	✔	✘ - H-bonding with side chain CONH$_2$ of Q89 - H-bonding with backbone CO of I92 **- Overall: Uncharged, polar contact**	✔ - NH$_2$ close to V93	4 ± 3	N/A	N/A
K94	Random coil	10.56	80.3	✔	✔ - Salt bridge with side chain COOH of E77 - H-bonding with backbone CO of W95 **- Overall: Charged, (-)**	✔ - NH$_2$ close to I92, M99	6 ± 2	N/A	N/A

Table A.2: DEPC modification percentages of nucleophilic residues in ubiquitin (Ub) intact protein and its proteolytic peptides. Each experiment was performed in triplicate or quadruplicate (n = 3 or 4). Error bars shown in a table are standard deviations. Listed along are structural features of each residue from PDB 1UBQ. ✔ and ✘ represent presence or absence of each structural element, respectively. (+) and (-) represent positively charged and negatively charged polar contacts, respectively.

Ub residue	Secondary structure	pKa‡	%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)	Peptide (DEPC:peptide = 4:1)	Peptide (DEPC:peptide = 50:1)
			Solvent accessibility*		**Microenvironment**		**% DEPC CL**		
K6	β-sheet	10.37	59.7	✔	✘ - Overall: Uncharged, nothing	✘	0.8 ± 0.6	N/A[b]	N/A
T7	β-sheet	N/A[a]	14.2	✘	✘ - H-bonding with backbone CO of T9, K11 - H-bonding with backbone NH of T9, K11 **- Overall: Uncharged, polar contact**	✔ - OH close to L8	13 ± 2	N.D.[c]	N.D.
T9	Random coil	N/A	85.5	✔	✘ - H-bonding with side chain OH of T7 **- Overall: Uncharged, polar contact**	✘	N.D.	N.D.	N.D.
K11	β-sheet	11.02	59.8	✔	✔ - Salt bridge with side chain COOH of E34 **- Overall: Charged, (-)**	✔ - NH$_2$ close to I13	N.D.	N/A	N/A
T12	β-sheet	N/A	39.4	✔	✘ **- Overall: Uncharged, nothing**	✘	N.D.	N.D.	N.D.
T14	β-sheet	N/A	54.8	✔	✘ **- Overall: Uncharged, nothing**	✘	N.D.	N.D.	N.D.
S20	Random coil	N/A	86.9	✔	✘ **- Overall: Uncharged, nothing**	✔ - OH close to P19	N.D.	N.D.	N.D.

[a] N/A = not applicable; [b] N-terminal blocking using SNHSA also results in side reaction (acetylation) at Lys and His residues in free peptides, so they are not discussed here.

[c] N.D. = not detected (less than DEPC labeling threshold of 0.01%[1])

* SASA calculated using GETAREA[2], as explained in detail in **Section 2.2.8**; ‡ pKa calculated using PROPKA[3]

Ub residue	Secondary structure	pK$_a$‡	Solvent accessibility*		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)	Peptide (DEPC:peptide = 4:1)	Peptide (DEPC:peptide = 50:1)
T22	Random coil	N/A	46.2	✔	✘ - H-bonding with backbone NH of N25 - **Overall: Uncharged, polar contact**	✘	N.D.	N.D.	N.D.
K27	α-helix	10.53	9.8	✘	✔ - Salt bridge with side chain COOH of D52 - H-bonding with backbone CO of Q41 - **Overall: Charged, (-)**	✔ - NH$_2$ close to I23	N.D.	N/A	N/A
K29	α-helix	10.40	37.3	✔	✘ - H-bonding with backbone CO of E16 - **Overall: Uncharged, polar contact**	✔ - NH$_2$ close to V17	N.D.	N/A	N/A
K33	α-helix	10.27	49.4	✔	✘ - H-bonding with backbone CO of T14 - **Overall: Uncharged, polar contact**	✔ - NH$_2$ close to I13, L15	20 ± 6	N/A	N/A
K48	Random coil	10.31	56.1	✔	✘ - H-bonding with backbone CO of A46 - **Overall: Uncharged, polar contact**	✔ - NH$_2$ close to A46	11 ± 4	N/A	N/A
T55	Random coil	N/A	32.3	✔	✘ - H-bonding with side chain OH of S57 - H-bonding with backbone NH of D58 - **Overall: Uncharged, polar contact**	✘	N.D.	N.D.	N.D.
S57	α-helix	N/A	59.6	✔	✘ - H-bonding with side chain OH of T55, backbone CO of P19 - **Overall: Uncharged, polar contact**	✘	N.D.	N.D.	N.D.

Ub residue	Secondary structure	pK$_a$‡	Solvent accessibility*		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)	Peptide (DEPC:peptide = 4:1)	Peptide (DEPC:peptide = 50:1)
Y59	α-helix	9.98	12.9	✘	✘ - H-bonding with backbone NH of E51 **- Overall: Uncharged, polar contact**	✔ - OH close to L50	N.D.	N.D.	N.D.
K63	Random coil	10.66	74.3	✔	✘ **- Overall: Uncharged, nothing**	✘	0.31 ± 0.08	N/A	N/A
S65	β-sheet	N/A	9.2	✘	✘ - H-bonding with backbone CO of Q62 - H-bonding with backbone NH of Q62 **- Overall: Uncharged, polar contact**	✔ - OH close to F45, I61	2.4 ± 0.2	N.D.	N.D.
T66	β-sheet	N/A	40.1	✔	✘ **- Overall: Uncharged, nothing**	✔ - OH close to F4	N.D.	N.D.	N.D.
H68	β-sheet	6.00	49.1	✔	✘ **- Overall: Uncharged, nothing**	✔ - NH close to I44	2.7 ± 0.5	N/A	N/A

195

Table A.3: DEPC modification percentages of nucleophilic residues in hGH intact protein and its proteolytic peptides.

Each experiment was performed in triplicate or quadruplicate (n = 3 or 4). Error bars shown in a table are standard deviations. Listed along are structural features of each residue from PDB 1HGU. ✔ and ✘ represent presence or absence of each structural element, respectively. (+) and (-) represent positively charged and negatively charged polar contacts, respectively.

hGH residue	Secondary structure	pKa‡	%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 5:1)	Peptide (DEPC:peptide = 5:1)	Peptide (DEPC:peptide = 50:1)
T4	Random coil	N/A[a]	100	✔	✘ - Overall: Uncharged, nothing	✔ - OH close to P3	N.D.[c]	N.D.	N.D.
S8	α-helix	N/A	40.3	✔	✘ - Overall: Uncharged, nothing	✔ - OH close to V186	N.D.	N.D.	N.D.
H19	α-helix	4.78	37.2	✔	✘ - Overall: Uncharged, nothing	✔ - imidazole amine close to M15	1.3 ± 0.8 (H19/H22)	N/A[b]	N/A
H22	α-helix	5.16	24.0	✔	✘ - Overall: Uncharged, nothing	✔ - imidazole amine close to F26, M171	N.D.	N/A	N/A
T28	α-helix	N/A	10.7	✘	✔ - H-bonding with side chain COOH of D27 - H-bonding with backbone CO of L24 - Overall: Charged, (-)	✔ - OH close to L24	1.3 ± 0.6	N.D.	N.D.
Y29	α-helix	13.59	4.1	✘	✔ - H-bonding with side chain NH$_2$ of K42 - Overall: Charged, (+)	✔ - OH close to Y161	0.2 ± 0.1	N.D.	N.D.
Y36	α-helix	10.35	37.2	✔	✘ - Overall: Uncharged, nothing	✔ - OH close to F32	0.4 ± 0.3	N.D.	N.D.
K39	**No information from a crystal structure about this residue**						0.06 ± 0.01	N/A	N/A

[a] N/A = not applicable; [b] N-terminal blocking using SNHSA also results in side reaction (acetylation) at Lys and His residues in free peptides, so they are not discussed here.

[c] N.D. = not detected (less than DEPC labeling threshold of 0.01%[1])

* SASA calculated using GETAREA[2], as explained in detail in **Section 2.2.8**; ‡ pKa calculated using PROPKA[3]

hGH residue	Secondary structure	pK$_a$‡	Solvent accessibility*		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 5:1)	Peptide (DEPC:peptide = 5:1)	Peptide (DEPC:peptide = 50:1)
K42	Random coil	11.22	36.2	✔	✘ - H-bonding with side chain OH of Y28 **- Overall: Uncharged, polar contact**	✘	0.07 ± 0.02	N/A	N/A
Y43	Random coil	10.05	87.7	✔	✘ **- Overall: Uncharged, nothing**	✘	N.D.	N.D.	N.D.
S44	Random coil	N/A	37.0	✔	✘ - H-bonding with backbone CO of Q41 - H-bonding with backbone NH of F45 **- Overall: Uncharged, polar contact**	✔ - OH close to F45, L46	N.D.	N.D.	N.D.
T51	No information from a crystal structure about this residue						N.D.	N.D.	N.D.
S52	Random coil	N/A	65.5	✔	✘ - H-bonding with backbone CO, NH of L53 **- Overall: Uncharged, polar contact**	✘	N.D.	N.D.	N.D.
S56	Random coil	N/A	20.8	✔	✘ - H-bonding with backbone CO of L53 **- Overall: Uncharged, polar contact**	✔ - OH close to L53	0.02 ± 0.01	N.D.	N.D.
S58	Random coil	N/A	3.3	✘	✘ - H-bonding with backbone CO, NH of I59 - H-bonding with backbone CO of F55 **- Overall: Uncharged, polar contact**	✔ - OH close to F177	64 ± 4	N.D.	N.D.
T61	Random coil	N/A	38.7	✔	✔ - H-bonding with side chain NH$_2$ of K73 **- Overall: Charged, (+)**	✔ - OH close to P62, F177	N.D.	N.D.	N.D.

hGH residue	Secondary structure	pK$_a$‡	Solvent accessibility*		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 5:1)	Peptide (DEPC:peptide = 5:1)	Peptide (DEPC:peptide = 50:1)
S63	Random coil	N/A	17.6	✘	✘ - H-bonding with backbone CO of P62 **- Overall: Uncharged, polar contact**	✔ - OH close to F177	N.D.	N.D.	N.D.
T68	No information from a crystal structure about this residue						N.D.	N.D.	N.D.
K71	Random coil	10.04	27.1	✔	✘ - H-bonding with backbone CO of Q69 **- Overall: Uncharged, polar contact**	✔ - NH$_2$ close to I59	N.D.	N/A	N/A
S72	Random coil	N/A	32.2	✔	✘ - H-bonding with backbone NH of L74 **- Overall: Uncharged, polar contact**	✔ - OH close to L74	N.D.	N.D.	N.D.
S80	α-helix	N/A	0.5	✘	✘ - H-bonding with backbone CO of L76 **- Overall: Uncharged, polar contact**	✔ - OH close to L83, V174, L178,	N.D.	N.D.	N.D.
S86	α-helix	N/A	0.1	✘	✘ - H-bonding with backbone CO of Y144 **- Overall: Uncharged, polar contact**	✔ - OH close to F55, W87	N.D.	N.D.	N.D.
S96	Random coil	N/A	49.7	✔	✘ **- Overall: Uncharged, nothing**	✘	N.D.	N.D.	N.D.
S101	Random coil	N/A	33.4	✔	✘ - H-bonding with backbone CO of V103 **- Overall: Uncharged, polar contact**	✘	N.D.	N.D.	N.D.
Y104	Random coil	10.05	82.5	✔	✘ **- Overall: Uncharged, nothing**	✘	N.D.	N.D.	N.D.

hGH residue	Secondary structure	pKₐ‡	Solvent accessibility*		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 5:1)	Peptide (DEPC:peptide = 5:1)	Peptide (DEPC:peptide = 50:1)
S107	Random coil	N/A	26.1	✔	✔ - H-bonding with side chain guanidinium of R20 - Overall: Charged, (+)	✔ - OH close to L24	N.D.	N.D.	0.19 ± 0.12
S109	Random coil	N/A	49.3	✔	✔ - H-bonding with sidechain COOH of D106 - Overall: Charged, (-)	✘	N.D.	N.D.	N.D.
Y112	α-helix	10.38	29.9	✔	✘ - Overall: Uncharged, nothing	✘	N.D.	N.D.	N.D.
K116	α-helix	10.63	40.9	✔	✘ - Overall: Uncharged, nothing	✔ - NH₂ close to L88, Y112	N.D.	N/A	N/A
T124	α-helix	N/A	47.2	✔	✘ - Overall: Uncharged, nothing	✔ - OH close to L10	N.D.	N.D.	N.D.
S133	Random coil	N/A	80.2	✔	✘ - H-bonding with backbone CO of G132 - Overall: Uncharged, polar contact	✘	N.D.	N.D.	N.D.
T136	Random coil	N/A	35.9	✔	✔ - H-bonding with side chain guanidinium of R78 - Overall: Charged, (+)	✔ - OH close to L74	0.15 ± 0.09	0.1 ± 0.1	0.08 ± 0.02
K141	Random coil	10.42	90.5	✔	✘ - Overall: Uncharged, nothing	✘	18 ± 8	N/A	N/A
T143	Random coil	N/A	96.4	✔	✘ - Overall: Uncharged, nothing	✔ - OH close to A145	N.D.	N.D.	N.D.
Y144	Random coil	10.41	31.1	✔	✘ - Overall: Uncharged, nothing	✔ - OH close to L82	0.66 ± 0.05	N.D.	0.02 ± 0.03
S145	No information from a crystal structure about this residue						N.D.	N.D.	N.D.

hGH residue	Secondary structure	pK$_a$‡	Solvent accessibility*		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 5:1)	Peptide (DEPC:peptide = 5:1)	Peptide (DEPC:peptide = 50:1)
K146	Random coil	9.4	1.1	✘	✘ - Overall: Uncharged, nothing	✔ - NH$_2$ close to L53, F55	0.09 ± 0.08	N/A	N/A
T149	No information from a crystal structure about this residue						N.D.	N.D.	N.D.
S151	Random coil	N/A	56.7	✔	✘ - H-bonding with backbone CO of N150 - Overall: Uncharged, polar contact	✘	N.D.	N.D.	N.D.
H152	Random coil	6.38	84.4	✔	✘ - Overall: Uncharged, nothing	✘	N.D.	N/A	N/A
K159	α-helix	11.43	70.9	✔	✔ - Salt bridge with sidechain COOH of E34 - Overall: Charged, (-)	✘	N.D.	N/A	N/A
Y161	α-helix	11.14	10.0	✘	✘ - Overall: Uncharged, nothing	✔ - OH close to L46	N.D.	N.D.	N.D.
T165	No information from a crystal structure about this residue						N.D.	N.D.	N.D.
K169	α-helix	10.00	42.8	✔	✘ - Overall: Uncharged, nothing	✘	N.D.	N/A	N/A
K173	α-helix	8.71	1.4	✘	✘ - H-bonding with backbone CO of K169 - Overall: Uncharged, polar contact	✔ - NH$_2$ close to F55	5.0 ± 0.4	N/A	N/A
T176	α-helix	N/A	19.0	✘	✔ - H-bonding with sidechain NH$_2$ of R179 - H-bonding with backbone CO of D172 - Overall: Charged, (+)	✘	N.D.	N.D.	N.D.

hGH residue	Secondary structure	pK$_a$‡	Solvent accessibility*		Microenvironment		% DEPC CL		
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 5:1)	Peptide (DEPC:peptide = 5:1)	Peptide (DEPC:peptide = 50:1)
S185	Random coil	N/A	24.0	✔	✘ - H-bonding with backbone CO of V181 - Overall: Uncharged, polar contact	✔ - OH close to F11	N.D.	N.D.	N.D.
S189	Random coil	N/A	96.1	✔	✘ - Overall: Uncharged, nothing	✘	N.D.	N.D.	N.D.

Table A.4: DEPC modification percentages of nucleophilic residues in bovine carbonic anhydrase (BCA) intact protein.
Each experiment was performed in triplicate or quadruplicate (n = 3 or 4). Error bars shown in a table are standard deviations. Listed along are structural features of each residue from PDB 1V9E. ✔ and ✘ represent presence or absence of each structural element, respectively. (+) and (-) represent positively charged and negatively charged polar contacts, respectively.

BCA residue	Secondary structure	pKa‡	Solvent accessibility*		Microenvironment		% DEPC CL
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 10:1)
S1	Random coil	N/A[a]	81.1	✔	✔ - H-bonding with side chain NH$_2$ of K168 **- Overall: Charged, (+)**	✘	N.D.[b]
H2	Random coil	4.54	11.3	✘	✘	✔ - imidazole amine close to W4	N.D.
H3	Random coil	6.33	81.0	✔	✘	✘	N.D.
Y6	Random coil	13.76	3.4	✘	✘ - H-bonding with backbone CO of N242	✔ - OH close to W15	15.4 ± 0.6
K8	Random coil	10.40	89.9	✔	✘	✘	13 ± 1
H9	Random coil	6.27	84.7	✔	✘	✘	6 ± 1
H14	α-helix	5.95	42.6	✔	✘ - H-bonding with backbone CO of K8	✘	4 ± 1
H16	α-helix	6.07	38.0	✔	✘	✘	3 ± 1
K17	α-helix	10.30	66.1	✔	✘	✘	N.D.
S28	Random coil	N/A	0.0	✘	✘ - H-bonding with side chain OH of Y192 - H-bonding with side chain aromatic amine of W207 - H-bonding with backbone CO and NH of S195	✔ - OH close to Y192, W207	2 ± 1

[a] N/A = not applicable; [b] N.D. = not detected (less than DEPC labeling threshold of 0.01%[1])

* SASA calculated using GETAREA[2], as explained in detail in **Section 2.2.8**; ‡ pKa calculated using PROPKA[3]

BCA residue	Secondary structure	pK$_a$‡	Solvent accessibility*		Microenvironment		% DEPC CL
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 10:1)
T34	α-helix	N/A	40.7	✔	✔ - H-bonding with side chain COOH of D109 - **Overall: Charged, (-)**	✔ - OH close to V37, V108	N.D.
K35	α-helix	10.67	100.0	✔	✗	✗	N.D.
K44	Random coil	10.65	66.3	✔	✗	✗	N.D.
Y50	Random coil	14.02	4.8	✗	✗ - H-bonding with side chain aromatic amine of H121	✔ - OH close to P179	0.2 ± 0.2
T54	Random coil	N/A	36.2	✔	✗ - H-bonding with backbone CO of E52	✔ - OH close to A53	0.6 ± 0.4
S55	β-sheet	N/A	0.0	✗	✗	✔ - OH close to T54, Y69, P175	0.02 ± 0.02
H63	Random coil	1.74	17.3	✗	✗	✔ - imidazole amine close to W4, Y6	23 ± 8
S64	Random coil	N/A	0.2	✗	✗ - H-bonding with side chain CONH$_2$ of N61	✔ - OH close to W96	0.9 ± 0.5
Y69	β-sheet	14.67	0.2	✗	✗ - H-bonding with backbone CO of F177	✔ - OH close to A53, V67, F92, F177	0.52 ± 0.06
S72	Random coil	N/A	74.8	✔	✗	✗	0.23 ± 0.02
K75	Random coil	10.53	32.3	✔	✔ - H-bonding with side chain COOH of D70	✗	0.22 ± 0.01
K79	β-sheet	10.49	48.4	✔	✗	✗	3.4 ± 0.3
T84	Random coil	N/A	100.0	✔	✗	✗	N.D.
T86	Random coil	N/A	26.7	✔	✗	✔ - OH close to V77	0.008 ± 0.002
Y87	β-sheet	12.20	2.1	✗	✗ - H-bonding with side chain CONH$_2$ of N123	✔ - OH close to L83	0.02 ± 0.01

BCA residue	Secondary structure	pK$_a$‡	%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 10:1)
H93	β-sheet	-2.62	12.9	✖	✖	✖	N.D.
H95	β-sheet	-1.24	0.1	✖	✖	✖	0.2 ± 0.2
S98	Random coil	N/A	48.5	✔	✖	✔ - OH close to L240, A241	0.02 ± 0.01
S99	Random coil	N/A	59.9	✔	✔ - H-bonding with side chain COOH of D101 - Overall: Charged, (-)	✖	N.D.
S104	Random coil	N/A	0.0	✖	✖ - H-bonding with backbone CO of H106 - H-bonding with backbone NH of Y113	✖	N.D.
H106	Random coil	2.72	0.0	✖	✖	✔ - imidazole amine close to P29, V30, W207, V209	0.1 ± 0.1
T107	β-sheet	N/A	4.4	✖	✔ - H-bonding with side chain COOH of D31 - Overall: Charged, (-)	✔ - OH close to A246	0.1 ± 0.1
K111	Random coil	9.40	53.6	✔	✖	✔ - NH$_2$ close to Y113	3.7 ± 0.9
K112	Random coil	10.09	44.2	✔	✖	✔ - NH$_2$ close to W243	0.2 ± 0.2
Y113	Random coil	11.30	7.7	✖	✖	✔ - OH close to P213	N.D.
H118	β-sheet	-5.13	1.4	✖	✖	✔ - imidazole amine close to V141	0.17 ± 0.06
H121	β-sheet	3.50	0.0	✖	✖	✔ - imidazole amine close to L119	N.D.
T124	Random coil	N/A	23.7	✔	✖ - H-bonding with backbone CO of T86	✖	N.D.
K125	Random coil	12.41	55.5	✔	✔ - H-bonding with side chain COOH of D137 - H-bonding with side chain OH of Y126 - Overall: Charged, (-)	✔ - NH$_2$ close to Y126	N.D.

BCA residue	Secondary structure	pK$_a$‡	Solvent accessibility*		Microenvironment		% DEPC CL
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 10:1)
Y126	α-helix	10.91	16.9	✗	✓ - H-bonding with side chain COOH of D137 - H-bonding with side chain NH$_2$ of K125	✗	N.D.
T131	α-helix	N/A	56.1	✓	✓ - H-bonding with side chain COOH of D128 **- Overall: Charged, (-)**	✗	N.D.
K147	β-sheet	10.31	59.0	✓	✗	✓ - NH$_2$ close to Y113, P213	4 ± 1
K157	α-helix	10.06	42.7	✓	✗	✓ - NH$_2$ close to F174, P175, F177	0.9 ± 0.3
S164	α-helix	N/A	43.4	✓	✗	✓ - OH close to A161	N.D.
K166	Random coil	10.47	29.3	✓	✗ - H-bonding with backbone CO of T86	✗	N.D.
T167	Random coil	N/A	21.2	✓	✗	✓ - OH close to I165, A231	N.D.
K168	Random coil	10.31	27.0	✓	✗ - H-bonding with side chain OH of S1	✓ - NH$_2$ close to A231	N.D.
K170	Random coil	10.45	53.7	✓	✗	✗	N.D.
S171	β-sheet	N/A	30.3	✓	✗	✗	N.D.
T172	β-sheet	N/A	32.7	✓	✗	✓ - OH close to M58, I165	N.D.
S181	α-helix	N/A	30.0	✓	✗	✓ - OH close to A154, F177, L182	N.D.
Y189	β-sheet	12.63	0.0	✗	✗ - H-bonding with backbone CO of K44	✓ - OH close to L43, L46, P82	N.D.
T191	β-sheet	N/A	0.4	✗	✗ - H-bonding with backbone CO of W207	✓ - OH close to P82, Y189, W207	N.D.
Y192	β-sheet	14.03	0.0	✗	✗ - H-bonding with side chain OH of S28	✓ - OH close to V30	N.D.
S195	Random coil	N/A	0.0	✗	✗ - H-bonding with backbone CO of L196 - H-bonding with backbone NH of S28	✗	N.D.

BCA residue	Secondary structure	pK_a‡	Solvent accessibility*		Microenvironment		% DEPC CL
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 10:1)
T197	Random coil	N/A	2.2	✘	✘	✘	0.03 ± 0.02
T198	Random coil	N/A	18.1	✘	✘ - H-bonding with backbone CO of P199	✔ - OH close to P199	0.014 ± 0.003
S204	Random coil	N/A	2.1	✘	✘ - H-bonding with backbone CO of A133	✔ - OH close to L202, V205	N.D.
T206	β-sheet	N/A	23.1	✔	✘	✔ - OH close to A140	N.D.
K211	Random coil	10.74	23.0	✔	✘ - H-bonding with backbone CO of K259	✔ - NH$_2$ close to W190	0.005 ± 0.003
S215	β-sheet	N/A	43.2	✔	✘ - H-bonding with backbone CO of D150	✔ - OH close to V216	N.D.
S217	α-helix	N/A	9.8	✘	✘ - H-bonding with sidechain CONH$_2$ of Q219 - H-bonding with backbone NH of Q220	✔ - OH close to V216	0.01 ± 0.01
S218	α-helix	N/A	54.7	✔	✔ - H-bonding with side chain COOH of D101 **- Overall: Charged, (-)**	✔ - OH close to V148	N.D.
K223	α-helix	10.46	32.4	✔	✔ - H-bonding with side chain COOH of D163 **- Overall: Charged, (-)**	✔ - NH$_2$ close to L162	N.D.
T226	Random coil	N/A	34.5	✔	✘	✘	N.D.
K250	Random coil	10.72	29.2	✔	✔ - H-bonding with side chain COOH of E25 - H-bonding with backbone CO of N23, G24 **- Overall: Charged, (-)**	✔ - NH$_2$ close to P248	14 ± 4
K259	Random coil	10.47	92.7	✔	✘	✘	66 ± 12

Table A.5: DEPC modification percentages of nucleophilic residues in equine holo-myoglobin (MYG) intact protein.
Each experiment was performed in triplicate or quadruplicate (n = 3 or 4). Error bars shown in a table are standard deviations. Listed along are structural features of each residue from PDB 1DWR. ✔ and ✗ represent presence or absence of each structural element, respectively. (+) and (-) represent positively charged and negatively charged polar contacts, respectively.

MYG residue	Secondary structure	pKa‡	Solvent accessibility*		Microenvironment		% DEPC CL
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)
S3	α-helix	N/A[a]	63.8	✔	✗ - H-bonding with backbone NH of E6	✔ - OH close to L2	39 ± 19 (N-term/S3)
K16	α-helix	10.98	22.1	✔	✔ - H-bonding with side chain COOH of D109 **- Overall: Charged, (-)**	✗	3 ± 2
H24	α-helix	5.75	2.7	✗	✗	✔ - imidazole amine close to V17, L115	0.01 ± 0.02
T34	α-helix	N/A	33.9	✔	✗ - H-bonding with backbone CO of I30, R31	✔ - OH close to I30	1.1 ± 0.5
H36	α-helix	5.90	32.6	✔	✗	✔ - imidazole amine close to L32, F106	1.1 ± 0.5
K42	Random coil	9.80	33.0	✔	✗ - H-bonding with backbone CO of K98	✔ - NH$_2$ close to P100	0.7 ± 0.4
K45	Random coil	10.74	54.9	✔	✔ - H-bonding with side chain COOH of D60 **- Overall: Charged, (-)**	✔ - NH$_2$ close to F46	0.8 ± 0.7
K47	Random coil	**Information from a crystal structure about this residue is not clear**					N.D.[b]
H48	Random coil	6.21	64.7	✔	✗	✔ - imidazole amine close to F46, L49	N.D.
K50	Random coil	10.46	83.4	✔	✗ - H-bonding with backbone CO of K47	✗	0.3 ± 0.3

[a] N/A = not applicable; [b] N.D. = not detected (less than DEPC labeling threshold of 0.01%[1])
* SASA calculated using GETAREA[2], as explained in detail in **Section 2.2.8**; ‡ pKa calculated using PROPKA[3]

MYG residue	Secondary structure	pK_a‡	Solvent accessibility*		Microenvironment		% DEPC CL
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)
T51	α-helix	N/A	61.8	✔	✔ - H-bonding with side chain COOH of E54 - H-bonding with backbone NH of E54 **- Overall: Charged, (-)**	✔ - OH close to A53	0.1 ± 0.1
K56	α-helix	10.19	58.3	✔	✖	✔ - NH$_2$ close to I30	0.1 ± 0.1
S58	α-helix	N/A	9.1	✖	✖ - H-bonding with backbone NH of L61	✔ - OH close to L61	0.2 ± 0.3
K62	α-helix	10.65	48.0	✔	✖	✖	0.8 ± 0.5
K63	α-helix	10.44	75.3	✔	✖	✖	0.1 ± 0.1
H64	α-helix	3.84	26.6	✔	✖	✔ - imidazole amine close to F46, L49	0.02 ± 0.02
T66	α-helix	N/A	40.2	✔	✖ - H-bonding with backbone CO of K62	✔ - OH close to I21, A22	0.2 ± 0.1
T70	α-helix	N/A	71.8	✔	✖ - H-bonding with backbone CO of T66	✔ - OH close to I21	N.D.
K77	α-helix	11.06	50.5	✔	✔ - H-bonding with side chain COOH of E18 **- Overall: Charged, (-)**	✔ - NH$_2$ close to L11, W14	7 ± 3
K78	Random coil	11.35	46.2	✔	✔ - H-bonding with side chain COOH of E85 - H-bonding with backbone CO of H81 **- Overall: Charged, (-)**	✖	6 ± 3
K79	Random coil	10.47	53.6	✔	✖	✖	3.1 ± 0.4
H81	Random coil	6.47	89.1	✔	✖	✖	1.0 ± 0.5
H82	α-helix	6.42	4.9	✖	✔ - H-bonding with side chain COOH of D141 **- Overall: Charged, (-)**	✔ - imidazole amine close to L86	1.0 ± 0.6
K87	α-helix	10.48	52.0	✔	✖	✖	1.0 ± 0.5
S92	α-helix	**Information from a crystal structure about this residue is not clear**					1 ± 1

MYG residue	Secondary structure	pKa‡	Solvent accessibility*		Microenvironment		% DEPC CL
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)
H93	α-helix	2.32	35.3	✔	✘ - H-bonding with side chain OH of S92 - Coordination with Fe(II) of heme	✔ - imidazole amine close to L89, I99, L104	0.11 ± 0.08
T95	α-helix	N/A	76.2	✔	✘	✔ - OH close to L149	0.3 ± 0.2
K96	α-helix	10.38	84.2	✔	✘	✘	1.2 ± 0.9
H97	Random coil	4.46	43.3	✔	✔ - H-bonding with COOH of heme **- Overall: Charged, (-)**	✔ - imidazole amine close to hydrocarbon chain of heme	0.2 ± 0.2
K98	Random coil	10.23	65.0	✔	✘ - H-bonding with side chain OH of T95	✘	0.8 ± 0.4
K102	α-helix	11.14	63.6	✔	✔ - H-bonding with side chain COOH of E105 **- Overall: Charged, (-)**	✔ - NH_2 close to F106	0.17 ± 0.09
Y103	α-helix	11.10	15.2	✘	✘	✔ - OH close to P100	0.24 ± 0.08
S108	α-helix	N/A	5.4	✘	✔ - H-bonding with side chain guanidinium of R139 - H-bonding with backbone CO of L104 **- Overall: Charged, (+)**	✔ - OH close to I111	0.16 ± 0.05
H113	α-helix	5.52	43.2	✔	✔ - H-bonding with side chain guanidinium of R31 **- Overall: Charged, (+)**	✘	2.7 ± 0.8
H116	α-helix	6.02	43.9	✔	✘	✔ - imidazole amine close to I112	0.9 ± 0.5
S117	α-helix	N/A	51.3	✔	✘	✘	8 ± 4
K118	α-helix	11.44	23.5	✔	✔ - H-bonding with side chain COOH of E27 **- Overall: Charged, (-)**	✔ - NH_2 close to V114	1.7 ± 0.7
H119	α-helix	5.26	20.0	✔	✘	✔ - imidazole amine close to L115	3 ± 3
T132	α-helix	N/A	45.8	✔	✘ - H-bonding with backbone CO of Q128	✔ - OH close to I112, M131, L135	N.D.

MYG residue	Secondary structure	pK$_a$‡	Solvent accessibility*		Microenvironment		% DEPC CL
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)
K133	α-helix	11.12	44.8	✔	✔ - H-bonding with side chain COOH of E6 - H-bonding with backbone CO of G1 - **Overall: Charged, (-)**	✘	N.D.
K145	α-helix	11.64	16.4	✘	✔ - H-bonding with side chain COOH of E83 - **Overall: Charged, (-)**	✔ - NH$_2$ close to L86	2 ± 1
Y146	α-helix	12.86	4.2	✘	✘ - H-bonding with backbone CO of I99	✔ - OH close to I142, F151	0.5 ± 0.2
K147	α-helix	10.45	97.0	✔	✘	✘	3.1 ± 0.7

Table A.6: DEPC modification percentages of nucleophilic residues in hen egg-white lysozyme (LYZ) intact protein.
Each experiment was performed in triplicate or quadruplicate (n = 3 or 4). Error bars shown in a table are standard deviations. Listed along are structural features of each residue from PDB 1DWR. ✔ and ✘ represent presence or absence of each structural element, respectively. (+) and (-) represent positively charged and negatively charged polar contacts, respectively.

LYZ residue	Secondary structure	pKa[‡]	Solvent accessibility*		Microenvironment		% DEPC CL
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)
K19	β-sheet	11.25	44.6	✔	✔ - H-bonding with side chain COOH of E25 **- Overall: Charged, (-)**	✔ - NH$_2$ close to F21	63 ± 10
K31	α-helix	11.60	42.1	✔	✔ - H-bonding with COOH of C-terminus **- Overall: Charged, (-)**	✔ - NH$_2$ close to L147	10 ± 3
H33	α-helix	6.42	19.7	✘	✘ - H-bonding with side chain OH of T89	✔ - imidazole amine close to I106, V110	0.17 ± 0.09
Y38	Random coil	9.24	32.4	✔	✘ - H-bonding with side chain OH of S118	✔ - OH close to V117	N.D.[b]
Y41	Random coil	10.10	23.5	✔	✘ - H-bonding with backbone NH of M123	✔ - OH close to V117, M123	0.04 ± 0.02
S42	Random coil	N/A[a]	41.2	✔	✘ - H-bonding with backbone NH of G26	✔ - OH close to L43	0.06 ± 0.03
K51	α-helix	10.15	33.7	✔	✘ - H-bonding with side chain CONH$_2$ of N55	✔ - NH$_2$ close to F56	9 ± 5
S54	Random coil	N/A	0.0	✘	✘ - H-bonding with backbone NH of I73	✔ - OH close to I73	N.D.
T58	Random coil	N/A	0.8	✘	✘	✔ - OH close to I73	N.D.
T61	β-sheet	N/A	44.2	✔	✘	✘	N.D.

[a] N/A = not applicable; [b] N.D. = not detected (less than DEPC labeling threshold of 0.01%[1])

* SASA calculated using GETAREA[2], as explained in detail in **Section 2.2.8**; [‡] pKa calculated using PROPKA[3]

LYZ residue	Secondary structure	pK$_a$‡	%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)
T65	Random coil	N/A	100	✔	✘	✘	N.D.
S68	β-sheet	N/A	1.6	✘	✘ - H-bonding with side chain CONH$_2$ of N64	✘	N.D.
T69	β-sheet	N/A	5.4	✘	✘	✔ - OH close to Y71	N.D.
Y71	β-sheet	11.87	9.8	✘	✘	✘	N.D.
S78	Random coil	N/A	23.5	✔	✘ - H-bonding with backbone NH of M123	✔ - OH close to V117, M123	0.04 ± 0.02
T87	Random coil	N/A	0.0	✘	✘ - H-bonding with side chain OH of T89 - H-bonding with backbone CO, NH of T69	✔ - OH close to Y71	0.010 ± 0.005
S90	Random coil	N/A	2.1	✘	✔ - H-bonding with side chain COOH of E84 - Overall: Charged, (-)	✔ - OH close to Y88	0.31 ± 0.08
S99	α-helix	N/A	15.9	✘	✘ - H-bonding with backbone CO of S78, T87	✘	0.04 ± 0.02
S103	Random coil	N/A	85.5	✔	✘ - H-bonding with side chain NH of A100	✔ - OH close to P97	0.19 ± 0.04
S104	Random coil	N/A	53.7	✔	✘ - H-bonding with side chain OH of T89 - H-bonding with backbone CO, NH of D105	✘	0.03 ± 0.01
T107	α-helix	N/A	77.0	✔	✘	✘	0.03 ± 0.01
S109	α-helix	N/A	53.2	✔	✔ - H-bonding with side chain COOH of D105 - Overall: Charged, (-)	✔ - OH close to I106	N.D.
K114	α-helix	10.15	2.0	✘	✘ - H-bonding with backbone CO of L101	✔ - OH close to I78, L101, I106	0.15 ± 0.02
K115	Random coil	10.45	27.3	✔	✘ - H-bonding with backbone CO of H33	✔ - NH$_2$ close to V110	5.7 ± 0.5

LYZ residue	Secondary structure	pK$_a$‡	Solvent accessibility*		Microenvironment		% DEPC CL
			%SASA ratio	> 20%	Charged polar contact [within 4 Å]	Hydrophobic neighbor [within 6 Å]	Intact protein (DEPC:protein = 4:1)
S118	Random coil	N/A	27.6	✔	✘	✔ - OH close to V117	0.21 ± 0.06
K134	Random coil	10.17	51.7	✔	✘ - H-bonding with side chain CONH$_2$ of N124	✔ - NH$_2$ close to W129	0.6 ± 0.1
T136	Random coil	N/A	50.4	✔	✘ - H-bonding with backbone CO of C133	✔ - OH close to V138	20 ± 5

References

1.   Limpikirati, P.; Hale, J. E.; Hazelbaker, M.; Huang, Y.; Jia, Z.; Yazdani, M.; Graban, E. M.; Vaughan, R. C.; Vachet, R. W., Covalent labeling and mass spectrometry reveal subtle higher order structural changes for antibody therapeutics. *mAbs* **2019,** *11* (3), 463-476.
2.   Fraczkiewicz, R.; Braun, W., Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J. Comput. Chem.* **1998,** *19* (3), 319-333.
3.   Li, H.; Robertson, A. D.; Jensen, J. H., Very fast empirical prediction and rationalization of protein pKa values. *Proteins: Struct., Funct., Bioinf.* **2005,** *61* (4), 704-721.

DEPC MODIFICATION PERCENTAGES FOR INDIVIDUAL RESIDUES OF

RITUXIMAB UNDER NATIVE AND STRESSED CONDITIONS

Table B.1: DEPC modification percentages for individual residues of rituximab under native conditions and after heating to 45 °C for 4 h.

A difference was considered significant if the p-value, calculated by performing an unpaired T-test, was less than 0.05 (corresponding to a 95% confidence level, n = 4). Error bars shown in a table are standard deviations. Color coding of the cells (introduced below) indicates the extent of covalent labeling (CL) change which falls within low (L), medium (M), and high (H) bins.

	High	Medium	Low	Low	Medium	High
Absolute difference (His, Lys)	≥ 20%	5% to 20%	≤ 5%	≥ -5%	-5% to -20%	≤ -20%
Ratio difference (Ser, Thr, Tyr)	≥ 70%	40% to 70%	≤ 40%	≥ -40%	-40% to -70%	≤ -70%
	Increased labeling				Decreased labeling	

Residues	Native	Stressed 45 °C 4 h	Significant?	Absolute difference (H, K)	Ratio difference (S, T, Y)
Light chain - V$_L$ domain					
S 5	0.03% ± 0.02%	0.03% ± 0.01%	No		
S 7					
S 12	0.07% ± 0.02%	0.05% ± 0.03%	No		
S 14	1.5% ± 0.7%	0.63% ± 0.30%	No		
K18	0.17% ± 0.11%	0.09% ± 0.09%	No		
T 20					
T 22					
S 26					
S 27	0.19% ± 0.10%	0.24% ± 0.29%	No		
S 28	0.04% ± 0.04%	0.10% ± 0.10%	No		
S 30					
Y 31	0.22% ± 0.13%	0.31% ± 0.32%	No		
H 33	0.51% ± 0.42%	0.09% ± 0.10%	No		
K 38	0.10% ± 0.06%	0.15% ± 0.08%	No		
S 41	0.01% ± 0.02%	0.01% ± 0.01%	No		
S 42	0.03% ± 0.02%	0.01% ± 0.01%	No		
K 44	0.13% ± 0.06%	0.13% ± 0.05%	No		
Y 48	0.05% ± 0.03%	0.07% ± 0.04%	No		
T 50	0.11% ± 0.05%	0.09% ± 0.07%	No		
S 51	0.05% ± 0.03%	0.03% ± 0.02%	No		
S 55	0.02% ± 0.02%	0.01% ± 0.00%	No		
S 62	3.6% ± 2.6%	1.8% ± 1.4%	No		
S 64	2.1% ± 1.8%	1.2% ± 1.0%	No		
S 66	1.3% ± 1.2%	0.79% ± 0.32%	No		
T 68	1.3% ± 1.2%	0.79% ± 0.32%	No		
S 69	0.02% ± 0.02%	0.02% ± 0.01%	No		
Y 70	0.02% ± 0.02%	0.02% ± 0.01%	No		
S 71	50% ± 7%	30% ± 19%	No		
T 73	7.5% ± 4.8%	5.3% ± 4.4%	No		
S 75	7.6% ± 4.9%	5.5% ± 4.6%	No		
T 84	0.08% ± 0.04%	0.07% ± 0.04%	No		
Y 85					
Y 86					
T 91					
S 92	1.7% ± 0.8%	3.7% ± 3.7%	No		

Residues	Native	Stressed 45 °C 4 h	Significant?	Absolute difference (H, K)	Ratio difference (S, T, Y)
T 96	1.9% ± 0.8%	3.9% ± 3.7%	No		
T 101					
K 102	0.00% ± 0.00%	0.00% ± 0.00%	No		
K 106					
Light chain - C_L domain					
T 108					
S 113	0.02% ± 0.01%	0.14% ± 0.11%	No		
S 120					
K 125					
S 126	0.06% ± 0.10%	0.02% ± 0.01%	No		
T 128					
S 130					
Y 139					
K 144					
K 148	1.3% ± 1.5%	1.3% ± 0.8%	No		
S 155	4.5% ± 1.5%	9.3% ± 3.6%	Yes	+4.9% ± 3.9%	+109% ± 95%
S 158	4.5% ± 1.5%	9.3% ± 3.6%	Yes	+4.9% ± 3.9%	+109% ± 95%
S 161	4.5% ± 1.5%	9.3% ± 3.6%	Yes	+4.9% ± 3.9%	+109% ± 95%
T 163	0.61% ± 0.55%	0.89% ± 0.51%	No		
S 167	0.64% ± 0.58%	0.99% ± 0.56%	No		
K 168					
S 170					
T 171					
Y 172					
S 173					
S 175					
S 176					
T 177					
T 179	0.53% ± 0.96%	0.06% ± 0.06%	No		
S 181	0.97% ± 0.52%	0.31% ± 0.06%	Yes	-0.67% ± 0.53%	-68% ± 65%
K 182	16% ± 3%	8.0% ± 5.3%	No		
Y 185					
K 187	90% ± 4%	92% ± 4%	No		
H 188	78% ± 9%	87% ± 10%	No		
K 189	77% ± 9%	86% ± 9%	No		
Y 191	3.4% ± 1.3%	6.1% ± 4.8%	No		
T 196	3.4% ± 1.5%	4.9% ± 1.0%	No		
H 197	3.7% ± 1.6%	5.3% ± 0.9%	No		
S 201					
S 202					
T 205					
K 206	0.92% ± 0.39%	1.3% ± 0.6%	No		
S 207					
Heavy chain – V_H domain					
K 13	2.5% ± 1.7%	2.7% ± 1.2%	No		
S 17	2.4% ± 1.6%	2.6% ± 1.2%	No		
K 19	0.20% ± 0.06%	0.22% ± 0.07%	No		
S 21					
K 23	0.82% ± 0.70%	2.5% ± 2.8%	No		
S 25					
Y 27	43% ± 1%	44% ± 4%	No		
T 28	41% ± 2%	44% ± 3%	No		
T 30	21% ± 8%	17% ± 8%	No		
S 31	0.30% ± 0.35%	0.52% ± 0.34%	No		
Y 32	0.30% ± 0.35%	0.52% ± 0.34%	No		
H 35	90% ± 5%	84% ± 7%	No		
K 38	56% ± 17%	57% ± 13%	No		
T 40					
Y 52	0.35% ± 0.11%	0.26% ± 0.17%	No		
T 58					
S 59					
Y 60					

Residues	Native	Stressed 45 °C 4 h	Significant?	Absolute difference (H, K)	Ratio difference (S, T, Y)
K 63	1.64% ± 0.64%	1.11% ± 0.58%	No		
K 65					
K 67	80% ± 11%	82% ± 15%	No		
T 69	4.0% ± 2.2%	25% ± 18%	No		
T 71	26% ± 5%	39% ± 21%	No		
K 74	32% ± 5%	36% ± 17%	No		
S 75					
S 76	26% ± 7%	6.5% ± 4.2%	Yes	-20% ± 8%	-76% ± 38%
S 77	14% ± 4%	2.3% ± 1.6%	Yes	-12% ± 4%	-84% ± 38%
T 78	0.55% ± 0.54%	0.45% ± 0.42%	No		
Y 80	1.1% ± 1.0%	1.4% ± 0.6%	No		
S 84	2.5% ± 0.4%	3.4% ± 5.8%	No		
S 85					
T 87					
S 88	3.6% ± 0.8%	4.4% ± 1.6%	No		
S 91					
Y 94	0.13% ± 0.08%	0.22% ± 0.12%	No		
Y 95	2.8% ± 1.7%	3.4% ± 1.7%	No		
S 99					
T 100					
Y 101					
Y 102					
Y 107					
T 115	0.11% ± 0.17%	0.57% ± 0.91%	No		
T 116	0.11% ± 0.17%	0.57% ± 0.91%	No		
T 118					
S 120	7.9% ± 4.4%	2.7% ± 0.9%	No		
Heavy chain – C$_H$1 domain					
S 123	9.5% ± 2.0%	5.6% ± 5.1%	No		
T 124	14% ± 4%	7.5% ± 6.0%	No		
K 125	42% ± 12%	25% ± 15%	No		
S 128	0.01% ± 0.01%	0.05% ± 0.03%	No		
S 135	0.05% ± 0.02%	0.25% ± 0.17%	No		
S 136	0.04% ± 0.02%	0.03% ± 0.02%	No		
K 137	0.05% ± 0.02%	0.22% ± 0.12%	Yes	+0.17% ± 0.13	
S 138	1.4% ± 0.9%	1.3% ± 0.7%	No		
T 139	0.03% ± 0.02%	0.02% ± 0.02%	No		
S 140	0.33% ± 0.19%	0.06% ± 0.03%	Yes	-0.26% ± 0.20%	-81% ± 77%
T 143	0.30% ± 0.18%	0.07% ± 0.04%	Yes	-0.24% ± 0.19%	-78% ± 78%
K 151					
Y 153	0.01% ± 0.00%	0.01% ± 0.01%	No		
T 159	4.6% ± 2.7%	2.8% ± 1.2%	No		
S 161	1.8% ± 0.5%	0.7% ± 0.8%	No		
S 164	14% ± 4%	8.2% ± 4.6%	No		
T 168	7.9% ± 3.8%	4.9% ± 3.2%	No		
S 169	0.38% ± 0.59%	0.53% ± 0.31%	No		
H 172	35% ± 4%	35% ± 8%	No		
T 173					
S 180	7.5% ± 1.6%	9.1% ± 10.6%	No		
S 181	7.3% ± 1.7%	8.9% ± 10.6%	No		
Y 184					
S 185					
S 187	0.86% ± 0.93%	1.2% ± 1.2%	No		
S 188	1.4% ± 1.1%	1.4% ± 1.3%	No		
T 191	1.5% ± 0.8%	0.39% ± 0.52%	No		
S 194	0.04% ± 0.02%	0.17% ± 0.10%	Yes	+0.13% ± 0.10%	+313% ± 281%
S 195	1.4% ± 0.7%	0.8% ± 0.5%	No		
S 196	1.6% ± 1.1%	0.95% ± 0.47%	No		
T 199					
T 201	0.55% ± 0.36%	2.4% ± 3.7%	No		
Y 202	0.55% ± 0.36%	2.4% ± 3.7%	No		

Residues	Native	Stressed 45 °C 4 h	Significant?	Absolute difference (H, K)	Ratio difference (S, T, Y)
H 208	1.5% ± 0.2%	1.0% ± 0.7%	No		
K 209	7.8% ± 3.1%	6.4% ± 2.9%	No		
S 211	0.00% ± 0.00%	0.00% ± 0.00%	No		
T 213					
K 214	0.30% ± 0.26%	0.20% ± 0.18%	No		
K 217					
K 218					
K 222	100% ± 0%	100% ± 0%	No		
S 223	0.80% ± 0.56%	1.3% ± 0.8%	No		
Heavy chain – Hinge region and C$_H$2 domain					
K 226	39% ± 12%	73% ± 20%	Yes	+34% ± 23%	
T 227	0.07% ± 0.07%	0.14% ± 0.12%	No		
H 228	16% ± 8%	14% ± 3%	No		
T 229	16% ± 8%	14% ± 3%	No		
S 243	0.06% ± 0.05%	0.04% ± 0.03%	No		
K 250					
K 252	100% ± 0%	100% ± 0%	No		
T 254					
S 258					
T 260					
T 264	0.04% ± 0.03%	0.03% ± 0.02%	No		
S 271	5.1% ± 2.5%	4.2% ± 1.9%	No		
H 272	5.1% ± 2.4%	4.2% ± 1.9%	No		
K 278	0.13% ± 0.07%	0.16% ± 0.07%	No		
Y 282	0.08% ± 0.07%	0.06% ± 0.02%	No		
H 289	0.06% ± 0.05%	0.04% ± 0.01%	No		
K 292	0.01% ± 0.01%	0.01% ± 0.01%	No		
T 293	7.4% ± 2.8%	23% ± 12%	Yes	+16% ± 12%	+220% ± 188%
K 294	7.4% ± 2.8%	23% ± 12%	Yes	+16% ± 12%	
Y 300					
S 302	83% ± 4%	75% ± 17%	No		
T 303	83% ± 4%	75% ± 17%	No		
Y 304					
S 308					
T 311	1.6% ± 0.5%	1.2% ± 0.7%	No		
H 314	0.29% ± 0.10%	0.25% ± 0.20%	No		
K 321	4.7% ± 1.6%	3.3% ± 1.7%	No		
Y 323					
K 324	94% ± 5%	92% ± 3%	No		
K 326					
S 328					
K 330	100% ± 0%	100% ± 0%	No		
K 338	0.53% ± 0.14%	0.48% ± 0.28%	No		
T 339					
S 341	27% ± 3%	22% ± 8%	No		
K 342					
K 344					
Heavy chain – C$_H$3 domain					
Y 353	27% ± 16%	39% ± 25%	No		
T 354	0.01% ± 0.00%	0.01% ± 0.01%	No		
S 358	0.11% ± 0.05%	0.17% ± 0.08%	No		
T 363	0.12% ± 0.06%	0.19% ± 0.08%	No		
K 364	0.11% ± 0.05%	0.17% ± 0.09%	No		
S 368	0.03% ± 0.03%	0.04% ± 0.03%	No		
T 370	7.5% ± 1.3%	14% ± 4%	Yes	+6.5% ± 3.9%	+87% ± 54%
K 374	7.5% ± 1.3%	14% ± 4%	Yes	+6.5% ± 3.9%	
Y 377	0.63% ± 0.33%	0.65% ± 0.22%	No		
S 379					
S 387					
Y 395					
K 396	0.20% ± 0.10%	0.05% ± 0.02%	Yes	-0.15% ± 0.10%	

Residues	Native	Stressed 45 °C 4 h	Significant?	Absolute difference (H, K)	Ratio difference (S, T, Y)
T 397					
T 398	9.6% ± 4.0%	6.4% ± 2.8%	No		
S 404	9.0% ± 4.0%	5.7% ± 2.9%	No		
S 407	1.6% ± 0.4%	0.9% ± 1.1%	No		
Y 411					
S 412					
K 413					
T 415					
K 418					
S 419					
S 428	0.13% ± 0.06%	0.14% ± 0.10%	No		
S 430	6.0% ± 0.8%	5.2% ± 1.0%	No		
H 433	13% ± 3%	15% ± 2%	No		
H 437	14% ± 2%	11% ± 2%	No		
H 439	1.4% ± 0.6%	1.4% ± 0.3%	No		
Y 440	0.00% ± 0.00%	0.05% ± 0.04%	No		
T 441					
K 443	1.1% ± 0.6%	1.3% ± 0.7%	No		
S 444					
S 446					
S 448	17% ± 2%	26% ± 4%	Yes	+9.1% ± 5.0%	+54% ± 31%
K 451					

Table B.2: DEPC modification percentages for individual residues of rituximab under native conditions and after heating to 55 °C for 4 h.

A difference was considered significant if the p-value, calculated by performing an unpaired T-test, was less than 0.05 (corresponding to a 95% confidence level, n = 4). Error bars shown in a table are standard deviations. Color coding of the cells (introduced below) indicates the extent of covalent labeling (CL) change which falls within low (L), medium (M), and high (H) bins.

	High	Medium	Low	Low	Medium	High
Absolute difference (His, Lys)	≥ 20%	5% to 20%	≤ 5%	≥ -5%	-5% to -20%	≤ -20%
Ratio difference (Ser, Thr, Tyr)	≥ 70%	40% to 70%	≤ 40%	≥ -40%	-40% to -70%	≤ -70%
	Increased labeling				Decreased labeling	

Residues	Native	Stressed 55 °C 4 h	Significant?	Absolute difference (H, K)	Ratio difference (S, T, Y)
Light chain - V_L domain					
S 5	0.03% ± 0.01%	0.04% ± 0.00%	No		
S 7					
S 12	0.03% ± 0.01%	0.02% ± 0.01%	No		
S 14					
K 18	0.56% ± 0.06%	0.66% ± 0.22%	No		
T 20	17% ± 4%	8% ± 2%	Yes	-10% ± 4%	-56% ± 26%
T 22					
S 26					
S 27					
S 28					
S 30					
Y 31					
H 33	0.08% ± 0.02%	0.21% ± 0.02%	Yes	+0.14% ± 0.03%	
K 38	0.15% ± 0.05%	0.24% ± 0.02%	No		
S 41	0.04% ± 0.01%	0.02% ± 0.02%	No		
S 42					
K 44	0.06% ± 0.02%	0.15% ± 0.01%	Yes	+0.08% ± 0.02%	
Y 48	0.09% ± 0.04%	0.13% ± 0.01%	No		
T 50	0.03% ± 0.03%	0.05% ± 0.01%	No		
S 51					
S 55					
S 62					
S 64					
S 66	18% ± 4%	15% ± 2%	No		
T 68	18% ± 4%	15% ± 2%	No		
S 69					
Y 70	14% ± 5%	11% ± 4%	No		
S 71					
T 73	53% ± 14%	86% ± 3%	Yes	+32% ± 15%	+61% ± 32%
S 75	53% ± 14%	86% ± 3%	Yes	+32% ± 15%	+61% ± 32%
T 84	1.3% ± 0.2%	2.6% ± 0.5%	Yes	+1.3% ± 0.6%	+103% ± 46%
Y 85	1.3% ± 0.2%	2.6% ± 0.5%	Yes	+1.3% ± 0.5%	+100% ± 43%
Y 86					
T 91	0.03% ± 0.02%	0.10% ± 0.01%	Yes	+0.07% ± 0.02%	+225% ± 196%
S 92	1.8% ± 0.6%	2.1% ± 0.6%	No		
T 96	2.0% ± 0.6%	2.2% ± 0.6%	No		
T 101	0.18% ± 0.01%	0.14% ± 0.02%	No		
K 102	0.08% ± 0.02%	0.10% ± 0.01%	No		
K 106					
Light chain - C_L domain					
T 108	1.3% ± 0.4%	1.1% ± 0.3%	No		
S 113	1.0% ± 0.3%	0.8% ± 0.1%	No		
S 120					
K 125					
S 126					

Residues	Native	Stressed 55 °C 4 h	Significant?	Absolute difference (H, K)	Ratio difference (S, T, Y)
T 128	5.4% ± 1.0%	3.1% ± 0.3%	Yes	-2.3% ± 1.1%	-42.4% ± 21.7%
S 130	5.4% ± 1.0%	3.1% ± 0.3%	Yes	-2.3% ± 1.1%	-42.4% ± 21.7%
Y 139					
K 144					
K 148	7.2% ± 1.7%	9.7% ± 0.8%	No		
S 155					
S 158					
S 161					
T 163	0.00% ± 0.00%	0.01% ± 0.01%	No		
S 167					
K 168	0.48% ± 0.19%	0.47% ± 0.16%	No		
S 170					
T 171					
Y 172					
S 173					
S 175					
S 176					
T 177					
T 179					
S 181					
K 182	5.0% ± 2.1%	19% ± 10%	No		
Y 185	23.5% ± 0.3%	27% ± 2%	Yes	+4% ± 2%	+16% ± 8%
K 187	85% ± 3%	86% ± 4%	No		
H 188	55% ± 2%	64% ± 3%	Yes	+10% ± 4%	
K 189	60% ± 3%	66% ± 3%	No		
Y 191	0.74% ± 0.09%	0.98% ± 0.33%	No		
T 196					
H 197	0.11% ± 0.03%	0.16% ± 0.03%	No		
S 201					
S 202					
T 205					
K 206	0.37% ± 0.13%	0.38% ± 0.10%	No		
S 207	4.2% ± 2.3%	2.8% ± 2.0%	No		
Heavy chain – V$_H$ domain					
K 13	2.7% ± 0.5%	3.7% ± 0.3%	Yes	+1.0% ± 0.6%	
S 17					
K 19	3.5% ± 0.7%	4.5% ± 0.4%	No		
S 21	24% ± 4%	11% ± 4%	Yes	-13% ± 6%	-55% ± 27%
K 23	86% ± 5%	92% ± 4%	No		
S 25					
Y 27					
T 28					
T 30					
S 31					
Y 32	0.05% ± 0.03%	0.40% ± 0.26%	No		
H 35	23% ± 3%	4% ± 1%	Yes	-19% ± 3%	
K 38	30% ± 2%	22% ± 3%	Yes	-8% ± 3%	
T 40					
Y 52	1.1% ± 0.2%	1.3% ± 0.2%	No		
T 58	0.12% ± 0.01%	0.15% ± 0.02%	No		
S 59	0.09% ± 0.01%	0.09% ± 0.03%	No		
Y 60	0.09% ± 0.01%	0.09% ± 0.03%	No		
K 63					
K 65					
K 67	23% ± 7%	33% ± 2%	No		
T 69					
T 71					
K 74	79% ± 4%	87% ± 1%	Yes	+8% ± 4%	
S 75					
S 76	2.4% ± 0.2%	1.5% ± 0.2%	Yes	-0.8% ± 0.3%	-35% ± 13%
S 77	0.01% ± 0.01%	0.01% ± 0.01%	No		
T 78	0.28% ± 0.10%	0.26% ± 0.23%	No		

Residues	Native	Stressed 55 °C 4 h	Significant?	Absolute difference (H, K)	Ratio difference (S, T, Y)
Y 80					
S 84	1.8% ± 0.7%	1.4% ± 0.6%	No		
S 85	0.32% ± 0.06%	0.28% ± 0.14%	No		
T 87					
S 88	0.02% ± 0.01%	0.01% ± 0.01%	No		
S 91	0.25% ± 0.04%	0.23% ± 0.09%	No		
Y 94	0.25% ± 0.04%	0.23% ± 0.09%	No		
Y 95					
S 99					
T 100					
Y 101					
Y 102					
Y 107					
T 115	14% ± 4%	12% ± 1%	No		
T 116					
T 118	71% ± 3%	59% ± 7%	Yes	-12% ± 7%	-17% ± 11%
S 120	84% ± 2%	70% ± 7%	Yes	-14% ± 7%	-17% ± 9%
Heavy chain – C$_H$1 domain					
S 123	1.2% ± 0.3%	1.6% ± 0.2%	No		
T 124					
K 125					
S 128	0.92% ± 0.44%	1.4% ± 1.0%	No		
S 135	1.1% ± 0.5%	1.3% ± 1.1%	No		
S 136					
K 137					
S 138					
T 139	0.13% ± 0.07%	0.12% ± 0.04%	No		
S 140	0.03% ± 0.02%	0.02% ± 0.01%	No		
T 143	0.01% ± 0.01%	0.01% ± 0.01%	No		
K 151					
Y 153	0.02% ± 0.02%	0.07% ± 0.05%	No		
T 159					
S 161	1.8% ± 0.2%	2.2% ± 0.4%	No		
S 164	0.002% ± 0.001%	0.016% ± 0.005%	Yes	+0.014% ± 0.005%	+890% ± 849%
T 168	1.3% ± 0.4%	2.4% ± 1.2%	No		
S 169	0.23% ± 0.13%	0.35% ± 0.14%	No		
H 172	0.23% ± 0.13%	0.35% ± 0.14%	No		
T 173	1.1% ±0.2%	2.2% ± 0.7%	No		
S 180					
S 181					
Y 184					
S 185					
S 187	10% ± 3%	11% ± 2%	No		
S 188	10% ± 3%	10% ± 2%	No		
T 191	4.6% ± 1.1%	4.3% ± 0.7%	No		
S 194					
S 195					
S 196	0.01% ± 0.00%	0.05% ± 0.03%	No		
T 199	0.01% ± 0.00%	0.05% ± 0.03%	No		
T 201					
Y 202					
H 208	0.00% ± 0.00%	0.00% ± 0.00%	No		
K 209	6.4% ± 2.0%	8.0% ± 0.9%	No		
S 211	0.00% ± 0.00%	0.00% ± 0.00%	No		
T 213					
K 214	0.50% ± 0.38%	0.40% ± 0.12%	No		
K 217					
K 218					
K 222	99% ± 1%	100% ± 0%	No		
S 223	0.06% ± 0.04%	0.04% ± 0.00%	No		
Heavy chain – Hinge region and C$_H$2 domain					
K 226	33% ± 7%	72% ± 5%	Yes	+39% ± 8%	

Residues	Native	Stressed 55 °C 4 h	Significant?	Absolute difference (H, K)	Ratio difference (S, T, Y)
T 227	1.4% ± 0.8%	1.6% ± 1.2%	No		
H 228	36% ± 7%	43% ± 15%	No		
T 229	36% ± 7%	42% ± 14%	No		
S 243					
K 250					
K 252	100% ± 0%	100% ± 0%	No		
T 254					
S 258	5.2% ± 1.8%	3.1% ± 0.8%	No		
T 260	0.70% ± 0.58%	0.81% ± 0.33%	No		
T 264	11% ± 2%	16% ± 11%	No		
S 271	10% ± 1%	15% ± 11%	No		
H 272	0.26% ± 0.08%	0.24% ± 0.06%	No		
K 278	0.16% ± 0.10%	0.28% ± 0.07%	No		
Y 282	0.39% ± 0.02%	0.33% ± 0.04%	No		
H 289	0.49% ± 0.13%	0.41% ± 0.03%	No		
K 292					
T 293	38% ± 6%	85% ± 8%	Yes	+46% ± 10%	+122% ± 32%
K 294					
Y 300	50% ± 1%	56% ± 7%	No		
S 302	38% ± 3%	16% ± 6%	Yes	-22% ± 7%	-58% ± 18%
T 303					
Y 304					
S 308					
T 311					
H 314	0.24% ± 0.08%	0.23% ± 0.06%	No		
K 321	0.29% ± 0.04%	0.28% ± 0.01%	No		
Y 323					
K 324					
K 326					
S 328					
K 330	100% ± 0%	100% ± 0%	No		
K 338	2.4% ± 1.7%	4.4% ± 0.5%	No		
T 339					
S 341					
K 342					
K 344					
Heavy chain – C$_H$3 domain					
Y 353	0.44% ± 0.62%	14% ± 2%	Yes	+14% ± 2%	+3118% ± 4418%
T 354					
S 358					
T 363	0.006% ± 0.002%	0.012% ± 0.003%	Yes	+0.006% ± 0.003%	+111% ± 67%
K 364	0.15% ± 0.07%	0.29% ± 0.21%	No		
S 368	0.02% ± 0.02%	0.18% ± 0.01%	Yes	+0.16% ± 0.03%	+1018% ± 1537%
T 370					
K 374	4.1% ± 1.6%	7.8% ± 2.0%	No		
Y 377	11% ± 2%	12% ± 7%	No		
S 379	13% ± 2%	13% ± 7%	No		
S 387	11% ± 2%	12% ± 7%	No		
Y 395	2.2% ± 0.8%	2.4% ± 1.0%	No		
K 396	0.82% ± 0.22%	0.58% ± 0.27%	No		
T 397	0.87% ± 0.29%	1.0% ± 0.2%	No		
T 398	12% ± 1%	17% ± 3%	No		
S 404	5.2% ± 1.3%	4.8% ± 1.8%	No		
S 407	15% ± 4%	30% ± 6%	Yes	+15% ± 7%	+101% ± 52%
Y 411	7.5% ± 5.1%	20% ± 4%	Yes	+12% ± 6%	+163% ± 138%
S 412					
K 413					
T 415					
K 418	29% ± 24%	46% ± 4%	No		
S 419	19% ± 15%	33% ± 2%	No		
S 428	0.86% ± 0.08%	1.2% ± 0.4%	No		

Residues	Native	Stressed 55 °C 4 h	Significant?	Absolute difference (H, K)	Ratio difference (S, T, Y)
S 430	6.8% ± 2.2%	12% ± 2%	No		
H 433	15% ± 2%	13% ± 4%	No		
H 437	5.8% ± 2.9%	7.4% ± 0.9%	No		
H 439	5.4% ± 0.9%	4.2% ± 3.2%	No		
Y 440	0.02% ± 0.01%	0.02% ± 0.01%	No		
T 441					
K 443	0.47% ± 0.24%	0.35% ± 0.07%	No		
S 444					
S 446					
S 448	71% ± 12%	92% ± 4%	Yes	+22% ± 13%	+30% ± 18%
K 451	41% ± 16%	79% ± 7%	Yes	+37% ± 18%	

223

Table B.3: DEPC modification percentages for individual residues of rituximab under native conditions and after heating to 65 °C for 4 h.

A difference was considered significant if the p-value, calculated by performing an unpaired T-test, was less than 0.05 (corresponding to a 95% confidence level, n = 4). Error bars shown in a table are standard deviations. Color coding of the cells (introduced below) indicates the extent of covalent labeling (CL) change which falls within low (L), medium (M), and high (H) bins.

	High	Medium	Low	Low	Medium	High
Absolute difference (His, Lys)	≥ 20%	5% to 20%	≤ 5%	≥ -5%	-5% to -20%	≤ -20%
Ratio difference (Ser, Thr, Tyr)	≥ 70%	40% to 70%	≤ 40%	≥ -40%	-40% to -70%	≤ -70%
	Increased labeling				Decreased labeling	

Residues	Native	Stressed 65 °C 4 h	Significant?	Absolute difference (H, K)	Ratio difference (S, T, Y)
Light chain - V_L domain					
S 5	0.03% ± 0.01%	0.01% ± 0.01%	Yes	-0.02% ± 0.01%	-62% ± 40%
S 7					
S 12	0.03% ± 0.01%	0.01% ± 0.01%	No		
S 14					
K 18	0.56% ± 0.06%	0.33% ± 0.27%	No		
T 20	17% ± 4%	13% ± 8%	No		
T 22					
S 26					
S 27					
S 28					
S 30					
Y 31					
H 33	0.08% ± 0.02%	0.61% ± 0.49%	No		
K 38	0.15% ± 0.05%	0.22% ± 0.06%	No		
S 41	0.04% ± 0.01%	0.12% ± 0.04%	No		
S 42					
K 44	0.06% ± 0.02%	0.09% ± 0.04%	No		
Y 48	0.09% ± 0.04%	0.06% ± 0.01%	No		
T 50	0.03% ± 0.03%	0.00% ± 0.00%	No		
S 51					
S 55					
S 62					
S 64					
S 66	18% ± 4%	12% ± 4%	No		
T 68	18% ± 4%	12% ± 4%	No		
S 69					
Y 70	14% ± 5%	1.8% ± 2.3%	Yes	-12% ± 5%	-87% ± 48%
S 71					
T 73	53% ± 14%	73% ± 13%	No		
S 75	53% ± 14%	73% ± 13%	No		
T 84	1.3% ± 0.2%	2.9% ± 1.3%	No		
Y 85	1.3% ± 0.2%	2.9% ± 1.3%	No		
Y 86					
T 91	0.03% ± 0.02%	0.01% ± 0.01%	No		
S 92	1.8% ± 0.6%	4.3% ± 2.2%	No		
T 96	2.0% ± 0.6%	4.3% ± 2.1%	No		
T 101	0.18% ± 0.01%	0.07% ± 0.05%	Yes	-0.11% ± 0.05	-60% ± 30%
K 102	0.08% ± 0.02%	0.04% ± 0.02%	No		
K 106					
Light chain - C_L domain					
T 108	1.3% ± 0.4%	0.66% ± 0.32%	No		
S 113	1.0% ± 0.3%	0.52% ± 0.26%	No		
S 120					
K 125					
S 126					

Residues	Native	Stressed 65 °C 4 h	Significant?	Absolute difference (H, K)	Ratio difference (S, T, Y)
T 128	5.4% ± 1.0%	3.2% ± 1.9%	No		
S 130	5.4% ± 1.0%	3.2% ± 1.9%	No		
Y 139					
K 144					
K 148	7.2% ± 1.7%	4.9% ± 1.0%	No		
S 155					
S 158					
S 161					
T 163	0.00% ± 0.00%	0.00% ± 0.00%	No		
S 167					
K 168	0.48% ± 0.19%	0.40% ± 0.21%	No		
S 170					
T 171					
Y 172					
S 173					
S 175					
S 176					
T 177					
T 179					
S 181					
K 182	5.0% ± 2.1%	13% ± 8%	No		
Y 185	23.5% ± 0.3%	25% ± 4%	No		
K 187	85% ± 3%	78% ± 4%	No		
H 188	55% ± 2%	62% ± 8%	No		
K 189	60% ± 3%	65% ± 6%	No		
Y 191	0.74% ± 0.09%	0.98% ± 0.24%	No		
T 196					
H 197	0.11% ± 0.03%	0.57% ± 0.22%	Yes	+0.46% ± 0.23%	
S 201					
S 202					
T 205					
K 206	0.37% ± 0.13%	0.61% ± 0.22%	No		
S 207	4.2% ± 2.3%	2.7% ± 0.6%	No		
Heavy chain – V_H domain					
K 13	2.7% ± 0.5%	3.5% ± 1.4%	No		
S 17					
K 19	3.5% ± 0.7%	3.4% ± 1.5%	No		
S 21	24% ± 4%	12% ± 8%	No		
K 23	86% ± 5%	96% ± 5%	No		
S 25					
Y 27					
T 28					
T 30					
S 31					
Y 32	0.05% ± 0.03%	0.27% ± 0.20%	No		
H 35	23% ± 3%	11% ± 3%	Yes	-12% ± 4%	
K 38	30% ± 2%	38% ± 10%	No		
T 40					
Y 52	1.1% ± 0.2%	1.4% ± 0.2%	No		
T 58	0.12% ± 0.01%	0.21% ± 0.15%	No		
S 59	0.09% ± 0.01%	0.19% ± 0.16%	No		
Y 60	0.09% ± 0.01%	0.19% ± 0.16%	No		
K 63					
K 65					
K 67	23% ± 7%	70% ± 17%	Yes	+48% ± 19%	
T 69					
T 71					
K 74	79% ± 4%	52% ± 8%	Yes	-27% ± 9%	
S 75					
S 76	2.4% ± 0.2%	4.0% ± 1.2%	No		
S 77	0.01% ± 0.01%	0.03% ± 0.01%	Yes	+0.02% ± 0.01%	+165% ± 127%
T 78	0.28% ± 0.10%	0.16% ± 0.09%	No		

Residues	Native	Stressed 65 °C 4 h	Significant?	Absolute difference (H, K)	Ratio difference (S, T, Y)
Y 80					
S 84	1.8% ± 0.7%	1.3% ± 0.5%	No		
S 85	0.32% ± 0.06%	0.26% ± 0.09%	No		
T 87					
S 88	0.02% ± 0.01%	0.03% ± 0.01%	No		
S 91	0.25% ± 0.04%	0.21% ± 0.06%	No		
Y 94	0.25% ± 0.04%	0.21% ± 0.06%	No		
Y 95					
S 99					
T 100					
Y 101					
Y 102					
Y 107					
T 115	14% ± 4%	14% ± 5%	No		
T 116					
T 118	71% ± 3%	55% ± 5%	Yes	-16% ± 6%	-22% ± 9%
S 120	84% ± 2%	69% ± 4%	Yes	-15% ± 4%	-18% ± 5%
Heavy chain – C$_H$1 domain					
S 123	1.2% ± 0.3%	1.2% ± 0.2%	No		
T 124					
K 125					
S 128	0.92% ± 0.44%	0.30% ± 0.16%	No		
S 135	1.1% ± 0.5%	0.35% ± 0.34%	No		
S 136					
K 137					
S 138					
T 139	0.13% ± 0.07%	0.23% ± 0.10%	No		
S 140	0.03% ± 0.02%	0.05% ± 0.01%	No		
T 143	0.01% ± 0.01%	0.02% ± 0.01%	No		
K 151					
Y 153	0.02% ± 0.02%	0.01% ± 0.01%	No		
T 159					
S 161	1.8% ± 0.2%	1.2% ± 0.2%	Yes	-0.66% ± 0.31%	-36% ± 18%
S 164	0.00% ± 0.00%	0.02% ± 0.02%	No		
T 168	1.3% ± 0.4%	2.2% ± 1.5%	No		
S 169	0.23% ± 0.13%	0.23% ± 0.18%	No		
H 172	0.23% ± 0.13%	0.23% ± 0.18%	No		
T 173	1.1% ±0.2%	1.9% ± 0.9%	No		
S 180					
S 181					
Y 184					
S 185					
S 187	10% ± 3%	11% ± 2%	No		
S 188	10% ± 3%	11% ± 2%	No		
T 191	4.6% ± 1.1%	4.4% ± 0.1%	No		
S 194					
S 195					
S 196	0.01% ± 0.00%	0.08% ± 0.02%	Yes	+0.07% ± 0.02%	+728% ± 276%
T 199	0.01% ± 0.00%	0.05% ± 0.02%	Yes	+0.05% ± 0.02%	+689% ± 396%
T 201					
Y 202					
H 208	0.00% ± 0.00%	0.04% ± 0.01%	Yes	+0.04% ± 0.01%	
K 209	6.4% ± 2.0%	4.0% ± 1.0%	No		
S 211	0.00% ± 0.00%	0.21% ± 0.02%	Yes	+0.21% ± 0.02%	+6074% ± 5773%
T 213					
K 214	0.50% ± 0.38%	0.19% ± 0.12%	No		
K 217					
K 218					
K 222	99% ± 1%	98% ± 2%	No		
S 223	0.06% ± 0.04%	0.05% ± 0.01%	No		
Heavy chain – Hinge region and C$_H$2 domain					
K 226	33% ± 7%	39% ± 17%	No		

Residues	Native	Stressed 65 °C 4 h	Significant?	Absolute difference (H, K)	Ratio difference (S, T, Y)
T 227	1.4% ± 0.8%	1.2% ± 0.8%	No		
H 228	36% ± 7%	40% ± 12%	No		
T 229	36% ± 7%	40% ± 12%	No		
S 243					
K 250					
K 252	100% ± 0%	100% ± 0%	No		
T 254					
S 258	5.2% ± 1.8%	1.2% ± 1.5%	Yes	-4.0% ± 2.3%	-77% ± 52%
T 260	0.70% ± 0.58%	0.16% ± 0.10%	No		
T 264	11% ± 2%	9.6% ± 2.3%	No		
S 271	10% ± 1%	9.3% ± 2.2%	No		
H 272	0.26% ± 0.08%	0.19% ± 0.10	No		
K 278	0.16% ± 0.10%	0.24% ± 0.03%	No		
Y 282	0.39% ± 0.02%	0.39% ± 0.08%	No		
H 289	0.49% ± 0.13%	0.34% ± 0.23%	No		
K 292					
T 293	38% ± 6%	82% ± 27%	No		
K 294					
Y 300	50% ± 1%	30% ± 13%	No		
S 302	38% ± 3%	10% ± 16%	Yes	-28% ± 16%	-74% ± 43%
T 303					
Y 304					
S 308					
T 311					
H 314	0.24% ± 0.08%	0.20% ± 0.03%	No		
K 321	0.29% ± 0.04%	0.37% ± 0.06%	No		
Y 323					
K 324					
K 326					
S 328					
K 330	100% ± 0%	100% ± 0%	No		
K 338	2.4% ± 1.7%	1.5% ± 0.4%	No		
T 339					
S 341					
K 342					
K 344					
Heavy chain – C_H3 domain					
Y 353	0.44% ± 0.62%	0.00% ± 0.00%	No		
T 354					
S 358					
T 363	0.006% ± 0.002%	0.002% ± 0.001%	Yes	-0.004 ± 0.002	-66% ± 40%
K 364	0.15% ± 0.07%	0.10% ± 0.08%	No		
S 368	0.02% ± 0.02%	0.00% ± 0.00%	No		
T 370					
K 374	4.1% ± 1.6%	3.5% ± 2.4%	No		
Y 377	11% ± 2%	8.5% ± 4.9%	No		
S 379	13% ± 2%	10% ± 5%	No		
S 387	11% ± 2%	8.5% ± 4.9%	No		
Y 395	2.2% ± 0.8%	2.3% ± 0.3%	No		
K 396	0.82% ± 0.22%	0.73% ± 0.19%	No		
T 397	0.87% ± 0.29%	1.0% ± 0.2%	No		
T 398	12% ± 1%	16% ± 10%	No		
S 404	5.2% ± 1.3%	4.9% ± 2.7%	No		
S 407	15% ± 4%	23% ± 12%	No		
Y 411	7.5% ± 5.1%	13% ± 6%	No		
S 412					
K 413					
T 415					
K 418	29% ± 24%	36% ± 18%	No		
S 419	19% ± 15%	25% ± 15%	No		
S 428	0.86% ± 0.08%	0.41% ± 0.23%	Yes	-0.45% ± 0.24%	-52% ± 29%
S 430	6.8% ± 2.2%	8.1% ± 4.6%	No		

Residues	Native	Stressed 65 °C 4 h	Significant?	Absolute difference (H, K)	Ratio difference (S, T, Y)
H 433	15% ± 2%	16% ± 3%	No		
H 437	5.8% ± 2.9%	5.2% ± 0.9%	No		
H 439	5.4% ± 0.9%	2.5% ± 0.3%	Yes	-2.9% ± 0.9%	
Y 440	0.02% ± 0.01%	0.01% ± 0.01%	Yes	-0.01% ± 0.01%	-60% ± 42%
T 441					
K 443	0.47% ± 0.24%	0.25% ± 0.15%	No		
S 444					
S 446					
S 448	71% ± 12%	47% ± 9%	No		
K 451	41% ± 16%	5.1% ± 5.4%	Yes	-36% ± 17%	

BIBLIOGRAPHY

2016 FDA Approved Drugs. http://www.centerwatch.com/drug-information/fda-approved-drugs/year/2016 (accessed Mar 29).

Abbott, W. M.; Damschroder, M. M.; Lowe, D. C., Current approaches to fine mapping of antigen-antibody interactions. *Immunology* **2014,** *142* (4), 526-535.

Ambrogelly, A.; Gozo, S.; Katiyar, A.; Dellatore, S.; Kune, Y.; Bhat, R.; Sun, J.; Li, N.; Wang, D.; Nowak, C.; Neill, A.; Ponniah, G.; King, C.; Mason, B.; Beck, A.; Liu, H., Analytical comparability study of recombinant monoclonal antibody therapeutics. *mAbs* **2018,** *10* (4), 513-538.

Andersen Christian, B.; Manno, M.; Rischel, C.; Thórólfsson, M.; Martorana, V., Aggregation of a multidomain protein: A coagulation mechanism governs aggregation of a model IgG1 antibody under weak thermal stress. *Protein Science* **2009,** *19* (2), 279-290.

Asuero, A. G.; Sayago, A.; González, A. G., The Correlation Coefficient: An Overview. *Critical Reviews in Analytical Chemistry* **2006,** *36* (1), 41-59.

Atkins, P.; de Paula, J., *Atkins' Physical Chemistry. 10th ed*. 6th ed.; Oxford University Press: Oxford, England, United Kingdom, 2014.

Aye, T. T.; Low, T. Y.; Sze, S. K., Nanosecond Laser-Induced Photochemical Oxidation Method for Protein Surface Mapping with Mass Spectrometry. *Analytical Chemistry* **2005,** *77* (18), 5814-5822.

Baslé, E.; Joubert, N.; Pucheault, M., Protein Chemical Modification on Endogenous Amino Acids. *Chemistry & Biology* **2010,** *17* (3), 213-227.

Bavro, Vassiliy N.; Gupta, S.; Ralston, C., Oxidative footprinting in the study of structure and function of membrane proteins: current state and perspectives. *Biochemical Society Transactions* **2015,** *43* (5), 983.

Beck, A.; Diemer, H.; Ayoub, D.; Debaene, F.; Wagner-Rousset, E.; Carapito, C.; Van Dorsselaer, A.; Sanglier-Cianférani, S., Analytical characterization of biosimilar antibodies and Fc-fusion proteins. *Trends Anal Chem* **2013,** *48*, 81-95.

Beck, A.; Wurch, T.; Bailly, C.; Corvaia, N., Strategies and challenges for the next generation of therapeutic antibodies. *Nat Rev Immunol* **2010,** *10* (5), 345-352.

Berkowitz, S. A.; Engen, J. R.; Mazzeo, J. R.; Jones, G. B., Analytical tools for characterizing biopharmaceuticals and the implications for biosimilars. *Nature Reviews Drug Discovery* **2012,** *11* (7), 527-540.

Bitan, G., Structural Study of Metastable Amyloidogenic Protein Oligomers by Photo-Induced Cross-Linking of Unmodified Proteins. In *Methods in Enzymology*, Academic Press: 2006; Vol. 413, pp 217-236.

Blencowe, A.; Hayes, W., Development and application of diazirines in biological and synthetic macromolecular systems. *Soft Matter* **2005,** *1* (3), 178-205.

Bommana, R.; Chai, Q.; Schöneich, C.; Weiss, W. F.; Majumdar, R., Understanding the Increased Aggregation Propensity of a Light-Exposed IgG1 Monoclonal Antibody Using Hydrogen Exchange Mass Spectrometry, Biophysical Characterization, and Structural Analysis. *Journal of Pharmaceutical Sciences* **2018,** *107* (6), 1498-1511.

Bonnington, L.; Lindner, I.; Gilles, U.; Kailich, T.; Reusch, D.; Bulau, P., Application of Hydrogen/Deuterium Exchange-Mass Spectrometry to Biopharmaceutical

Development Requirements: Improved Sensitivity to Detection of Conformational Changes. *Analytical Chemistry* **2017,** *89* (16), 8233-8237.

Borotto, N. B.; Degraan-Weber, N.; Zhou, Y.; Vachet, R. W., Label Scrambling During CID of Covalently Labeled Peptide Ions. *J. Am. Soc. Mass Spectrom.* **2014,** *25* (10), 1739-1746.

Borotto, N. B.; Zhou, Y.; Hollingsworth, S. R.; Hale, J. E.; Graban, E. M.; Vaughan, R. C.; Vachet, R. W., Investigating Therapeutic Protein Structure with Diethylpyrocarbonate Labeling and Mass Spectrometry. *Analytical Chemistry* **2015,** *87* (20), 10627-10634.

Bridgewater, J. D.; Lim, J.; Vachet, R. W., Transition Metal−Peptide Binding Studied by Metal-Catalyzed Oxidation Reactions and Mass Spectrometry. *Analytical Chemistry* **2006,** *78* (7), 2432-2438.

Bridgewater, J. D.; Lim, J.; Vachet, R. W., Using metal-catalyzed oxidation reactions and mass spectrometry to identify amino acid residues within 10 Å of the metal in Cu-binding proteins. *Journal of the American Society for Mass Spectrometry* **2006,** *17* (11), 1552-1559.

Brunner, J.; Senn, H.; Richards, F. M., 3-Trifluoromethyl-3-phenyldiazirine. A new carbene generating group for photolabeling reagents. *Journal of Biological Chemistry* **1980,** *255* (8), 3313-3318.

Cammarata, M.; Lin, K.-Y.; Pruet, J.; Liu, H.-w.; Brodbelt, J., Probing the Unfolding of Myoglobin and Domain C of PARP-1 with Covalent Labeling and Top-Down Ultraviolet Photodissociation Mass Spectrometry. *Analytical Chemistry* **2014,** *86* (5), 2534-2542.

Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T. A.; Brusniak, M.-Y.; Paulse, C.; Creasy, D.; Flashner, L.; Kani, K.; Moulding, C.; Seymour, S. L.; Nuwaysir, L. M.; Lefebvre, B.; Kuhlmann, F.; Roark, J.; Rainer, P.; Detlev, S.; Hemenway, T.; Huhmer, A.; Langridge, J.; Connolly, B.; Chadick, T.; Holly, K.; Eckels, J.; Deutsch, E. W.; Moritz, R. L.; Katz, J. E.; Agus, D. B.; MacCoss, M.; Tabb, D. L.; Mallick, P., A cross-platform toolkit for mass spectrometry and proteomics. *Nat Biotech* **2012,** *30* (10), 918-920.

Chantalat, L.; Jones, N. D.; Korber, F.; Navaza, J.; Pavlovsky, A. G., The Crystal Structure of Wild-type Growth Hormone at 2.5 Å Resolution. *Protein Pept. Lett.* **1995,** *2* (2), 333-340.

Charvátová, O.; Foley, B. L.; Bern, M. W.; Sharp, J. S.; Orlando, R.; Woods, R. J., Quantifying Protein Interface Footprinting by Hydroxyl Radical Oxidation and Molecular Dynamics Simulation: Application to Galectin-1. *Journal of the American Society for Mass Spectrometry* **2008,** *19* (11), 1692-1705.

Chen, J.; Cui, W.; Giblin, D.; Gross, M. L., New Protein Footprinting: Fast Photochemical Iodination Combined with Top-down and Bottom-up Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2012,** *23* (8), 1306-1318.

Chen, J.; Rempel, D. L.; Gau, B. C.; Gross, M. L., Fast Photochemical Oxidation of Proteins and Mass Spectrometry Follow Submillisecond Protein Folding at the Amino-Acid Level. *J. Am. Chem. Soc.* **2012,** *134* (45), 18724-18731.

Chen, J.; Rempel, D. L.; Gross, M. L., Temperature Jump and Fast Photochemical Oxidation Probe Submillisecond Protein Folding. *J. Am. Chem. Soc.* **2010,** *132* (44), 15502-15504.

Cheng, M.; Zhang, B.; Cui, W.; Gross, M. L., Laser-Initiated Radical Trifluoromethylation of Peptides and Proteins: Application to Mass-Spectrometry-Based Protein Footprinting. *Angewandte Chemie International Edition* **2017,** *56* (45), 14007-14010.

Coales, S. J.; Tuske, S. J.; Tomasso, J. C.; Hamuro, Y., Epitope mapping by amide hydrogen/deuterium exchange coupled with immobilization of antibody, on-line proteolysis, liquid chromatography and mass spectrometry. *Rapid Communications in Mass Spectrometry* **2009,** *23* (5), 639-647.

Cohen, F. E.; Kelly, J. W., Therapeutic approaches to protein-misfolding diseases. *Nature* **2003,** *426*, 905.

Courtois, F.; Schneider, C. P.; Agrawal, N. J.; Trout, B. L., Rational Design of Biobetters with Enhanced Stability. *Journal of Pharmaceutical Sciences* **2015,** *104* (8), 2433-2440.

Cragg, M. S.; Bayne, M. B.; Tutt, A. L.; French, R. R.; Beers, S.; Glennie, M. J.; Illidge, T. M., A new anti-idiotype antibody capable of binding rituximab on the surface of lymphoma cells. *Blood* **2004,** *104* (8), 2540.

Craig, P. O.; Ureta, D. B.; Delfino, J. M., Probing protein conformation with a minimal photochemical reagent. *Protein Science : A Publication of the Protein Society* **2002,** *11* (6), 1353-1366.

Craig, R.; Beavis, R. C., A method for reducing the time required to match protein sequences with tandem mass spectra. *Rapid Communications in Mass Spectrometry* **2003,** *17* (20), 2310-2316.

Craig, R.; Beavis, R. C., TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* **2004,** *20* (9), 1466-1467.

Das, J., Aliphatic Diazirines as Photoaffinity Probes for Proteins: Recent Developments. *Chemical Reviews* **2011,** *111* (8), 4405-4417.

Deperalta, G.; Alvarez, M.; Bechtel, C.; Dong, K.; McDonald, R.; Ling, V., Structural analysis of a therapeutic monoclonal antibody dimer by hydroxyl radical footprinting. *mAbs* **2013,** *5* (1), 86-101.

Dingman, R.; Balu-Iyer, S. V., Immunogenicity of Protein Pharmaceuticals. *Journal of Pharmaceutical Sciences* **2019,** *108* (5), 1637-1654.

Dobson, C. M., Protein folding and misfolding. *Nature* **2003,** *426*, 884.

Du, J.; Wang, H.; Zhong, C.; Peng, B.; Zhang, M.; Li, B.; Huo, S.; Guo, Y.; Ding, J., Structural Basis for Recognition of CD20 by Therapeutic Antibody Rituximab. *Journal of Biological Chemistry* **2007,** *282* (20), 15073-15080.

Ecker, D. M.; Jones, S. D.; Levine, H. L., The therapeutic monoclonal antibody market. *mAbs* **2015,** *7* (1), 9-14.

Engen, J. R.; Smith, D. L., Peer Reviewed: Investigating Protein Structure and Dynamics by Hydrogen Exchange MS. *Analytical Chemistry* **2001,** *73* (9), 256 A-265 A.

Espino, J. A.; Mali, V. S.; Jones, L. M., In Cell Footprinting Coupled with Mass Spectrometry for the Structural Analysis of Proteins in Live Cells. *Analytical Chemistry* **2015,** *87* (15), 7971-7978.

European Medicines Agency. Guideline on development, production, characterisation and specification for monoclonal antibodies and related products. 2016 Aug 4 [accessed 2017 Sep 25]. http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000867.jsp&mid=WC0b01ac058002956b.

European Medicines Agency. Guideline on similar biological medicinal products containing biotechnology-derived proteins as active substance: quality issues (revision 1). 2014 Jun 3 [accessed 2017 Sep 25]. http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000886.jsp&mid=WC0b01ac058002956b.

EvaluatePharma® *World Preview 2017, Outlook to 2022*; Evaluate Ltd: London, England, United Kingdom, 2017; p 8.

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

Finkelstein, A. V.; Ptitsyn, O. B., *Protein Physics: A Course of Lectures*. Academic Press: London, England, United Kingdom, 2002; pp 33-42.

Floege, J.; Ketteler, M., β2-Microglobulin–derived amyloidosis: An update. *Kidney International* **2001,** *59*, S164-S171.

Flores-Ortiz, L. F.; Campos-García, V. R.; Perdomo-Abúndez, F. C.; Pérez, N. O.; Medina-Rivero, E., Physicochemical Properties of Rituximab. *Journal of Liquid Chromatography & Related Technologies* **2014,** *37* (10), 1438-1452.

Fraczkiewicz, R.; Braun, W., Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J. Comput. Chem.* **1998,** *19* (3), 319-333.

Frokjaer, S.; Otzen, D. E., Protein drug stability: a formulation challenge. *Nature Reviews Drug Discovery* **2005,** *4* (4), 298-306.

Gau, B. C.; Chen, J.; Gross, M. L., Fast photochemical oxidation of proteins for comparing solvent-accessibility changes accompanying protein folding: Data processing and application to barstar. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2013,** *1834* (6), 1230-1238.

Gau, B. C.; Sharp, J. S.; Rempel, D. L.; Gross, M. L., Fast Photochemical Oxidation of Protein Footprints Faster than Protein Unfolding. *Analytical Chemistry* **2009,** *81* (16), 6563-6571.

Gazzano-Santoro, H.; Ralph, P.; Ryskamp, T. C.; Chen, A. B.; Mukku, V. R., A non-radioactive complement-dependent cytotoxicity assay for anti-CD20 monoclonal antibody. *Journal of Immunological Methods* **1997,** *202* (2), 163-171.

Geer, L. Y.; Markey, S. P.; Kowalak, J. A.; Wagner, L.; Xu, M.; Maynard, D. M.; Yang, X.; Shi, W.; Bryant, S. H., Open Mass Spectrometry Search Algorithm. *Journal of Proteome Research* **2004,** *3* (5), 958-964.

Gerega, S. K.; Downard, K. M., PROXIMO—a new docking algorithm to model protein complexes using data from radical probe mass spectrometry (RP-MS). *Bioinformatics* **2006,** *22* (14), 1702-1709.

Goldsmith, S. C.; Guan, J.-Q.; Almo, S. C.; Chance, M. R., Synchrotron Protein Footprinting: A Technique to Investigate Protein-Protein Interactions. *Journal of Biomolecular Structure and Dynamics* **2001,** *19* (3), 405-418.

Gómez, G. E.; Cauerhff, A.; Craig, P. O.; Goldbaum, F. A.; Delfino, J. M., Exploring protein interfaces with a general photochemical reagent. *Protein Science* **2006**, *15* (4), 744-752.

Götte, M.; Marquet, R.; Isel, C.; Anderson, V. E.; Keith, G.; Gross, H. J.; Ehresmann, C.; Ehresmann, B.; Heumann, H., Probing the higher order structure of RNA with peroxonitrous acid. *FEBS Letters* **1996**, *390* (2), 226-228.

Guan, J.-Q.; Vorobiev, S.; Almo, S. C.; Chance, M. R., Mapping the G-Actin Binding Surface of Cofilin Using Synchrotron Protein Footprinting. *Biochemistry* **2002**, *41* (18), 5765-5775.

Guo, C.; Cheng, M.; Gross, M. L., Protein-Metal-Ion Interactions Studied by Mass Spectrometry-Based Footprinting with Isotope-Encoded Benzhydrazide. *Anal. Chem.* **2019**, *91* (2), 1416-1423.

Guo, X.; Bandyopadhyay, P.; Schilling, B.; Young, M. M.; Fujii, N.; Aynechi, T.; Guy, R. K.; Kuntz, I. D.; Gibson, B. W., Partial Acetylation of Lysine Residues Improves Intraprotein Cross-Linking. *Anal. Chem.* **2008**, *80* (4), 951-960.

Gupta, S.; Celestre, R.; Petzold, C. J.; Chance, M. R.; Ralston, C., Development of a microsecond X-ray protein footprinting facility at the Advanced Light Source. *Journal of Synchrotron Radiation* **2014**, *21* (4), 690-699.

Hambly, D. M.; Gross, M. L., Laser Flash Photolysis of Hydrogen Peroxide to Oxidize Protein Solvent-Accessible Residues on the Microsecond Timescale. *Journal of the American Society for Mass Spectrometry* **2005**, *16* (12), 2057-2063.

Hambly, D.; Gross, M., Laser flash photochemical oxidation to locate heme binding and conformational changes in myoglobin. *International Journal of Mass Spectrometry* **2007**, *259* (1), 124-129.

Hampel, K. J.; Burke, J. M., Time-Resolved Hydroxyl-Radical Footprinting of RNA Using Fe(II)-EDTA. *Methods* **2001**, *23* (3), 233-239.

Hampson, G.; Ward, T. H.; Cummings, J.; Bayne, M.; Tutt, A. L.; Cragg, M. S.; Dive, C.; Illidge, T. M., Validation of an ELISA for the determination of rituximab pharmacokinetics in clinical trials subjects. *Journal of Immunological Methods* **2010**, *360* (1), 30-38.

Harris, T. K.; Turner, G. J., Structural Basis of Perturbed pKa Values of Catalytic Groups in Enzyme Active Sites. *IUBMB Life* **2002**, *53* (2), 85-98.

Hazenberg, B. P. C., Amyloidosis: A Clinical Overview. *Rheumatic Disease Clinics of North America* **2013**, *39* (2), 323-345.

Herrmann, A.; Svangård, E.; Claeson, P.; Gullbo, J.; Bohlin, L.; Göransson, U., Key role of glutamic acid for the cytotoxic activity of the cyclotide cycloviolacin O2. *Cellular and Molecular Life Sciences CMLS* **2006**, *63* (2), 235-245.

Hickey, J. M.; Toprani, V. M.; Kaur, K.; Mishra, R. P. N.; Goel, A.; Oganesyan, N.; Lees, A.; Sitrin, R.; Joshi, S. B.; Volkin, D. B., Analytical Comparability Assessments of 5 Recombinant CRM197 Proteins From Different Manufacturers and Expression Systems. *Journal of Pharmaceutical Sciences* **2018**, *107* (7), 1806-1819.

Holding, A. N., XL-MS: Protein cross-linking coupled with mass spectrometry. *Methods* **2015**, *89*, 54-63.

Holman, J. D.; Tabb, D. L.; Mallick, P., Employing ProteoWizard to Convert Raw Mass Spectrometry Data. In *Current Protocols in Bioinformatics*, John Wiley & Sons, Inc.: 2002.

Houde, D.; Berkowitz, S. A.; Engen, J. R., The Utility of Hydrogen/Deuterium Exchange Mass Spectrometry in Biopharmaceutical Comparability Studies. *Journal of Pharmaceutical Sciences* **2011,** *100* (6), 2071-2086.

Huang, R. Y. C.; O'Neil, S. R.; Lipovšek, D.; Chen, G., Conformational Assessment of Adnectin and Adnectin-Drug Conjugate by Hydrogen/Deuterium Exchange Mass Spectrometry. *Journal of The American Society for Mass Spectrometry* **2018,** *29* (7), 1524-1531.

Huang, W.; Ravikumar, K. M.; Parisien, M.; Yang, S., Theoretical modeling of multiprotein complexes by iSPOT: Integration of small-angle X-ray scattering, hydroxyl radical footprinting, and computational docking. *Journal of structural biology* **2016,** *196* (3), 340-349.

Huang, W.; Ravikumar, Krishnakumar M.; Chance, Mark R.; Yang, S., Quantitative Mapping of Protein Structure by Hydroxyl Radical Footprinting-Mediated Structural Mass Spectrometry: A Protection Factor Analysis. *Biophysical Journal* **2015,** *108* (1), 107-115.

Iacob, R. E.; Bou-Assaf, G. M.; Makowski, L.; Engen, J. R.; Berkowitz, S. A.; Houde, D., Investigating Monoclonal Antibody Aggregation Using a Combination of H/DX-MS and Other Biophysical Measurements. *Journal of Pharmaceutical Sciences* **2013,** *102* (12), 4315-4329.

ICH Harmonised Tripartite Guideline. Q5E Comparability of Biotechnological/Biological Products Subject to Changes in their Manufacturing Process. 2004 Nov [accessed 2016 Mar 12]. http://www.ich.org/products/guidelines/quality/quality-single/article/comparability-of-biotechnologicalbiological-products-subject-to-changes-in-their-manufacturing-proc.html.

ICH Harmonised Tripartite Guideline. Q6B Specifications: Test Procedures and Acceptance Criteria for Biotechnological/Biological Products.1999 Mar [accessed 2016 Mar 12]. http://www.ich.org/products/guidelines/quality/quality-single/article/specifications-test-procedures-and-acceptance-criteria-for-biotechnologicalbiological-products.html.

Isom, D. G.; Castañeda, C. A.; Cannon, B. R.; García-Moreno E, B., Large shifts in pKa values of lysine residues buried inside a protein. *Proceedings of the National Academy of Sciences* **2011,** *108* (13), 5260.

Ivankov, D. N.; Finkelstein, A. V., Prediction of protein folding rates from the amino acid sequence-predicted secondary structure. *Proceedings of the National Academy of Sciences of the United States of America* **2004,** *101* (24), 8942-8944.

Jhan, S.-Y.; Huang, L.-J.; Wang, T.-F.; Chou, H.-H.; Chen, S.-H., Dimethyl Labeling Coupled with Mass Spectrometry for Topographical Characterization of Primary Amines on Monoclonal Antibodies. *Analytical Chemistry* **2017,** *89* (7), 4255-4263.

Jones, L. M.; B. Sperry, J.; A. Carroll, J.; Gross, M. L., Fast Photochemical Oxidation of Proteins for Epitope Mapping. *Analytical Chemistry* **2011,** *83* (20), 7657-7661.

Jumper, C. C.; Bomgarden, R.; Rogers, J.; Etienne, C.; Schriemer, D. C., High-Resolution Mapping of Carbene-Based Protein Footprints. *Anal. Chem.* **2012,** *84* (10), 4411-4418.

Jumper, C. C.; Schriemer, D. C., Mass Spectrometry of Laser-Initiated Carbene Reactions for Protein Topographic Analysis. *Anal. Chem.* **2011,** *83* (8), 2913-2920.

Jung, S. K.; Lee, K. H.; Jeon, J. W.; Lee, J. W.; Kwon, B. O.; Kim, Y. J.; Bae, J. S.; Kim, D.-I.; Lee, S. Y.; Chang, S. J., Physicochemical characterization of Remsima®. *mAbs* **2014,** *6* (5), 1163-1177.

Kaltashov, I. A.; Bobst, C. E.; Abzalimov, R. R.; Berkowitz, S. A.; Houde, D., Conformation and Dynamics of Biopharmaceuticals: Transition of Mass Spectrometry-Based Tools from Academe to Industry. *Journal of the American Society for Mass Spectrometry* **2010,** *21* (3), 323-337.

Kaltashov, I. A.; Bobst, C. E.; Abzalimov, R. R.; Wang, G.; Baykal, B.; Wang, S., Advances and challenges in analytical characterization of biotechnology products: Mass spectrometry-based approaches to study properties and behavior of protein therapeutics. *Biotechnology Advances* **2012,** *30* (1), 210-222.

Kaltashov, I. A.; Eyles, S. J., Studies of biomolecular conformations and conformational dynamics by mass spectrometry. *Mass Spectrometry Reviews* **2002,** *21* (1), 37-71.

Kamal, J. K. A.; Chance, M. R., Modeling of protein binary complexes using structural mass spectrometry data. *Protein Science* **2008,** *17* (1), 79-94.

Kaplon, H.; Reichert, J. M., Antibodies to watch in 2019. *mAbs* **2018**, In Press, DOI: 10.1080/19420862.2018.1556465.

Katajamaa, M.; Miettinen, J.; Orešič, M., MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data. *Bioinformatics* **2006,** *22* (5), 634-636.

Katta, V.; Chait, B. T.; Carr, S., Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry. *Rapid Communications in Mass Spectrometry* **1991,** *5* (4), 214-217.

Kaur, P.; Kiselar, J. G.; Chance, M. R., Integrated Algorithms for High-Throughput Examination of Covalently Labeled Biomolecules by Structural Mass Spectrometry. *Analytical Chemistry* **2009,** *81* (19), 8141-8149.

Kaur, P.; Kiselar, J.; Shi, W.; Yang, S.; Chance, M. R., Covalent Labeling Techniques for Characterizing Higher Order Structure of Monoclonal Antibodies. In *State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Volume 3. Defining the Next Generation of Analytical and Biophysical Techniques*, American Chemical Society: 2015; Vol. 1202, pp 45-73.

Kaur, P.; Tomechko, S. E.; Kiselar, J.; Shi, W.; Deperalta, G.; Wecksler, A. T.; Gokulrangan, G.; Ling, V.; Chance, M. R., Characterizing monoclonal antibody structure by carboxyl group footprinting. *mAbs* **2015,** *7* (3), 540-552.

Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W., Top Down versus Bottom Up Protein Characterization by Tandem High-Resolution Mass Spectrometry. *Journal of the American Chemical Society* **1999,** *121* (4), 806-812.

Khurana, R.; Udgaonkar, J. B., Equilibrium unfolding studies of barstar: Evidence for an alternative conformation which resembles a molten globule. *Biochemistry* **1994,** *33* (1), 106-115.

Kim, S.; Pevzner, P. A., MS-GF+ makes progress towards a universal database search tool for proteomics. **2014,** *5*, 5277.

King, P. A.; Jamison, E.; Strahs, D.; Anderson, V. E.; Brenowitz, M., 'Footprinting' proteins on DNA with peroxonitrous acid. *Nucleic Acids Research* **1993,** *21* (10), 2473-2478.

Kiselar, J. G.; Janmey, P. A.; Almo, S. C.; Chance, M. R., Structural Analysis of Gelsolin Using Synchrotron Protein Footprinting. *Molecular & Cellular Proteomics* **2003,** *2* (10), 1120.

Kiselar, J. G.; Maleknia, S. D.; Sullivan, M.; Downard, K. M.; Chance, M. R., Hydroxyl radical probe of protein surfaces using synchrotron X-ray radiolysis and mass spectrometry. *International Journal of Radiation Biology* **2002,** *78* (2), 101-114.

Kiselar, J.; Chance, M. R., High-Resolution Hydroxyl Radical Protein Footprinting: Biophysics Tool for Drug Discovery. *Annu. Rev. Biophys.* **2018,** *47* (1), 315-333.

Klinger, A. L.; Kiselar, J.; Ilchenko, S.; Komatsu, H.; Chance, M. R.; Axelsen, P. H., A Synchrotron-Based Hydroxyl Radical Footprinting Analysis of Amyloid Fibrils and Prefibrillar Intermediates with Residue-Specific Resolution. *Biochemistry* **2014,** *53* (49), 7724-7734.

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.

Konermann, L.; Stocks, B. B.; Pan, Y.; Tong, X., Mass spectrometry combined with oxidative labeling for exploring protein structure and folding. *Mass Spectrometry Reviews* **2010,** *29* (4), 651-667.

Konermann, L.; Pan, Y., Exploring membrane protein structural features by oxidative labeling and mass spectrometry. *Expert Review of Proteomics* **2012,** *9* (5), 497-504.

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

Lee, K. H.; Lee, J.; Bae, J. S.; Kim, Y. J.; Kang, H. A.; Kim, S. H.; Lee, S. J.; Lim, K. J.; Lee, J. W.; Jung, S. K.; Chang, S. J., Analytical similarity assessment of rituximab biosimilar CT-P10 to reference medicinal product. *mAbs* **2018,** *10* (3), 380-396.

Li, H.; Rahimi, F.; Sinha, S.; Maiti, P.; Bitan, G.; Murakami, K., Amyloids and Protein Aggregation—Analytical Methods. In *Encyclopedia of Analytical Chemistry*, John Wiley & Sons, Ltd: 2006.

Li, H.; Robertson, A. D.; Jensen, J. H., Very fast empirical prediction and rationalization of protein pKa values. *Proteins: Struct., Funct., Bioinf.* **2005,** *61* (4), 704-721.

Li, J.; Wei, H.; Krystek, S. R.; Bond, D.; Brender, T. M.; Cohen, D.; Feiner, J.; Hamacher, N.; Harshman, J.; Huang, R. Y. C.; Julien, S. H.; Lin, Z.; Moore, K.; Mueller, L.; Noriega, C.; Sejwal, P.; Sheppard, P.; Stevens, B.; Chen, G.; Tymiak, A. A.; Gross, M. L.; Schneeweis, L. A., Mapping the Energetic Epitope of an Antibody/Interleukin-23 Interaction with Hydrogen/Deuterium Exchange, Fast Photochemical Oxidation of Proteins Mass Spectrometry, and Alanine Shave Mutagenesis. *Analytical Chemistry* **2017,** *89* (4), 2250-2258.

Li, K. S.; Chen, G.; Mo, J.; Huang, R. Y. C.; Deyanova, E. G.; Beno, B. R.; O'Neil, S. R.; Tymiak, A. A.; Gross, M. L., Orthogonal Mass Spectrometry-Based Footprinting for Epitope Mapping and Structural Characterization: The IL-6 Receptor upon Binding of Protein Therapeutics. *Analytical Chemistry* **2017,** *89* (14), 7742-7749.

Li, K. S.; Rempel, D. L.; Gross, M. L., Conformational-Sensitive Fast Photochemical Oxidation of Proteins and Mass Spectrometry Characterize Amyloid Beta 1–42 Aggregation. *J. Am. Chem. Soc.* **2016,** *138* (37), 12090-12098.

Li, K. S.; Shi, L.; Gross, M. L., Mass Spectrometry-Based Fast Photochemical Oxidation of Proteins (FPOP) for Higher Order Structure Characterization. *Acc. Chem. Res.* **2018,** *51* (3), 736-744.

Li, X.; Li, Z.; Xie, B.; Sharp, J. S., Improved Identification and Relative Quantification of Sites of Peptide and Protein Oxidation for Hydroxyl Radical Footprinting. *Journal of The American Society for Mass Spectrometry* **2013,** *24* (11), 1767-1776.

Li, Z.; Moniz, H.; Wang, S.; Ramiah, A.; Zhang, F.; Moremen, K. W.; Linhardt, R. J.; Sharp, J. S., High Structural Resolution Hydroxyl Radical Protein Footprinting Reveals an Extended Robo1-Heparin Binding Interface. *Journal of Biological Chemistry* **2015,** *290* (17), 10729-10740.

Lilie, J.; Behar, D.; Sujdak, R. J.; Schuler, R. H., Lifetime of trifluoromethyl radical in aqueous solution. *The Journal of Physical Chemistry* **1972,** *76* (18), 2517-2520.

Limpikirati, P.; Hale, J. E.; Hazelbaker, M.; Huang, Y.; Jia, Z.; Yazdani, M.; Graban, E. M.; Vaughan, R. C.; Vachet, R. W., Covalent labeling and mass spectrometry reveal subtle higher order structural changes for antibody therapeutics. *mAbs* **2019,** *11* (3), 463-476.

Limpikirati, P.; Liu, T.; Vachet, R. W., Covalent labeling-mass spectrometry with non-specific reagents for studying protein structure and interactions. *Methods* **2018,** *144*, 79-93.

Limpikirati, P.; Pan, X.; Vachet, R. W., Covalent Labeling with Diethylpyrocarbonate: Sensitive to the Residue Microenvironment, Providing Improved Analysis of Protein Higher Order Structure by Mass Spectrometry. *Analytical Chemistry* **2019,** *91* (13), 8516-8523.

Lindsley, C. W., New 2017 Data and Statistics for Pharmaceutical Products. *ACS Chemical Neuroscience* **2018,** *9* (7), 1518-1519.

Liu, T.; Limpikirati, P.; Vachet, R. W., Synergistic Structural Information from Covalent Labeling and Hydrogen–Deuterium Exchange Mass Spectrometry for Protein–Ligand Interactions. *Analytical Chemistry* **2019**.

Liu, T.; Marcinko, T. M.; Kiefer, P. A.; Vachet, R. W., Using Covalent Labeling and Mass Spectrometry To Study Protein Binding Sites of Amyloid Inhibiting Molecules. *Analytical Chemistry* **2017,** *89* (21), 11583-11591.

Lu, Y.; Zhang, H.; Niedzwiedzki, D. M.; Jiang, J.; Blankenship, R. E.; Gross, M. L., Fast Photochemical Oxidation of Proteins Maps the Topology of Intrinsic Membrane Proteins: Light-Harvesting Complex 2 in a Nanodisc. *Analytical Chemistry* **2016,** *88* (17), 8827-8834.

Luheshi, L. M.; Crowther, D. C.; Dobson, C. M., Protein misfolding and disease: from the test tube to the organism. *Current Opinion in Chemical Biology* **2008,** *12* (1), 25-31.

Mädler, S.; Bich, C.; Touboul, D.; Zenobi, R., Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. *J. Mass Spectrom.* **2009,** *44* (5), 694-706.

Madsen, J. A.; Yin, Y.; Qiao, J.; Gill, V.; Renganathan, K.; Fu, W.-Y.; Smith, S.; Anderson, J., Covalent Labeling Denaturation Mass Spectrometry for Sensitive

Localized Higher Order Structure Comparisons. *Analytical Chemistry* **2016,** *88* (4), 2478-2488.

Magnenat, L.; Palmese, A.; Fremaux, C.; D'Amici, F.; Terlizzese, M.; Rossi, M.; Chevalet, L., Demonstration of physicochemical and functional similarity between the proposed biosimilar adalimumab MSB11022 and Humira®. *mAbs* **2017,** *9* (1), 127-139.

Majumdar, R.; Middaugh, C. R.; Weis, D. D.; Volkin, D. B., Hydrogen-Deuterium Exchange Mass Spectrometry as an Emerging Analytical Tool for Stabilization and Formulation Development of Therapeutic Monoclonal Antibodies. *Journal of Pharmaceutical Sciences* **2015,** *104* (2), 327-345.

Maleknia, S. D.; Brenowitz, M.; Chance, M. R., Millisecond Radiolytic Modification of Peptides by Synchrotron X-rays Identified by Mass Spectrometry. *Analytical Chemistry* **1999,** *71* (18), 3965-3973.

Maleknia, S. D.; Chance, M. R.; Downard, K. M., Electrospray-assisted modification of proteins: a radical probe of protein structure. *Rapid Communications in Mass Spectrometry* **1999,** *13* (23), 2352-2358.

Maleknia, S. D.; Ralston, C. Y.; Brenowitz, M. D.; Downard, K. M.; Chance, M. R., Determination of Macromolecular Folding and Structure by Synchrotron X-Ray Radiolysis Techniques. *Analytical Biochemistry* **2001,** *289* (2), 103-115.

Maleknia, S. D.; Wong, J. W. H.; Downard, K. M., Photochemical and electrophysical production of radicals on millisecond timescales to probe the structure, dynamics and interactions of proteins. *Photochemical & Photobiological Sciences* **2004,** *3* (8), 741-748.

Malito, E.; Faleri, A.; Lo Surdo, P.; Veggi, D.; Maruggi, G.; Grassi, E.; Cartocci, E.; Bertoldi, I.; Genovese, A.; Santini, L.; Romagnoli, G.; Borgogni, E.; Brier, S.; Lo Passo, C.; Domina, M.; Castellino, F.; Felici, F.; van der Veen, S.; Johnson, S.; Lea, S. M.; Tang, C. M.; Pizza, M.; Savino, S.; Norais, N.; Rappuoli, R.; Bottomley, M. J.; Masignani, V., Defining a protective epitope on factor H binding protein, a key meningococcal virulence factor and vaccine antigen. *Proceedings of the National Academy of Sciences* **2013,** *110* (9), 3304.

Manzi, L.; Barrow, A. S.; Scott, D.; Layfield, R.; Wright, T. G.; Moses, J. E.; Oldham, N. J., Carbene footprinting accurately maps binding sites in protein–ligand and protein–protein interactions. *Nature Communications* **2016,** *7* (1), 13288.

Marcinko, T. M.; Dong, J.; LeBlanc, R.; Daborowski, K. V.; Vachet, R. W., Small molecule-mediated inhibition of β-2-microglobulin-based amyloid fibril formation. *Journal of Biological Chemistry* **2017,** *292* (25), 10630-10638.

Martz, E. Atomic Coordinate Files for Whole IgG1; 1996 Feb 7 [accessed 2018 Aug 4]. https://www.umass.edu/microbio/rasmol/padlan.htm.

McClintock, C.; Kertesz, V.; Hettich, R. L., Development of an Electrochemical Oxidation Method for Probing Higher Order Protein Structure with Mass Spectrometry. *Analytical Chemistry* **2008,** *80* (9), 3304-3317.

McLafferty, F. W.; Breuker, K.; Jin, M.; Han, X.; Infusini, G.; Jiang, H.; Kong, X.; Begley, T. P., Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics. *FEBS Journal* **2007,** *274* (24), 6256-6268.

Mehler, E. L.; Fuxreiter, M.; Simon, I.; Garcia-Moreno E, B., The role of hydrophobic microenvironments in modulating pKa shifts in proteins. *Proteins: Struct., Funct., Bioinf.* **2002,** *48* (2), 283-292.

Mendoza, V. L.; Antwi, K.; Barón-Rodríguez, M. A.; Blanco, C.; Vachet, R. W., Structure of the Preamyloid Dimer of β-2-Microglobulin from Covalent Labeling and Mass Spectrometry. *Biochemistry* **2010,** *49* (7), 1522-1532.

Mendoza, V. L.; Barón-Rodríguez, M. A.; Blanco, C.; Vachet, R. W., Structural Insights into the Pre-Amyloid Tetramer of β-2-Microglobulin from Covalent Labeling and Mass Spectrometry. *Biochemistry* **2011,** *50* (31), 6711-6722.

Mendoza, V. L.; Vachet, R. W., Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spectrometry Reviews* **2009,** *28* (5), 785-815.

Mendoza, V. L.; Vachet, R. W., Protein Surface Mapping Using Diethylpyrocarbonate with Mass Spectrometric Detection. *Analytical Chemistry* **2008,** *80* (8), 2895-2904.

Meng, F.; Forbes, A. J.; Miller, L. M.; Kelleher, N. L., Detection and localization of protein modifications by high resolution tandem mass spectrometry. *Mass Spectrometry Reviews* **2005,** *24* (2), 126-134.

Miles, E. W., Modification of histidyl residues in proteins by diethylpyrocarbonate. In *Methods in Enzymology*, Academic Press: Cambridge, Massachusetts, United States, 1977; Vol. 47, pp 431-442.

Miller, B. T.; Collins, T. J.; Nagle, G. T.; Kurosky, A., The occurrence of O-acylation during biotinylation of gonadotropin-releasing hormone and analogs. Evidence for a reactive serine. *J. Biol. Chem.* **1992,** *267* (8), 5060-9.

Miller, B. T.; Collins, T. J.; Rogers, M. E.; Kurosky, A., Peptide Biotinylation with Amine-Reactive Esters: Differential Side Chain Reactivity. *Peptides* **1997,** *18* (10), 1585-1595.

Miller, B. T.; Kurosky, A., Elevated Intrinsic Reactivity of Seryl Hydroxyl Groups within the Linear Peptide Triads His-Xaa-Ser or Ser-Xaa-His. *Biochem. Biophys. Res. Commun.* **1993,** *196* (1), 461-467.

Miller, J. N.; Miller, J. C., *Statistics and chemometrics for analytical chemistry*. 6th ed.; Pearson/Prentice Hall: Harlow, England, United Kingdom, 2005.

Minkoff, B. B.; Blatz, J. M.; Choudhury, F. A.; Benjamin, D.; Shohet, J. L.; Sussman, M. R., Plasma-Generated OH Radical Production for Analyzing Three-Dimensional Structure in Protein Therapeutics. *Scientific Reports* **2017,** *7* (1), 12946.

Misra, S. K.; Sood, A.; Soares, P. A.; Pomin, V. H.; Woods, R. J.; Sharp, J. S., Mapping of the Fondaparinux Binding Site of JR-FL gp120 by High Resolution Hydroxyl Radical Protein Footprinting and Computational Docking. *bioRxiv* **2017**.

Myers, J. K.; Pace, C. N.; Scholtz, J. M., Helix Propensities Are Identical in Proteins and Peptides. *Biochemistry* **1997,** *36* (36), 10923-10929.

Naganathan, A. N.; Muñoz, V., Scaling of Folding Times with Protein Size. *Journal of the American Chemical Society* **2005,** *127* (2), 480-481.

Negi, S.; Zhu, H.; Fraczkiewicz, R.; Braun, W. Solvent Accessible Surface Areas, Atomic Solvation Energies, and Their Gradients for Macromolecules. 2015 April 17 [accessed 2019 Feb 9]. http://curie.utmb.edu/area_man.html.

Nejadnik, M. R.; Randolph, T. W.; Volkin, D. B.; Schöneich, C.; Carpenter, J. F.; Crommelin, D. J. A.; Jiskoot, W., Postproduction Handling and Administration of

Protein Pharmaceuticals and Potential Instability Issues. *Journal of Pharmaceutical Sciences* **2018,** *107* (8), 2013-2019.

Neugebauer, J. M., Detergents: An overview. In *Methods in Enzymology*, Deutscher, M. P., Ed. Academic Press: 1990; Vol. 182, pp 239-253.

Nölting, B.; Golbik, R.; Neira, J. L.; Soler-Gonzalez, A. S.; Schreiber, G.; Fersht, A. R., The folding pathway of a protein at high resolution from microseconds to seconds. *Proceedings of the National Academy of Sciences* **1997,** *94* (3), 826-830.

Nölting, B.; Golbik, R.; Fersht, A. R., Submillisecond events in protein folding. *Proceedings of the National Academy of Sciences of the United States of America* **1995,** *92* (23), 10668-10672.

Nupur, N.; Chhabra, N.; Dash, R.; Rathore, A. S., Assessment of structural and functional similarity of biosimilar products: Rituximab as a case study. *mAbs* **2018,** *10* (1), 143-158.

Opuni Kwabena, F. M.; Al-Majdoub, M.; Yefremova, Y.; El-Kased Reham, F.; Koy, C.; Glocker Michael, O., Mass spectrometric epitope mapping. *Mass Spectrometry Reviews* **2016,** *37* (2), 229-241.

Padayatti, P. S.; Wang, L.; Gupta, S.; Orban, T.; Sun, W.; Salom, D.; Jordan, S. R.; Palczewski, K.; Chance, M. R., A Hybrid Structural Approach to Analyze Ligand Binding by the Serotonin Type 4 Receptor (5-HT4). *Molecular & Cellular Proteomics : MCP* **2013,** *12* (5), 1259-1271.

Padlan, E. A., Anatomy of the antibody molecule. *Molecular Immunology* **1994,** *31* (3), 169-217.

Pan, L. Y.; Salas-Solano, O.; Valliere-Douglass, J. F., Antibody Structural Integrity of Site-Specific Antibody-Drug Conjugates Investigated by Hydrogen/Deuterium Exchange Mass Spectrometry. *Analytical Chemistry* **2015,** *87* (11), 5669-5676.

Pan, L. Y.; Salas-Solano, O.; Valliere-Douglass, J. F., Conformation and Dynamics of Interchain Cysteine-Linked Antibody-Drug Conjugates as Revealed by Hydrogen/Deuterium Exchange Mass Spectrometry. *Analytical Chemistry* **2014,** *86* (5), 2657-2664.

Pan, L. Y.; Salas-Solano, O.; Valliere-Douglass, J. F., Localized conformational interrogation of antibody and antibody-drug conjugates by site-specific carboxyl group footprinting. *mAbs* **2017,** *9* (2), 307-318.

Pan, Y.; Brown, L.; Konermann, L., Kinetic Folding Mechanism of an Integral Membrane Protein Examined by Pulsed Oxidative Labeling and Mass Spectrometry. *Journal of Molecular Biology* **2011,** *410* (1), 146-158.

Pan, Y.; Piyadasa, H.; O'Neil, J. D.; Konermann, L., Conformational Dynamics of a Membrane Transport Protein Probed by H/D Exchange and Covalent Labeling: The Glycerol Facilitator. *Journal of Molecular Biology* **2012,** *416* (3), 400-413.

Pan, Y.; Ruan, X.; Valvano, M. A.; Konermann, L., Validation of Membrane Protein Topology Models by Oxidative Labeling and Mass Spectrometry. *Journal of The American Society for Mass Spectrometry* **2012,** *23* (5), 889-898.

Peck, M. T.; Ortega, G.; De Luca-Johnson, J. N.; Schlessman, J. L.; Robinson, A. C.; García-Moreno E, B., Local Backbone Flexibility as a Determinant of the Apparent pKa Values of Buried Ionizable Groups in Proteins. *Biochemistry* **2017,** *56* (40), 5338-5346.

Petsko, G. A.; Ringe, D., *Protein Structure and Function*. New Science Press Ltd: London, England, United Kingdom, 2004; pp 10-11.

Pirrone, G. F.; Iacob, R. E.; Engen, J. R., Applications of Hydrogen/Deuterium Exchange MS from 2012 to 2014. *Analytical Chemistry* **2015,** *87* (1), 99-118.

Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M., MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **2010,** *11* (1), 395.

Poor, T. A.; Jones, L. M.; Sood, A.; Leser, G. P.; Plasencia, M. D.; Rempel, D. L.; Jardetzky, T. S.; Woods, R. J.; Gross, M. L.; Lamb, R. A., Probing the paramyxovirus fusion (F) protein-refolding event from pre- to postfusion by oxidative footprinting. *Proceedings of the National Academy of Sciences* **2014,** *111* (25), E2596-E2605.

Rashidzadeh, H.; Khrapunov, S.; Chance, M. R.; Brenowitz, M., Solution Structure and Interdomain Interactions of the Saccharomyces cerevisiae "TATA Binding Protein" (TBP) Probed by Radiolytic Protein Footprinting. *Biochemistry* **2003,** *42* (13), 3655-3665.

Rathore, D.; Faustino, A.; Schiel, J.; Pang, E.; Boyne, M.; Rogstad, S., The role of mass spectrometry in the characterization of biologic protein products. *Expert Review of Proteomics* **2018,** *15* (5), 431-449.

Richards, F. M.; Lamed, R.; Wynn, R.; Patel, D.; Olack, G., Methylene as a possible universal footprinting reagent that will include hydrophobic surface areas: overview and feasibility: properties of diazirine as a precursor. *Protein Science* **2000,** *9* (12), 2506-2517.

Rinas, A.; Espino, J. A.; Jones, L. M., An efficient quantitation strategy for hydroxyl radical-mediated protein footprinting using Proteome Discoverer. *Analytical and Bioanalytical Chemistry* **2016,** *408* (11), 3021-3031.

Robinson, A. C.; Majumdar, A.; Schlessman, J. L.; García-Moreno E, B., Charges in Hydrophobic Environments: A Strategy for Identifying Alternative States in Proteins. *Biochemistry* **2017,** *56* (1), 212-218.

Rogstad, S.; Faustino, A.; Ruth, A.; Keire, D.; Boyne, M.; Park, J., A Retrospective Evaluation of the Use of Mass Spectrometry in FDA Biologics License Applications. *Journal of The American Society for Mass Spectrometry* **2017,** *28* (5), 786-794.

Ruan, Q.; Zhao, C.; Ramsay, C. S.; Tetin, S. Y., Characterization of Fluorescently Labeled Protein with Electrospray Ionization-MS and Fluorescence Spectroscopy: How Random is Random Labeling? *Anal. Chem.* **2018,** *90* (16), 9695-9699.

Schmidt, C.; Macpherson, J. A.; Lau, A. M.; Tan, K. W.; Fraternali, F.; Politis, A., Surface Accessibility and Dynamics of Macromolecular Assemblies Probed by Covalent Labeling Mass Spectrometry and Integrative Modeling. *Analytical Chemistry* **2017,** *89* (3), 1459-1468.

Schorzman, A. N.; Perera, L.; Cutalo-Patterson, J. M.; Pedersen, L. C.; Pedersen, L. G.; Kunkel, T. A.; Tomer, K. B., Modeling of the DNA-binding site of yeast Pms1 by mass spectrometry *DNA repair* **2011,** *10* (5), 454-465.

Schrodinger, LLC *The PyMOL Molecular Graphics System, Version 1.8*, 2015.

Sengupta, U.; Nilson, A. N.; Kayed, R., The Role of Amyloid-β Oligomers in Toxicity, Propagation, and Immunotherapy. *EBioMedicine* **2016,** *6*, 42-49.

Seo, N.; Polozova, A.; Zhang, M.; Yates, Z.; Cao, S.; Li, H.; Kuhns, S.; Maher, G.; McBride, H. J.; Liu, J., Analytical and functional similarity of Amgen biosimilar ABP 215 to bevacizumab. *mAbs* **2018,** *10* (4), 678-691.

Shankar, G. M.; Li, S.; Mehta, T. H.; Garcia-Munoz, A.; Shepardson, N. E.; Smith, I.; Brett, F. M.; Farrell, M. A.; Rowan, M. J.; Lemere, C. A.; Regan, C. M.; Walsh, D. M.; Sabatini, B. L.; Selkoe, D. J., Amyloid-β protein dimers isolated directly from Alzheimer's brains impair synaptic plasticity and memory. *Nature Medicine* **2008,** *14*, 837.

Shannon, D. A.; Weerapana, E., Covalent protein modification: the current landscape of residue-specific electrophiles. *Current Opinion in Chemical Biology* **2015,** *24* (Supplement C), 18-26.

Sharp, J. S.; Becker, J. M.; Hettich, R. L., Analysis of Protein Solvent Accessible Surfaces by Photochemical Oxidation and Mass Spectrometry. *Analytical Chemistry* **2004,** *76* (3), 672-683.

Sharp, J. S.; Becker, J. M.; Hettich, R. L., Protein surface mapping by chemical oxidation: Structural analysis by mass spectrometry. *Analytical Biochemistry* **2003,** *313* (2), 216-225.

Sharp, J. S.; Tomer, K. B., Analysis of the Oxidative Damage-Induced Conformational Changes of Apo- and Holocalmodulin by Dose-Dependent Protein Oxidative Surface Mapping. *Biophysical Journal* **2007,** *92* (5), 1682-1692.

Shastry, M. C. R.; Udgaonkar, J. B., The Folding Mechanism of Barstar: Evidence for Multiple Pathways and Multiple Intermediates. *Journal of Molecular Biology* **1995,** *247* (5), 1013-1027.

Shen, G.; Li, S.; Cui, W.; Liu, S.; Yang, Y.; Gross, M.; Li, W., Membrane Protein Structure in Live Cells: Methodology for Studying Drug Interaction by Mass Spectrometry-Based Footprinting. *Biochemistry* **2018**, Article ASAP.

Shrivastava, A.; Gupta, V., Methods for the determination of limit of detection and limit of quantitation of the analytical methods. *Chronicles of Young Scientists* **2011,** *2* (1), 21-25.

Sinz, A., Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein–protein interactions. *Mass Spectrometry Reviews* **2006,** *25* (4), 663-682.

Siuti, N.; Kelleher, N. L., Decoding protein modifications using top-down mass spectrometry. *Nature Methods* **2007,** *4*, 817.

Srikanth, R.; Mendoza, V. L.; Bridgewater, J. D.; Zhang, G.; Vachet, R. W., Copper Binding to β-2-Microglobulin and Its Pre-Amyloid Oligomers. *Biochemistry* **2009,** *48* (41), 9871-9881.

Srikanth, R.; Wilson, J.; Bridgewater, J. D.; Numbers, J. R.; Lim, J.; Olbris, M. R.; Kettani, A.; Vachet, R. W., Improved Sequencing of Oxidized Cysteine and Methionine Containing Peptides Using Electron Transfer Dissociation. *Journal of the American Society for Mass Spectrometry* **2007,** *18* (8), 1499-1506.

Srikanth, R.; Wilson, J.; Burns, C. S.; Vachet, R. W., Identification of the Copper(II) Coordinating Residues in the Prion Protein by Metal-Catalyzed Oxidation Mass Spectrometry: Evidence for Multiple Isomers at Low Copper(II) Loadings. *Biochemistry* **2008,** *47* (35), 9258-9268.

Srikanth, R.; Wilson, J.; Vachet, R. W., Correct identification of oxidized histidine residues using electron-transfer dissociation. *Journal of Mass Spectrometry* **2009,** *44* (5), 755-762.

Stocks, B. B.; Konermann, L., Structural Characterization of Short-Lived Protein Unfolding Intermediates by Laser-Induced Oxidative Labeling and Mass Spectrometry. *Anal. Chem.* **2009,** *81* (1), 20-27.

Suchanek, M.; Radzikowska, A.; Thiele, C., Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells. *Nature Methods* **2005,** *2*, 261.

Tabb, D. L.; Fernando, C. G.; Chambers, M. C., MyriMatch: Highly Accurate Tandem Mass Spectral Peptide Identification by Multivariate Hypergeometric Analysis. *Journal of Proteome Research* **2007,** *6* (2), 654-661.

Takamoto, K.; Chance, M. R., Radiolytic protein footprinting with mass spectrometry to probe the structure of macromolecular complexes. *Annual Review of Biophysics and Biomolecular Structure* **2006,** *35* (1), 251-276.

The Global Proteome Machine Organization. cRAP protein sequences. 2012 Jan 1 [accessed 2017 Oct 18]. http://www.thegpm.org/crap/index.html.

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

Thurlkill, R. L.; Grimsley, G. R.; Scholtz, J. M.; Pace, C. N., pK values of the ionizable groups of proteins. *Protein Sci.* **2006,** *15* (5), 1214-1218.

Tong, X.; Wren, J. C.; Konermann, L., Effects of Protein Concentration on the Extent of γ-Ray-Mediated Oxidative Labeling Studied by Electrospray Mass Spectrometry. *Analytical Chemistry* **2007,** *79* (16), 6376-6382.

Trinquier, G.; Sanejouand, Y. H., Which effective property of amino acids is best preserved by the genetic code? *Protein Engineering, Design and Selection* **1998,** *11* (3), 153-169.

Tullius, T. D.; Greenbaum, J. A., Mapping nucleic acid structure by hydroxyl radical cleavage. *Current Opinion in Chemical Biology* **2005,** *9* (2), 127-134.

U.S. Food and Drug Administration. Guidance concerning demonstration of comparability of human biological products, including therapeutic biotechnologyderived products. 1996 Apr [accessed 2018 Jul 21]. https://www.fda.gov/drugs/guidancecomplianceregulatoryinformation/guidances/ucm122879.htm.

U.S. Food and Drug Administration. Points to Consider in the Manufacture and Testing of Monoclonal Antibody Products for Human Use. 1997 Feb 28 [accessed 2017 Sep 23]. https://www.fda.gov/downloads/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/OtherRecommendationsforManufacturers/UCM153182.pdf.

U.S. Food and Drug Administration. Quality Considerations in Demonstrating Biosimilarity of a Therapeutic Protein Product to a Reference Product to a Reference Product; 2015 Apr [accessed 2017 Sep 25]. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291134.pdf.

Ureta, D. B.; Craig, P. O.; Gómez, G. E.; Delfino, J. M., Assessing Native and Non-native Conformational States of a Protein by Methylene Carbene Labeling: The Case of Bacillus licheniformis β-Lactamase. *Biochemistry* **2007,** *46* (50), 14567-14577.

Vahidi, S.; Stocks, B. B.; Liaghati-Mobarhan, Y.; Konermann, L., Mapping pH-Induced Protein Structural Changes Under Equilibrium Conditions by Pulsed Oxidative Labeling and Mass Spectrometry. *Analytical Chemistry* **2012,** *84* (21), 9124-9130.

Vahidi, S.; Stocks, B. B.; Liaghati-Mobarhan, Y.; Konermann, L., Submillisecond Protein Folding Events Monitored by Rapid Mixing and Mass Spectrometry-Based Oxidative Labeling. *Analytical Chemistry* **2013,** *85* (18), 8618-8625.

Vahidi, S.; Konermann, L., Probing the Time Scale of FPOP (Fast Photochemical Oxidation of Proteins): Radical Reactions Extend Over Tens of Milliseconds. *Journal of The American Society for Mass Spectrometry* **2016,** *27* (7), 1156-1164.

van der Kant, R.; Karow-Zwick, A. R.; Van Durme, J.; Blech, M.; Gallardo, R.; Seeliger, D.; Aßfalg, K.; Baatsen, P.; Compernolle, G.; Gils, A.; Studts, J. M.; Schulz, P.; Garidel, P.; Schymkowitz, J.; Rousseau, F., Prediction and Reduction of the Aggregation of Monoclonal Antibodies. *Journal of Molecular Biology* **2017,** *429* (8), 1244-1261.

Vaudel, M.; Barsnes, H.; Berven, F. S.; Sickmann, A.; Martens, L., SearchGUI: An open-source graphical user interface for simultaneous OMSSA and X!Tandem searches. *Proteomics* **2011,** *11* (5), 996-999.

Vaudel, M.; Burkhart, J. M.; Zahedi, R. P.; Oveland, E.; Berven, F. S.; Sickmann, A.; Martens, L.; Barsnes, H., PeptideShaker enables reanalysis of MS-derived proteomics data sets. *Nat Biotech* **2015,** *33* (1), 22-24.

Verdone, G.; Corazza, A.; Viglino, P.; Pettirossi, F.; Giorgetti, S.; Mangione, P.; Andreola, A.; Stoppini, M.; Bellotti, V.; Esposito, G., The solution structure of human β2-microglobulin reveals the prodromes of its amyloid transition. *Protein Sci.* **2002,** *11* (3), 487-499.

Verma, M.; Vats, A.; Taneja, V., Toxic species in amyloid disorders: Oligomers or mature fibrils. *Annals of Indian Academy of Neurology* **2015,** *18* (2), 138-145.

Vermeer, A. W.; Norde, W., The thermal stability of immunoglobulin: unfolding and aggregation of a multi-domain protein. *Biophysical Journal* **2000,** *78* (1), 394-404.

Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J., Structure of ubiquitin refined at 1.8 Å resolution. *J. Mol. Biol.* **1987,** *194* (3), 531-544.

Wales, T. E.; Engen, J. R., Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrometry Reviews* **2006,** *25* (1), 158-170.

Wang, L.; Qin, Y.; Ilchenko, S.; Bohon, J.; Shi, W.; Cho, M. W.; Takamoto, K.; Chance, M. R., Structural Analysis of a Highly Glycosylated and Unliganded gp120-Based Antigen Using Mass Spectrometry. *Biochemistry* **2010,** *49* (42), 9032-9045.

Wang, L.; Chance, M. R., Protein footprinting comes of age: mass spectrometry for biophysical structure assessment. *Molecular & Cellular Proteomics* **2017**.

Wang, L.; Chance, M. R., Structural Mass Spectrometry of Proteins Using Hydroxyl Radical Based Protein Footprinting. *Analytical Chemistry* **2011,** *83* (19), 7234-7241.

Wang, W.; Singh, S.; Zeng, D. L.; King, K.; Nema, S., Antibody Structure, Instability, and Formulation. *Journal of Pharmaceutical Sciences* **2007,** *96* (1), 1-26.

Wang, X.; Das, T. K.; Singh, S. K.; Kumar, S., Potential aggregation prone regions in biotherapeutics: A survey of commercial monoclonal antibodies. *mAbs* **2009,** *1* (3), 254-267.

Watkinson, T. G.; Calabrese, A. N.; Ault, J. R.; Radford, S. E.; Ashcroft, A. E., FPOP-LC-MS/MS Suggests Differences in Interaction Sites of Amphipols and Detergents with Outer Membrane Proteins. *Journal of The American Society for Mass Spectrometry* **2017,** *28* (1), 50-55.

Watson, C.; Janik, I.; Zhuang, T.; Charvátová, O.; Woods, R. J.; Sharp, J. S., Pulsed Electron Beam Water Radiolysis for Submicrosecond Hydroxyl Radical Protein Footprinting. *Analytical Chemistry* **2009,** *81* (7), 2496-2505.

Watson, C.; Sharp, J. S., Conformational Analysis of Therapeutic Proteins by Hydroxyl Radical Protein Footprinting. *The AAPS Journal* **2012,** *14* (2), 206-217.

Wecksler, A. T.; Kalo, M. S.; Deperalta, G., Mapping of Fab-1:VEGF Interface Using Carboxyl Group Footprinting Mass Spectrometry. *Journal of The American Society for Mass Spectrometry* **2015,** *26* (12), 2077-2080.

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.

Weiss, W. F.; Gabrielson, J. P.; Al-Azzam, W.; Chen, G.; Davis, D. L.; Das, T. K.; Hayes, D. B.; Houde, D.; Singh, S. K., Technical Decision Making With Higher Order Structure Data: Perspectives on Higher Order Structure Characterization From the Biopharmaceutical Industry. *Journal of Pharmaceutical Sciences* **2016,** *105* (12), 3465-3470.

Whitford, D., *Proteins: Structure and Function*. John Wiley & Sons Ltd: West Sussex, England, United Kingdom, 2005; pp 53-58.

Wong, J. W. H.; Maleknia, S. D.; Downard, K. M., Hydroxyl radical probe of the calmodulin-melittin complex interface by electrospray ionization mass spectrometry. *Journal of the American Society for Mass Spectrometry* **2005,** *16* (2), 225-233.

Wong, J. W. H.; Maleknia, S. D.; Downard, K. M., Study of the Ribonuclease−S-Protein−Peptide Complex Using a Radical Probe and Electrospray Ionization Mass Spectrometry. *Analytical Chemistry* **2003,** *75* (7), 1557-1563.

Wong, K.-B.; Freund, S. M. V.; Fersht, A. R., Cold Denaturation of Barstar:1H,15N and13C NMR Assignment and Characterisation of Residual Structure. *Journal of Molecular Biology* **1996,** *259* (4), 805-818.

Xie, B.; Sood, A.; Woods, R. J.; Sharp, J. S., Quantitative Protein Topography Measurements by High Resolution Hydroxyl Radical Protein Footprinting Enable Accurate Molecular Model Selection. *Scientific Reports* **2017,** *7* (1), 4552.

Xu, G.; Kiselar, J.; He, Q.; Chance, M. R., Secondary Reactions and Strategies To Improve Quantitative Protein Footprinting. *Analytical Chemistry* **2005,** *77* (10), 3029-3037.

Xu, G.; Chance, M. R., Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics. *Chemical Reviews* **2007,** *107* (8), 3514-3543.

Xu, G.; Chance, M. R., Radiolytic Modification and Reactivity of Amino Acid Residues Serving as Structural Probes for Protein Footprinting. *Analytical Chemistry* **2005,** *77* (14), 4549-4555.

Yu, C.; Huang, L., Cross-Linking Mass Spectrometry: An Emerging Technology for Interactomics and Structural Biology. *Analytical Chemistry* **2018,** *90* (1), 144-165.

Zhang, A.; Singh, S. K.; Shirts, M. R.; Kumar, S.; Fernandez, E. J., Distinct Aggregation Mechanisms of Monoclonal Antibody Under Thermal and Freeze-Thaw Stresses Revealed by Hydrogen Exchange. *Pharmaceutical Research* **2012,** *29* (1), 236-250.

Zhang, B.; Cheng, M.; Rempel, D.; Gross, M. L., Implementing fast photochemical oxidation of proteins (FPOP) as a footprinting approach to solve diverse problems in structural biology. *Methods* **2018,** *144*, 94-103.

Zhang, B.; Rempel, D. L.; Gross, M. L., Protein Footprinting by Carbenes on a Fast Photochemical Oxidation of Proteins (FPOP) Platform. *Journal of The American Society for Mass Spectrometry* **2016,** *27* (3), 552-555.

Zhang, H.; Cui, W.; Gross, M. L., Mass spectrometry for the biophysical characterization of therapeutic monoclonal antibodies. *FEBS Letters* **2014,** *588* (2), 308-317.

Zhang, H.; Gau, B. C.; Jones, L. M.; Vidavsky, I.; Gross, M. L., Fast Photochemical Oxidation of Proteins for Comparing Structures of Protein−Ligand Complexes: The Calmodulin−Peptide Model System. *Anal. Chem.* **2011,** *83* (1), 311-318.

Zhang, H.; Song, L.; Ye, H.; Hu, L.; Liang, W.; Liu, D., Characterization of a Novel Humanized Anti-CD20 Antibody with Potent Anti-Tumor Activity against Non-Hodgkin's Lymphoma. *Cellular Physiology and Biochemistry* **2013,** *32* (3), 645-654.

Zhang, Q.; Willison, L. N.; Tripathi, P.; Sathe, S. K.; Roux, K. H.; Emmett, M. R.; Blakney, G. T.; Zhang, H.-M.; Marshall, A. G., Epitope Mapping of a 95 kDa Antigen in Complex with Antibody by Solution-Phase Amide Backbone Hydrogen/Deuterium Exchange Monitored by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Analytical Chemistry* **2011,** *83* (18), 7129-7136.

Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M.-C.; Yates, J. R., Protein Analysis by Shotgun/Bottom-up Proteomics. *Chemical reviews* **2013,** *113* (4), 2343-2394.

Zhang, Y.; Wecksler, A. T.; Molina, P.; Deperalta, G.; Gross, M. L., Mapping the Binding Interface of VEGF and a Monoclonal Antibody Fab-1 Fragment with Fast Photochemical Oxidation of Proteins (FPOP) and Mass Spectrometry. *Journal of The American Society for Mass Spectrometry* **2017,** *28* (5), 850-858.

Zhang, Z.; Browne, S. J.; Vachet, R. W., Exploring Salt Bridge Structures of Gas-Phase Protein Ions using Multiple Stages of Electron Transfer and Collision Induced Dissociation. *J. Am. Soc. Mass Spectrom.* **2014,** *25* (4), 604-613.

Zhao, Bo, "PROTEIN DETECTION AND STRUCTURAL CHARACTERIZATION BY MASS SPECTROMETRY USING SUPRAMOLECULAR ASSEMBLIES AND SMALL MOLECULES" (2019). Doctoral Dissertations. 1761.

Zheng, X.; Wintrode, P. L.; Chance, M. R., Complementary Structural Mass Spectrometry Techniques Reveal Local Dynamics in Functionally Important Regions of a Metastable Serpin. *Structure* **2008,** *16* (1), 38-51.

Zhou, Y.; Wu, Y.; Yao, M.; Liu, Z.; Chen, J.; Chen, J.; Tian, L.; Han, G.; Shen, J.-R.; Wang, F., Probing the Lysine Proximal Microenvironments within Membrane

Protein Complexes by Active Dimethyl Labeling and Mass Spectrometry. *Anal. Chem.* **2016,** *88* (24), 12060-12065.

Zhou, Y.; Vachet, R. W., Diethylpyrocarbonate Labeling for the Structural Analysis of Proteins: Label Scrambling in Solution and How to Avoid It. *Journal of The American Society for Mass Spectrometry* **2012,** *23* (5), 899-907.

Zhou, Y.; Vachet, R. W., Increased Protein Structural Resolution from Diethylpyrocarbonate-based Covalent Labeling and Mass Spectrometric Detection. *Journal of The American Society for Mass Spectrometry* **2012,** *23* (4), 708-717.

Ziemianowicz, D. S.; Bomgarden, R.; Etienne, C.; Schriemer, D. C., Amino Acid Insertion Frequencies Arising from Photoproducts Generated Using Aliphatic Diazirines. *Journal of The American Society for Mass Spectrometry* **2017,** *28* (10), 2011-2021.