PROTEASE-CONTAINING MEMBRANES FOR RAPID, CONTROLLED ANTIBODY DIGESTION PRIOR TO MASS SPECTROMETRY ANALYSIS

By

Yongle Pang

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Chemistry-Doctor of Philosophy

2017

ABSTRACT

PROTEASE-CONTAINING MEMBRANES FOR RAPID, CONTROLLED ANTIBODY DIGESTION PRIOR TO MASS SPECTROMETRY ANALYSIS

By

Yongle Pang

Monoclonal antibodies are the fastest growing class of therapeutic drugs because of their high specificities to targeQt cells. Facile analysis of therapeutic mAbs and their post-translational modifications (PTMs) is essential for quality control, and mass spectrometry (MS) is the most powerful tool for antibody characterization. Conventional antibody characterization workflows contain an in-solution digestion step, which is labor-intensive and time-consuming. Protease-containing membranes are an attractive alternative platform for protein digestion because of their high local enzyme concentrations, short radial diffusion distances, rapid convection in pores, simple fabrication and low cost. Additionally, variation of protein residence time in the membrane gives control over the size of proteolytic peptides. This research focuses on developing workflows for monoclonal antibody characterization using functionalized porous membranes.

Sequential adsorption of poly (styrene sulfonate) and pepsin in a porous nylon membrane forms a pepsin membrane reactor. Pepsin is inexpensive and catalyzes proteolysis in acidic solutions, which avoids the need to alkylate cysteine residues and limits antibody deamidation. Variation of the residence times (3 ms to 3 s) of antibody solutions in pepsin-containing membranes yields “bottom-up” (1-2 kDa) to “middle-down” (5-15 kDa) peptides in less than 10 min. These peptic peptides cover the entire sequences of Herceptin and a Waters^{TM} antibody. Compared with the performance of bottom-up (in-solution tryptic digestion) and top-down (intact protein

fragmentation) analysis of an antibody light chain, middle-down (in-membrane peptic digestion) analysis gives the highest bond cleavage (99%). In-membrane digestion also facilitates detection of PTMs such as oxidation, deamidation, N-terminal pyroglutamic acid formation and glycosylation.

Recently developed protease-containing spin membranes provide an excellent platform for rapid, membrane-based protein digestion prior to ultrahigh-resolution Orbitrap MS analysis. Centrifugation of 100-200 μL of pretreated protein solutions through the pepsin- or trypsin-containing membranes takes less than 1 min and gives nearly 100% coverage of the protein sequences in subsequent direct infusion MS analysis of digests of apomyoglobin and four commercial monoclonal antibodies (Herceptin, Avastin, Rituxan and Vectibix). MS analysis of peptic and tryptic peptides also reveals mAb PTMs such as N-terminal pyroglutamate formation, C-terminal Lysine clipping and glycosylation. Liquid chromatography coupled to tandem mass spectrometry analysis of tryptic spin digests and subsequent MaxQuant data searching show 100% sequence coverage of all four antibody light chains, and 75.1%-98.4% coverage of the heavy chains. Compared to in-solution tryptic digestion of mAbs, spin digestion yields higher sequence coverage and a larger number of unique peptides.

In-membrane digestion also facilitates protein sequence comparison. Rapid peptic in-membrane digestion of two antibodies with direct infusion MS analysis accurately reveals the antibody modification site in less than 1 h. Overall, membrane-based protein digestion uses minimal sample preparation time and yields high peptide and sequence coverages for identification of protein PTMs.

I dedicate this dissertation to my parents, Chunming Pang and Jinlan Han,
for their love and support.

ACKNOWLEDGEMENTS

First I would like to thank my advisor Prof. Merlin Bruening for his continuous guidance and support over the past five years. He is very knowledgeable and helpful. His patience and trust encouraged me to overcome the obstacles along the way of pursuing my Ph.D. Merlin has excellent presentation skills, and he is always eager to share his experience of delivering complicated content to the audiences with different backgrounds. It is my honor to conduct my Ph.D. research with him, and I am very proud to be one of the alumni of the Bruening group.

I want to give special thanks to Prof. Gavin Reid, who taught me the fundamental knowledge of mass spectrometry. I still remember the tips and tricks he taught me for mass spectrometry. Most of my MS data collection was conducted in Gavin's lab. His enthusiasm for doing research inspired me to learn mass spectrometry step by step. I would also like to thank Prof. Daniel Jones, who gave me a lot of nice suggestions for doing research, getting prepared for conferences, and job hunting. I wish to thank Prof. Liangliang Sun for training me how to use data searching software. I also wish to thank Prof. Dana Spence for suggestions on this project.

I would like to thank my collaborators, Prof. Donald Hunt and Dr. Weihan Wang for the antibody characterization project. Weihan was graduated from Bruening's lab and gave me so much help in solving technical problems. I want to thank Dr. Gia Jokhadze from Clontech for manufacturing the spin columns, Dr. Mohammad Muhsin Chisti for providing the antibody samples, and Weijing Liu for the contribution to this dissertation. Without these collaborators' help, I cannot finish the projects described here.

I am also very grateful to all my fellow group members, including those who graduated before me and those who are currently working in the lab. In my early research stages, Dr. Yujing Tan and Dr. Jinlan Dong provided me so much help and brought me up to speed with my research. Yujing was very smart and hardworking. He is always a good example for me. Dr. Chao Cheng picked me up from the airport when I arrived in the US. I will always remember the the fun time we had together, and the sincere help from Bruening's lab members.

Lastly, I wish to express my deepest appreciation to my family and friends for their love and support. I am so fortunate to have my wonderful parents in my life. They always support me to chase my dream and do what I love to do. I would also like to thank my friends from Chinese Students and Scholar Association. The spectacular events we held and the wonderful time we had is a valuable memory for me. For my friends who have accompanied me in these years, I truly appreciate your continuous support and understanding. Our friendship will last no matter where we are.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

ACN	Acetonitrile
ADCC	Antibody-dependent cellular cytotoxicity
ADCs	Antibody-drug conjugates
BLA	Biologics License Application
CAD	Collision activation dissociation
CDC	Complement-dependent cytotoxicity
CDRs	Complementarity determining regions
CESI-MS/MS	Capillary electrophoresis-tandem mass spectrometry
CH1, CH2, and CH3	heavy chain constant region 1, 2 and 3
CHAMPS	Complete Homology-Assisted MS/MS Protein Sequencing
CHO	Chinese hamster ovary
CI	Chemical ionization
CID	Collision-induced dissociation
CL	light chain constant region
CNBr	Cyanogen bromide
CRBPs	Cellular retinol binding proteins
CRM	Charged residue model
CS	Chitosan

DHB	2, 5-dihydroxy benzoic acid
DTT	Dithiothreitol
ECD	Electron capture dissociation
EDC	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride
EI	Electron ionization
ESI	Electrospray ionization
ETD	Electron transfer dissociation
FA	Formic acid
FDA	Food and Drug Administration
FTICR	Fourier transform ion cyclotron resonance
HA	Hyaluronic acid
Hc	Heavy chain
HCD	Higher-energy C-trap dissociation
HCl	Hydrochloric acid
HDX	Hydrogen deuterium exchange
HOAc	acetic acid
IAM	Iodoacetamide
IdeS	*Streptococcus pyogenes*
IEM	Ion evaporation model
IgG	Immunoglobulin G

IMER	Immobilized enzyme reactors
INN	International Nonproprietary Name
IPA	Isopropyl Alcohol
Lc	Light chain
LC/MS	Liquid chromatography coupled to mass spectrometry
LIT	Linear ion trap
m/z	Mass to charge
MAA	Marketing Authorization Application
mAb	Monoclonal antibody
MALDI	Matrix-assisted laser desorption/ionization
MS	Mass spectrometry
MS/MS	tandem mass spectrometry
nanoESI	Nanoelectrospray
NEM	N-ethylmaleimide
NHS	N-hydroxysuccinimide
NTCB	2-nitro-5-thiocyanobenzoate
OmpT	Outer membrane protease T
PAA	Poly (acrylic acid)
PNGaseF	N-Glycosidase F
PSS	Polystyrene sulfonate

PTM	Post-translational modifications
Q	Quadrupole
RP-HPLC	Reverse-phase liquid chromatography
S/N	Signal to noise
SPS	Shotgun protein sequencing
TCEP	Tris (2-carboxyethyl) phosphine hydrochloride
TFA	Trifluoroacetic acid
TOF	Time of flight
UPLC-ESI-QTOF-MS	Ultra-performance liquid chromatography coupled with electrospray ionization quadrupole time-of-flight mass spectrometry
UVPD	ultraviolet photodissociation
VH	heavy chain variable region
VL	light chain variable region

Chapter 1 . Introduction

This dissertation describes the fabrication and use of functionalized porous membranes for protein digestion prior to mass spectrometry (MS) analysis. More specifically, my research focuses on developing a workflow that uses pepsin-containing membranes to controllably digest monoclonal antibodies (mAbs) prior to MS analysis of their posttranslational modifications (PTMs) and sequences. Additionally, I investigate enzyme-containing spin membranes for rapid protein digestion and apply controlled membrane-based digestion to analyze protein sequences. To give a foundation for my work, this chapter provides important background on both MS and protein digestion. The introduction starts with an overview of MS-based protein analysis methods, summarizing ionization methods, mass analyzers and tandem mass spectrometry techniques that I used in my research. Then, I briefly review the mAb analysis methods (bottom-up, top-down and middle-down strategies). The third part of the introduction describes protein digestion techniques, including traditional enzymes for protein digestion, techniques for accelerating protein digestion, and membrane-based protein digestion technology. Finally, this chapter outlines the subsequent chapters in this dissertation.

1.1 Mass spectrometry for protein analysis

Over the past two decades, MS has advanced tremendously and become one of the most powerful analytical methods for characterizing biomolecules.[1-4] MS exhibits high sensitivity, mass accuracy and resolution, along with high throughput and wide dynamic ranges.[5] With the completion of the human genome project in 2003, researchers began to focus on the protein set produced or modified by an organism (proteomics), and the metabolites within cells, biofluids and tissues (metabolomics). MS plays a crucial role in proteomics studies because it can give

high-accuracy protein masses. Moreover, tandem mass spectrometry (MS/MS) can fragment the ions of interests to provide information on the protein sequence and PTMs.[6] Additionally, current MS methods can quantify thousands of proteins from complex samples, which is crucial for proteomic studies.[7] However, the extreme complexity of a proteome makes MS analysis challenging and continues to push the development of new MS instrumentation, robust analytical methodologies, laborsaving sample-preparation techniques, and user-friendly bioinformatics tools.

Generally speaking, MS gives information on the masses of peptides and proteins. However, it does not directly determine the masses of molecules, but rather gives the mass to charge ratio (*m/z*) of ions. A MS experiment includes conversion of molecules to ions in the ionization source, separation of these ions (according to their *m/z* value) in the mass analyzer, detection of ions via electrical signals that depend on the ion abundance, and finally, processing the signals and producing a mass spectrum. This section briefly introduces ionization processes and mass analyzers that are fundamental to my work.

1.1.1 Ionization techniques

The inventions of matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI) in the late 1980s greatly enhanced the scope of protein analysis using MS.[8] Before that time, MS techniques such as electron ionization (EI) and chemical ionization (CI) could only ionize volatile analytes. Desorption/ionization techniques, including plasma desorption, fast atom bombardment or laser desorption, solve the ion production problem but suffer from poor signal to noise (S/N) ratios and show low intensities for compounds with molecular masses above 10,000. The development of two soft ionization techniques, MALDI

and ESI, reshaped the MS field. The method developers Koichi Tanaka and John Fenn received the Nobel Prize in chemistry in 2002 "*for their development of soft desorption ionization methods for mass spectrometric analyses of biological macromolecules*".

1.1.1.1 Matrix-assisted laser desorption/ionization

In 1988 Karas and Hillenkamp reported that matrix-ultraviolet laser desorption generates intact molecular ions of lysozyme, *β*-lactoglobulin A, porcine trypsin and albumin.[9] The molar ratio of the organic molecule matrix to analyte in such experiments is around 5000:1. The excess matrix minimizes intramolecular interactions among the analyte molecules as well as damage to the analyte during the ionization process. Matrices such as 2, 5-dihydroxy benzoic acid (DHB) rapidly absorb the laser energy at a certain wavelength, normally 337 or 355 nm, to give an explosive breakdown of the analyte-matrix mixture and send these molecules into the gas phase. Initially, I analyzed proteins using a Thermo LTQ XL ion-trap mass spectrometer equipped with a vMALDI source. However, matrix molecules ionize along with the analyte, usually in the low mass range (<500 m/z), which makes characterization of small peptides challenging.[10] Also, the LTQ XL contains a linear ion trap, which has a maximum *m/z* cutoff of 4000. Considering that MALDI gives predominant singly charged peptide ions,[11,12] and many of the peptides I analyze have molecular masses greater than 4000, I subsequently chose to use ESI as the primary ionization method. Nevertheless, MALDI has better tolerance towards salts and other contaminants than ESI.[13]

1.1.1.2 Electrospray ionization

Also in 1988, John Fenn published his seminal work on using electrospray to form large-molecule ions with multiple charged states.[14,15] In ESI, dissolved analyte enters a capillary, and

the electric field (~10^6 V/m) between the spray capillary and a counter electrode creates fine, highly charged droplets containing the analyte of interest. The dimensions of the droplets decrease with solvent evaporation, which increases the repulsion force between the charges in the droplet. The droplet continues to decrease in size until the repulsion force overcomes the surface tension, which results in droplet fission. At this point, the droplet releases a series of highly charged, tiny droplets. The Rayleigh equation, Eq (1-1), describes the maximum charge, Q_{Ry}, at which a droplet is stable.

$$Q_{Ry} = 8\pi(\varepsilon_0 \gamma R^3)^{1/2} \quad (1\text{-}1)$$

In this equation, ε_0 is the electrical permittivity; γ is the surface tension of the solvent, and R is the radius of the droplet.[16] The fission process continues until very small droplets form, and these are the precursors of the gas-phase ions. There are two mechanisms proposed for the formation of gas-phase ions from small droplets, the ion evaporation model (IEM) and the charged residue model (CRM).[17] Figure 1.1 shows the two models. The IEM predicts that ion emission will occur when the droplet radius decreases to 10 nm, because at that point the field strength at the droplet surface is sufficiently large that Coulombic repulsion overcomes the energy required to increase the droplet surface area and expel an ion. Data for small ionic analytes support the IEM, but large analytes such as proteins likely form ions through the CRM.[16]

In the CRM, solvent evaporation results in an increase in the electric field strength, which is large enough (at the highest surface curvature) to form a Taylor Cone that emits small, highly charged droplets. Repetition of this process results in droplets containing only one analyte, and gas-phase ions form through evaporation and declustering of these final droplets.

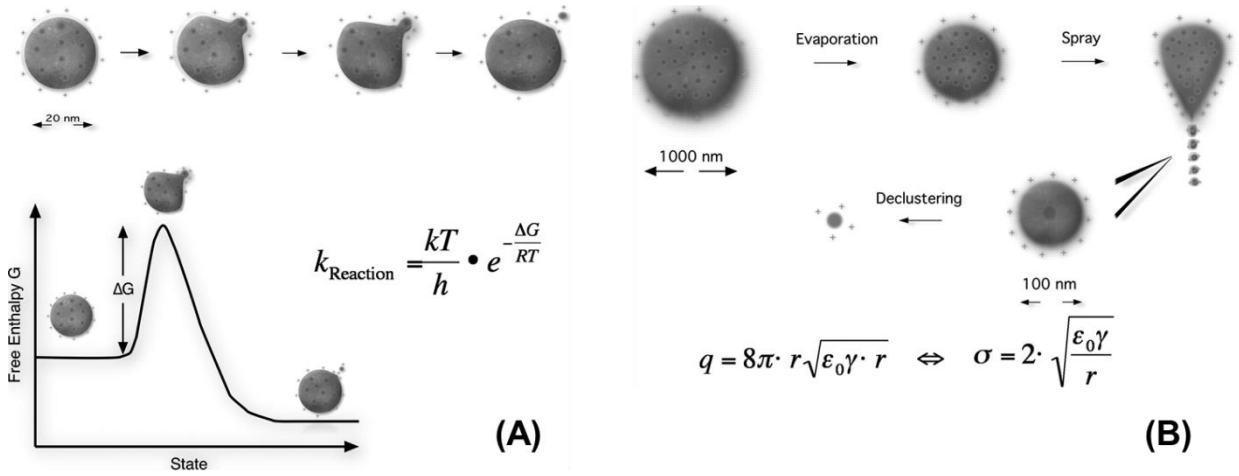


Figure 1.1. The ion-evaporation model and the charged residue model. (A) The ion-evaporation model. An individual ion leaves the charged droplet in a solvated state. [Acronyms: k_{Reaction}, reaction rate constant; k, Boltzmann constant; T, temperature; h, Planck's constant, and R, ideal gas constant.] (B) The charged residue model. As solvent evaporates, smaller and smaller droplets form from a Taylor cone. Finally, droplets contain only one ion, and declustering or evaporation lead to the desolvated ion. [Acronyms: q, droplet charge at the Rayleigh instability limit; r, droplet radius; ε_0, electric permittivity of the surrounding medium; γ, surface tension, and σ, surface charge density.] Figure copied (with permission) from Matthias Wilm. Principles of electrospray ionization. *Mol. & Cell Proteomics*. 2011; Vol 10: M111.009407. © the American Society for Biochemistry and Molecular Biology.

ESI normally operates in three modes: direct infusion, nanospray infusion, or electrospray infusion coupled with liquid chromatography (LC/MS).[10] Direct infusion ESI uses a syringe pump to introduce the sample to the ion source at a flow rate of several μL/min. Nanospray employs a much lower flow rate with a pressurized, special nozzle.[18] Commercially, the TriVersa NanoMate from Advion (Ithaca, NY) is a chip-based nanospray device. I conducted most of my work using the NanoMate, because a few μL of sample can last for more than 30 min of

electrospray. The chip contains an array of nanoelectrospray (nanoESI) nozzles, and each nozzle is one-fifth the diameter of a hair.

The most common mode for analyte introduction into the mass spectrometer, particular for mixtures, is LC-ESI-MS, where a protein or peptide mixture separates during flow through an LC column prior to ESI. This online sample analysis is advantageous for proteomics studies because offline fractionation causes significant sample loss. MS/MS of the peptides in the mass spectrometer provides sequence information, and modern software can match MS/MS data to peptide sequences in a database. As a well-established workflow, LC-ESI-MS/MS provides fast and robust qualitative and quantitative (with appropriate isotopic labeling or label-free methods) analysis of peptides.

In contrast to MALDI, which predominantly forms singly charged ions, ESI gives multiply charged peptide or protein ions to enable characterization of peptides or proteins in the mass range of common mass spectrometers. The maximum charge state of a peptide typically corresponds well with the number of amino acids that can accept a proton (Lysine, Arginine, Histidine, and the N-terminus).[10] Most of my data result from nanoESI-MS or LC-MS/MS analyses.

1.1.2 Mass analyzers

After the ionization process, mass analyzers separate gas-phase ions based on their *m/z* values. The mass analyzer determines the resolution, accuracy, mass range, scan speed, ion transmission, and MS/MS capabilities of a mass spectrometer.[19] The two main categories of mass analyzers include the scanning type in quadrupole (Q) and time of flight (TOF) instruments, and the trapping type in linear ion trap (LIT), Fourier transform ion cyclotron resonance (FTICR) and

Orbitrap instruments. State-of-the-art mass spectrometers often combine multiple mass analyzers to perform different types of experiments in one instrument. Also the arrangement of mass analyzers results in different instrument performance. In the following subsections, I briefly introduce the TOF, LIT and Orbitrap mass analyzers that I used to obtain MS and MS/MS data.

1.1.2.1 Time of flight mass analyzers

Time of flight analyzers differentiate ions according to their velocities. An electric field first accelerates ions to convert their electrical potential energy into kinetic energy. Eq (1-2) describes the acceleration process

$$E_k = \frac{mv^2}{2} = qV = zeV = E_e \qquad (1\text{-}2)$$

where E_k is the ion kinetic energy; m is the mass of the ion; v is the ion velocity; q is the total charge on the ion; z is the ion charge state; V is the electrical potential drop, and E_e is the electric potential energy. Eq (1-2) shows that the ion velocity depends on it mass-to-charge ratio. The TOF detector separates ions in time, as species with small m/z ratios reach the detector first.

The first TOF analyzers were linear instruments that suffered from low resolution because of the range of kinetic energies gained by the same ions. Further development of TOF instruments with delayed pulsed extraction and reflectrons greatly enhanced resolution.[10] TOF analyzers can couple with quadrupoles (QTOF) to give a robust hybrid instrument with the high resolution of the TOF analyzer and the MS/MS capability of the quadrupole.

1.1.2.2 Linear ion traps

The linear ion trap is also known as a 2D ion trap. Thermo Scientific terms it an LTQ. In a LIT, the radial quadrupolar oscillating electric field and the axial DC electric field control ion trajectories.[20] After the ions fly into the LIT, cooling occurs through collisions with inert gas, and the ions collect between the two ends of the trap (*z* axis). At the same time, radiofrequency potentials applied to the rods cause the ions to oscillate in the *xy* plane. Compared with a 3D ion trap (Paul ion trap), the LIT has ten times the ion-trapping capacity. The trapping efficiency increases from 5% for a 3D ion trap to 50% for LITs.[19] Also, the larger internal space in the LIT solves the space-charge problem in the 3D ion trap. LIT analyzers have excellent MS/MS capabilities, and multiple-stage tandem mass spectrometry (MS^n) is also possible, where a fragment ion is isolated for further fragmentation, which usually occurs through collision-induced dissociation (CID).[21] I will talk more fragmentation methods in later sections. Though the resolution of the LIT is not high, its relatively low price, robust operation, and easy maintenance, make it a very popular mass analyzer.[22]

1.1.2.3 Orbitrap analyzers

Developed by Alexander Makarov, the Orbitrap is now a leading high resolution mass spectrometry analyzer in the MS field.[23] FTICR instruments were the most common ultra-high resolution mass spectrometers before the invention of the Orbitrap. However, FTICR requires a superconducting magnet, which results in a high cost for operation and routine maintenance. Orbitrap analyzers can give a mass accuracy of sub parts per million,[24] and have a maximum resolving power of 500,000 at *m/z* 200 (for Orbitrap Fusion Lumos Tribid mass spectrometers). Figure 1.2 illustrates the advantage of an LTQ-Orbitrap analyzer versus an LTQ analyzer in

resolving a 6+ charged ion. With the LTQ, one can determine a centroid mass of the ion, but not the charge state. The Orbitrap spectrum enables determination of the monoisotopic mass as well as the charge state (from the separation of the isotopic peaks). An accurate monoisotopic mass is vital for accurately identifying peptides without MS/MS data.

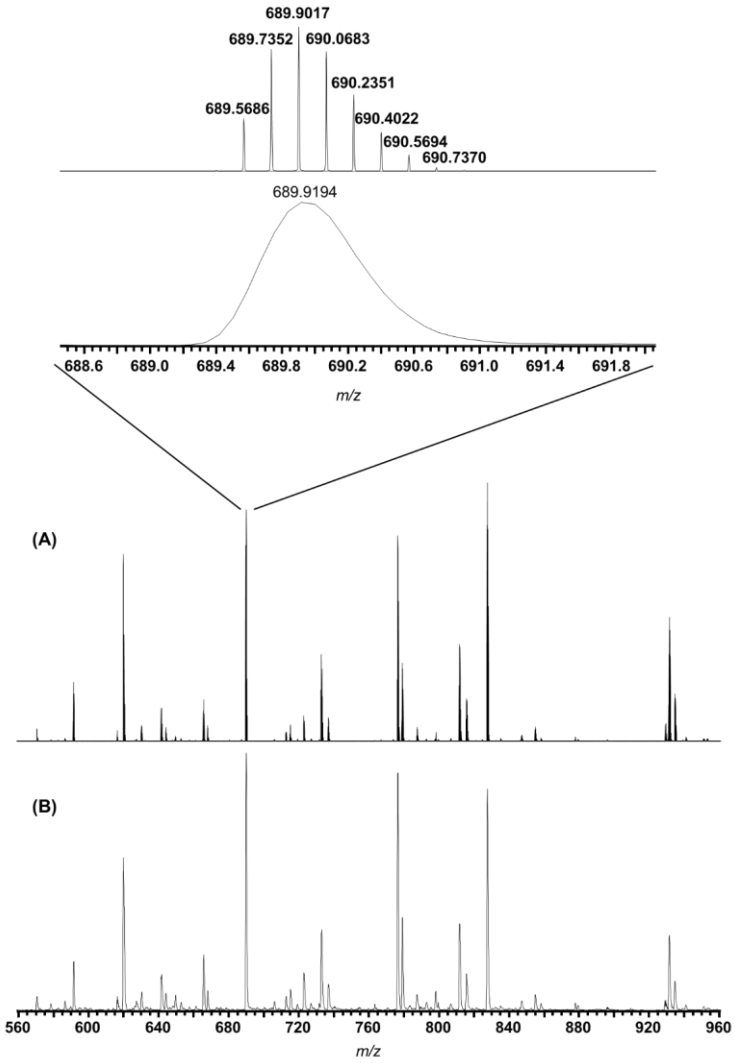


Figure 1.2. Part of the mass spectra of apomyoglobin peptic digests. (A) Mass spectrum collected with an Orbitrap mass analyzer; (B) Mass spectrum collected with just the LTQ. The inserts (expanded regions above spectrum (A)) show the mass spectra collected with the Orbitrap (top) or just the LTQ (bottom insert) for a 6+ charged peptide.

The extraordinary performance of the Orbitrap results from its unique design. These analyzers contain two cup-shaped outer electrodes, and a spindle-like central electrode.[19] Application of a voltage between the outer and central electrodes gives a linear electric field along the axis, which creates harmonic oscillations of ions. The radial electric field drags ions to the central electrode, and creates bent ion trajectories. When a packet of ions accumulated in the C-trap (a curved radiofrequency-only quadrupole ion trap between the LIT and orbitrap[25]) are injected into the Orbitrap, they begin rotating around the central electrode and oscillating between the outer electrodes. The outer electrodes detect the image currents created by the oscillations, to form the time-domain digital signals. The Fourier-transform of the time domain data yields frequency domain signals. As the frequency of the ion oscillation is inversely proportion to the square root of the *m/z* ratio of the ions, finally, a mass spectrum results. One of the limitations of the Orbitrap analyzer is that it cannot conduct MS/MS analysis, so normally it couples with other mass analyzers. Figure 1.3 shows the schematic of the LTQ Orbitrap Velos MS instrument that I used to acquire most of my infusion and LC-MS data.

The Orbitrap family of mass spectrometers has grown rapidly in the past ten years. Normally, a new version of an Orbitrap instrument is introduced to the general community during the American Society of Mass Spectrometry annual conference. In the future, I expect to see Orbitrap analyzers with even higher acquisition speeds, resolving powers, mass accuracies and sensitivities.

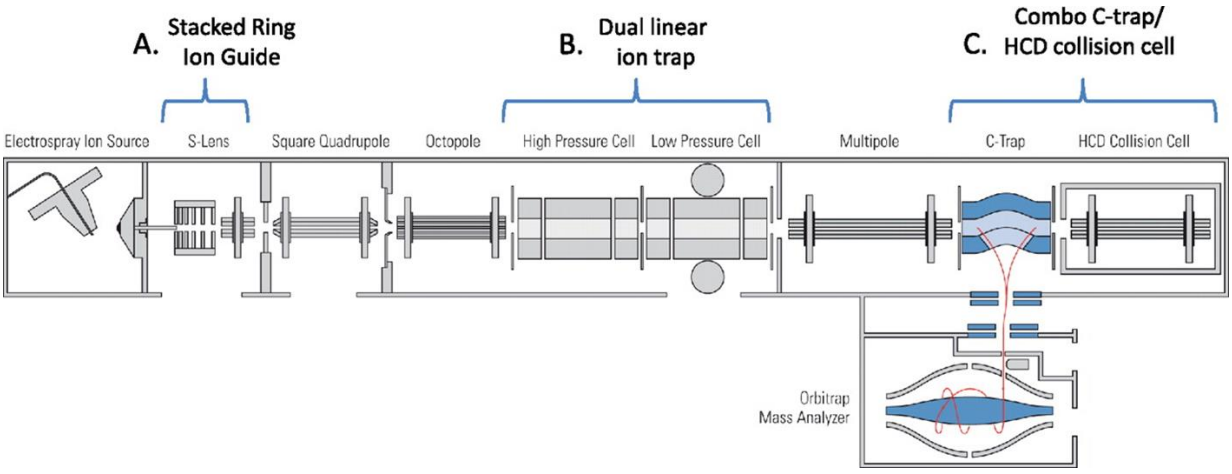


Figure 1.3. Schematic drawing of the LTQ Orbitrap Velos MS instrument with three new features compared with the old model LTQ Orbitrap. "A, the stacked ring ion guide (S-Lens) increases the ion flux from the ESI source into the instrument by a factor 5–10; B, the dual linear ion trap design enables efficient trapping and activation in the high-pressure cell (*left*) and fast scanning and detection in the low pressure cell (*right*). C, the combo C-trap and HCD collision cell with an applied axial field with improved fragment ion extraction and trapping capabilities."[26] Figure taken (with permission) from Jesper V. Olsen, Jae C. Schwartz, Jens Griep-Raming, Michael L. Nielsen, Eugen Damoc, Eduard Denisov, Oliver Lange, Philip Remes, Dennis Taylor, Maurizio Splendore, Eloy R. Wouters, Michael Senko, Alexander Makarov, Matthias Mann, and Stevan Horning. A Dual Pressure Linear Ion Trap Orbitrap Instrument with Very High Sequencing Speed. *Mol. Cell Proteomics.* 2009; Vol 8:2759-2769. © the American Society for Biochemistry and Molecular Biology.

1.1.3 Tandem mass spectrometry methods

Proper selection of the ionization method and mass analyzer is vital for obtaining MS results that solve a specific problem. Determination of detailed protein or peptide sequence information

often requires MS/MS analysis of the target protein or peptide ions, and the amount of sequence data one obtains often depends on the fragmentation method. Below I discuss the fragmentation methods that I employ in this dissertation.

1.1.3.1 Collision-induced activation/dissociation

Collision-induced dissociation (CID), also called collision activation dissociation (CAD), is the most common fragmentation method in MS/MS. It usually takes place in the collision cell of a mass spectrometer, the LIT or the Q2 of the triple quadrupole, for example.[10] After applying a supplemental resonance excitation voltage to the x-axis, ions gain energy and collide with inert gas. A fraction of the ion translational energy transfers to internal energy, which brings the ion to an excited state. Subsequent unimolecular decomposition of the activated ion gives the products ions (*b*-, *y*-type ions). CID is an ergodic ion activation method, and redistribution of the internal energy results in fragmentation at the weakest bonds. Ions with a high charge state obtain more kinetic energy in a given electric field than low-charge-state ions, and thus have a higher probability of fragmenting.[27] McLuckey et al. summarized the key experimental parameters in the CID of peptides and proteins ions.[28]

However, MS/MS spectra from a LIT suffer from low resolution and low mass accuracy. Also, fragment ions in the low *m/z* range are lost because of the low-mass cutoff that results from the radio frequency amplitude.[29] Additionally, the fragmentation preference for the weakest bond makes locating labile protein PTMs, such as phosphorylation, challenging. Moreover, CID has limited value for intact protein fragmentation.[19] CID includes an energy redistribution process, and for large ions such as proteins, energy redistribution goes through a large number of bonds. This limits the reaction rate for protein fragmentation.

1.1.3.2 Higher-energy C-trap dissociation

The term higher-energy C-trap dissociation (HCD), coined by Olsen et al., specifically describes fragmentation in the octopole collision cell of the Orbitrap instrument.[30] Different from CID MS/MS spectra of peptides, corresponding HCD spectra contains a_2, b_2 ion pairs and y_1 and y_2 ions, which helps the identification of reporter ions in the low *m/z* region. HCD is more efficient for intact protein fragmentation than CID.[31]

1.1.3.3. Electron Capture Dissociation and Electron Transfer Dissociation

McLafferty and his coworkers developed electron capture dissociation (ECD) as a fragmentation method.[32] A heated filament source outside the FTICR magnet produces low energy electrons (<0.2 eV), and a multiply charged positive ion captures an electron and forms a radical positive ion with reduced charge. An increase in ion internal energy because of dissociative recombination of an electron and the positive ion allows bond fission. ECD is a non-ergodic process, where no vibrational energy redistribution occurs.[33] As such, ECD cleaves more bonds than CID, product ions come from single-bond cleavage, and labile PTMs and non-covalent bonds remain after ECD. I see many possible applications for protein fragmentation using ECD. However, the size of the electron beam is small compared to the volume of the ion trapping chamber, so the fragmentation efficiency is low. Moreover, ECD occurs mainly with the FTICR, not the LIT because the electric field in the LIT expels electrons, which limits their reaction with ions.[19]

To overcome this problem, the Hunt lab developed electron transfer dissociation (ETD). The mechanism of ETD involves electron transfer from a singly charged radical anion to multiply charged cations.[34] This process releases 4-5.5 eV of energy, which triggers release of a hydrogen

radical. The fragmentation pathway is then the same as in ECD. ECD and ETD form *c*-, *z*-type ions. Figure 1.4 shows the production of *b*-, *y*-type ions and *c*-, *z*-type ions.

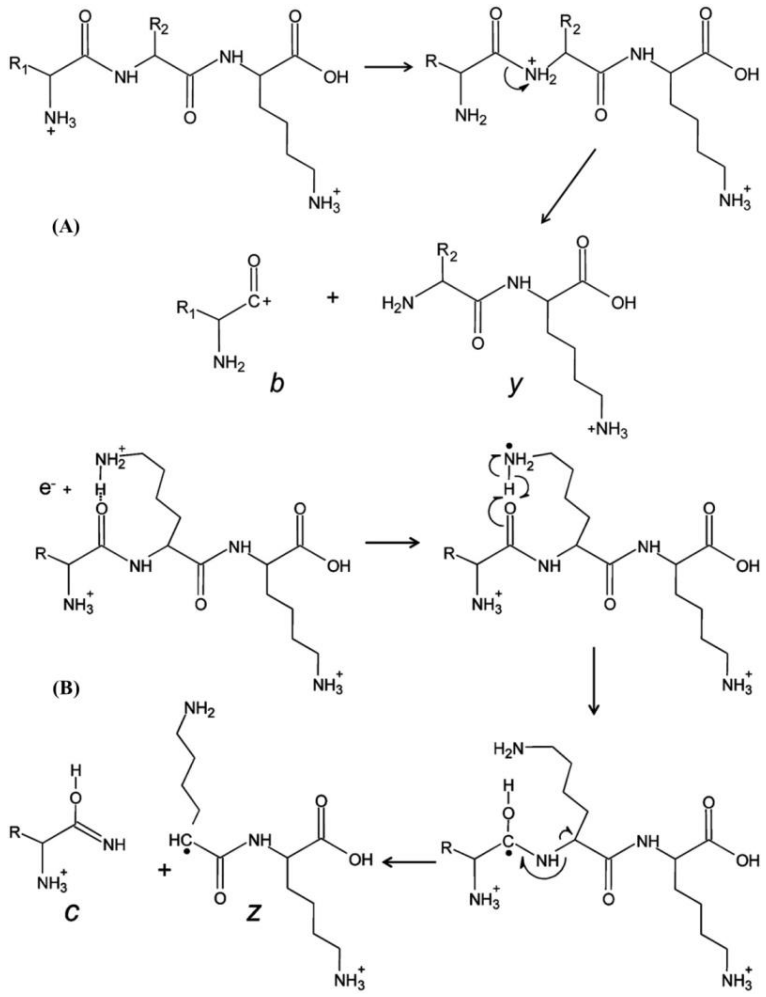


Figure 1.4. Mechanisms of CID and ETD. (A) CID of a multiply protonated peptide for production of *b*- and *y*-type ions. (B) Reaction of a low-energy electron with a multiply protonated peptide produces *c*- and *z*-type ions. The picture is reprinted with permission from: John E. P. Syka, Joshua J. Coon, Melanie J. Schroeder, Jeffrey Shabanowitz, and Donald F. Hunt. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* 2004, 101 (26), 9528-33. Copyright (2004) © National Academy of Sciences.

1.2 Monoclonal antibody analysis

Antibodies are large glycoproteins produced by B lymphocytes in our immune system to identify and eliminate foreign objects such as viruses.[35] In 1890, Behring and Kitasato published an article that showed that serum from an animal actively immunized against diphtheria toxin could neutralize even a fatal dose of the toxin in another animal.[36] Since then, antibodies have been the subject of intense research and applications. In 1975, Kohler and Milstein developed the hybridoma technique,[37] which makes the production of highly pure and specific mAbs possible *in vivo*[38] and *in vitro*.[39] Different from polyclonal antibodies, monoclonal antibodies are monospecific and homogeneous because they come from a single clone of immune cells. The general procedure for *in vivo* production of mAbs begins with immunization of an animal followed by isolation of B cells from the animal's spleen. Cultivated myeloma cells are then fused with the isolated B cells to create hybridoma cells. Subsequently, the hybridomas that produce antibodies of desired specificities are selected and cloned to produce identical daughter clones.[40]

However, adverse human immune reactivity was an initial problem with mAbs derived from hybridoma cells. Most mAbs were initially developed from animals such as rabbits. When given to humans, these "foreign" bodies often evoked an immune reaction and were eliminated before reaching their targets. One way to solve this problem is to produce genetically engineered antibodies, such as chimeric antibodies and humanized antibodies.[41] The three hypervariable loops, the complementarity determining regions (CDRs), on the light chain and heavy chain of an antibody determine the affinity and specificity of the antibody to a specific antigen. Thus, to build a recombinant humanized antibody, genes that encode the variable regions are fused with

genes that encode the remaining parts from a human antibody. This process removes most of the potentially immunogenic portions of the mAB, but its specificity for the intended therapeutic target does not change.

1.2.1 Monoclonal antibody market

At the same time that advances in MS technology were reshaping protein analysis, monoclonal antibodies were becoming the fastest growing class of therapeutic drugs. In 1986, the US Food and Drug Administration (FDA) approved Muromonab-CD3 (Orthoclone OKT3), which became the first therapeutic antibody. Since then, the biotherapeutic market has continually increased, due in large part to the development of therapeutic mAbs, Fc-fusion proteins, antibody fragments, and antibody-drug conjugates (ADCs). More than 40 antibody-related drugs are commercially available, including the "big 5": Rituximab (Rituxan, for Non-Hodgkin's lymphoma treatment), Infliximab (Remicade, for Crohn disease treatment), Trastuzumab (Herceptin, for breast cancer treatment), Adalimumab (Humira, for Rheumatoid arthritis treatment), and Bevacizumab (Avastin, for colorectal cancer treatment).[42] Antibodies have a higher success rate (25-29%) from clinical phase I trials to approval than small molecules (11%).[43] Their high specificity and low side effects are especially attractive. Table 1 shows marketed antibody products. In 2013, global sales for antibody products were estimated as $75 billion.[44]

Table 1.1. Marketed therapeutic monoclonal antibody products.

Brand name (INN)	Original BLA/MAA Applicant	Company Reporting US Sales	Company Reporting EU Sales	Year of First Approval	2013 Global Sales ($M)[a]
Abthrax (raxibacumab)	Human Genome Sciences	GlaxoSmithKline	N/A[b]	2012	23
Actemra (tocilizumab)	Roche	Roche	Roche	2009	1,119
Adcetris[c] (brentuximab vedotin)	Seattle Genetics	Seattle Genetics	Takeda Pharmaceutical Co.	2011	253
AlprolIX[d] (Factor IX Fc fusion protein)	Biogen Idec	Biogen Idec	N/A	2014	NoM[e]
Arcalyst[f] (rilonacept)	Regeneron Pharmaceuticals	Regeneron Pharmaceuticals	N/A	2008	17
Arzerra (ofatumumab)	GlaxoSmithKline	GlaxoSmithKline	GlaxoSmithKline	2009	117
Avastin (bevacizumab)	Genentech	Roche	Roche	2004	6,748
Benlysta (belimumab)	Human Genome Sciences	GlaxoSmithKline	GlaxoSmithKline	2011	228
Cimzia[g] (certolizumab pegol)	UCB	UCB	UCB	2008	789
Cyramza (ramucirumab)	Eli Lilly and Co.	Eli Lilly and Co.	N/A	2014	NoM[e]
Eloctate[h] (Factor VIII Fc fusion protein)	Biogen Idec	Biogen Idec	N/A	2014	NoM[e]
Enbrel[i] (etanercept)	Immunex	Amgen	Pfizer	1998	8,325
Entyvio (vedolizumab)	Takeda Pharmaceuticals U.S.A., Inc	Takeda Pharmaceutical Co.	Takeda Pharmaceutical Co.	2014	NoM[e]
Erbitux (cetuximab)	ImClone Systems	Bristol-Myers Squibb	Merck KGaA	2004	1,926
Eylea[j] (aflibercept)	Regeneron Pharmaceuticals	Regeneron Pharmaceuticals	Bayer Healthcare Pharmaceuticals	2011	1,851
Gazyva (obinutuzumab)	Genentech	Roche	Roche	2013	3
Herceptin (trastuzumab)	Genentech	Roche	Roche	1998	6,559
Humira (adalimumab)	Abbott Laboratories	AbbVie	AbbVie	2002	10,659
Ilaris (canakinumab)	Novartis Pharmaceuticals	Novartis Pharmaceuticals	Novartis Pharmaceuticals	2009	119
Inflectra[k l] (infliximab [biosimilar])	Hospira	N/A	Hospira	2013	<1[m]
Kadcyla[n] (ado-trastuzumab emtansine)	Genentech	Roche	Roche	2013	252
Keytruda (pembrolizumab)	Merck & Co.	Merck & Co.	N/A	2014	NoM[e]
Lemtrada (alemtuzumab)	Genzyme Therapeutics	N/A	Sanofi	2013	3

Table 1.1 (cont'd)

Lucentis[o] (ranibizumab)	Genentech	Roche	Novartis Pharmaceuticals	2006	4,205
Nplate[p] (romiplostim)	Amgen	Amgen	Amgen	2008	427
Nulojix[q] (belatacept)	Bristol-Myers Squibb	Bristol-Myers Squibb	Bristol-Myers Squibb	2011	26
Orencia[r] (abatacept)	Bristol-Myers Squibb	Bristol-Myers Squibb	Bristol-Myers Squibb	2005	1,444
Perjeta(pertuzumab)	Genentech	Roche	Roche	2012	352
Prolia[s] (denosumab)	Amgen	Amgen	GlaxoSmithKline	2011	824
Remicade (infliximab)	Centocor	Johnson & Johnson	Merck & Co.	1998	8,944
Removab[t] (catumaxomab)	Fresenius Biotech	N/A	NeoPharm Group	2009	5
Remsima[k l] (infliximab [biosimilar])	Celltrion	N/A	Celltrion	2013	<1[m]
ReoPro[u] (abciximab)	Centocor	Lilly	N/A	1994	127
Rituxan (rituximab)	Genentech	Roche	Roche	1997	7,500
Simponi/ Simponi Aria (golimumab)	Centocor Ortho Biotech	Johnson & Johnson	Merck & Co.	2009	1,432
Simulect (basiliximab)	Novartis Pharmaceuticals	Novartis Pharmaceuticals	Novartis Pharmaceuticals	1998	30[v]
Soliris (eculizumab)	Alexion Pharmaceuticals	Alexion Pharmaceuticals	Alexion Pharmaceuticals	2007	1,551
Stelara (ustekinumab)	Janssen-Cilag International	Johnson & Johnson	Johnson & Johnson	2009	1,504
Sylvant (siltuximab)	Janssen Biotech	Johnson & Johnson	Johnson & Johnson	2014	NoM[e]
Synagis (palivizumab)	Abbott Laboratories	AstraZeneca	Abbvie	1998	1,887
Tysabri (natalizumab)	Biogen Idec	Biogen Idec	Biogen Idec	2004	1,527
Vectibix (panitumumab)	Amgen	Amgen	Amgen	2006	389
Xgeva[s] (denosumab)	Amgen	Amgen	Amgen	2010	1,030
Xolair (omalizumab)	Genentech	Roche	Novartis	2003	1,465
Yervoy (ipilimumab)	Bristol-Myers Squibb	Bristol-Myers Squibb	Bristol-Myers Squibb	2011	960
Zaltrap[w] (ziv-aflibercept)	Sanofi Aventis	Sanofi	Sanofi	2012	70
Zevalin[x] (ibritumomab tiuxetan)	IDEC Pharmaceuticals	Spectrum Pharmaceuticals	Spectrum Pharmaceuticals	2002	29

[Acronyms: INN, International Nonproprietary Name; BLA, Biologics License Application; MAA, Marketing Authorization Application]. [a] Sales information obtained from company annual reports and other publically available sources. [b] N/A denotes product not available in this region. [c] Antibody-Drug Conjugate. [d] Fc Fusion Protein, Fc-Factor IX. [e] Product approval in 2014; no sales in 2013. [f] Fc Fusion Protein, Fc-IL1R. [g] Fab Conjugate. [h] Fc Fusion Protein, Fc-Factor VIII. [i] Fc Fusion Protein, Fc-TNFR (p75). [j] Fc Fusion Protein, Fc-VEGFR (1,2). [k] Biosimilar Antibody, Remicade Originator. [l] Inflectra and Remsima are considered as two individual products. [m] Product approval in late 2013; no annual sales disclosed, bioTRAK estimate of global sales. [n] Antibody-Drug Conjugate. [o] Fab. [p] Fc Fusion Protein, Fc-TPO-R binding peptide. [q] Fc Fusion Protein, Fc-CTLA-4 with amino acid substitutions. [r] Fc Fusion Protein, Fc-CTLA-4. [s] Prolia and Xgeva are considered as two individual products even though they contain the same bulk monoclonal antibody. [t] Bispecific, Tri-functional Antibody. [u] Sales data not disclosed, small patient market, bioTRAK® estimate of global sales. [v] Fab, produced by papain digestion of full length monoclonal antibody. [w] Fc Fusion Protein, Fc-VEGFR. [x] Antibody Conjugate. Table is reprinted (with permission) from Dawn M Ecker, Susan Dana Jones, Howard L Levine. The therapeutic monoclonal antibody market. *MAbs*, 2015, 7 (1), 9-14. Copyright © 2015 Taylor & Francis.

An antibody (in this dissertation antibody refers to immunoglobulin G, IgG) can potentially kill cancer cells by antibody-dependent cellular cytotoxicity (ADCC) or complement-dependent cytotoxicity (CDC). Figure 1.5 shows the mechanism for the antibody enhancing the effector function. ADCC is triggered by an interaction of the antibody Fc region and the Fcγ receptors (FcγRs) on immune effector cells, such as neutrophils, macrophages and natural killer cells. The tumor cell is killed by phagocytosis or lysis. In CDC, recruitment of the complement component C1q by IgG triggers a proteolytic cascade to activate the complement, which can lead to the formation of a membrane attack complex that kills the target cell by fracturing its cell membrane.[41]

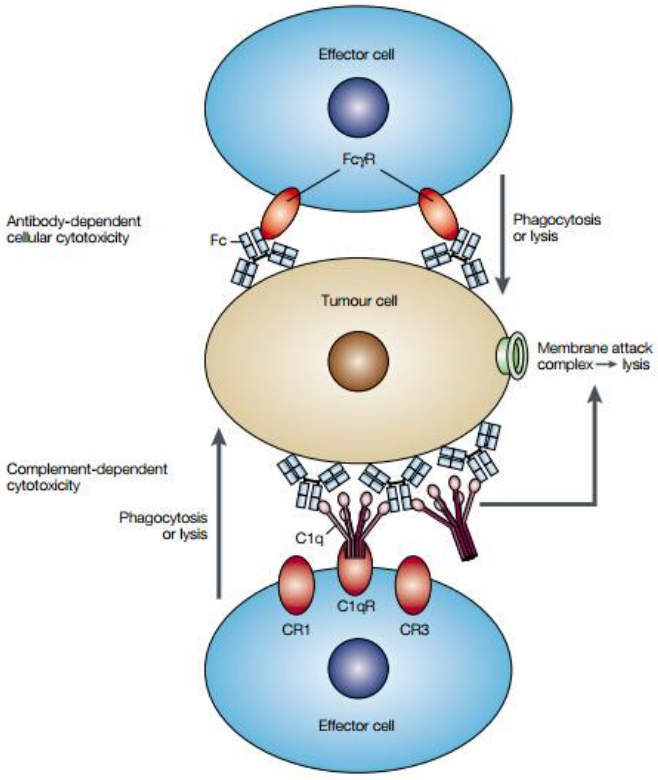


Figure 1.5. Mechanisms of ADCC and CDC. This figure is reprinted with permission from Paul Carter. Improving the efficacy of antibody-based cancer therapies. *Nat. Rev. Cancer*, 1, 118-129. Copyright © 2001, Rights Managed by Nature Publishing Group.

1.2.2 MS in mAb analysis

mAbs, typically immunoglobulin G (IgG), have a molecular weight of approximately 150 kDa. As Figure 1.6 shows, they contain two identical light chains (Lc, ~25 kDa) and two identical heavy chains (Hc, ~50 kDa). The Lc contains a constant region (CL), and a variable region (VL), whereas the Hc has three constant regions (CH1, CH2, and CH3) and one variable region (VH). Inter- and intra-disulfide bonds connect the Lc and Hc.

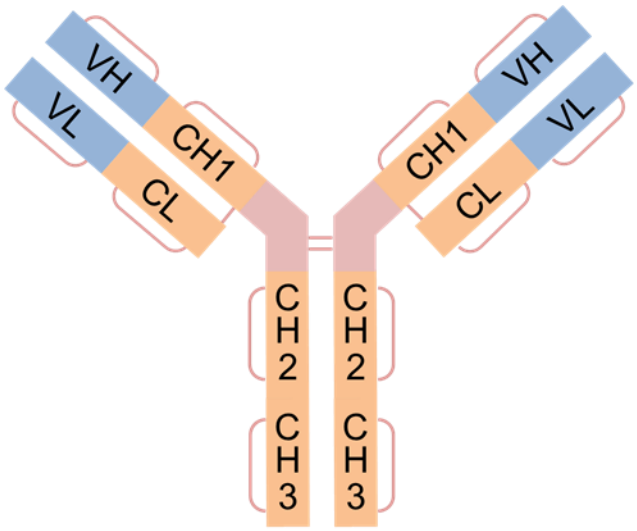


Figure 1.6. Structure of a monoclonal antibody. [Acronyms: VH, variable region of the heavy chain; CH1, CH2, and CH3, different constant regions of the heavy chain; CL, constant region of the light chain; VL, variable region of the light chain; Lc, light chain; Hc, heavy chain.]

Manufacture of these large biomolecules is very different from synthesizing small molecules. Normally, therapeutic antibodies are expressed in mammalian host cell lines, including NS0 murine myeloma cells, PER.C6® human cells, and Chinese hamster ovary (CHO) cells.[45] PTMs such as glycosylation, N-terminal pyroglutamate formation, asparagine deamidation, C-terminal Lysine clipping, aspartic acid isomerization, oxidation, and degradation may result from

intracellular and extracellular processes during the expression, purification and storage. A recent review summarized the heterogeneity of monoclonal antibodies.[46] Production of biosimilar antibodies (the same sequence) is a trend in the biopharmaceutical field; however, the inherent variability of antibody production makes exact copies of an effective antibody drug nearly impossible. The unintended byproducts may compromise product activity and stability. For example, oxidation of Met may cause an antibody conformational change, which affects binding to the antigen or Fc receptors.[47] Deamidation of Asn on the variable region could results in loss of binding affinity.[48] Thus, manufacturers must characterize the mAbs in detail after the production of each batch. Both the US FDA and the European Medicines Agency (EMA) require such control.[49]

The aforementioned heterogeneities of mAbs make their characterization challenging. Among the possible analytical techniques, MS is the most common and plays important roles throughout all the stages of mAb production due to its high accuracy and high-throughput capabilities. MS provides reliable information on clone selection, purification development, stability studies and comparability studies, and gives information related to primary sequences, PTMs, higher order structures and conformations of antibodies.[50] Generally speaking, there are four strategies for antibody analysis: bottom-up, top-down, middle-up and middle-down approaches.[51,52] Figure 1.7 summarizes different MS techniques for mAb characterization.

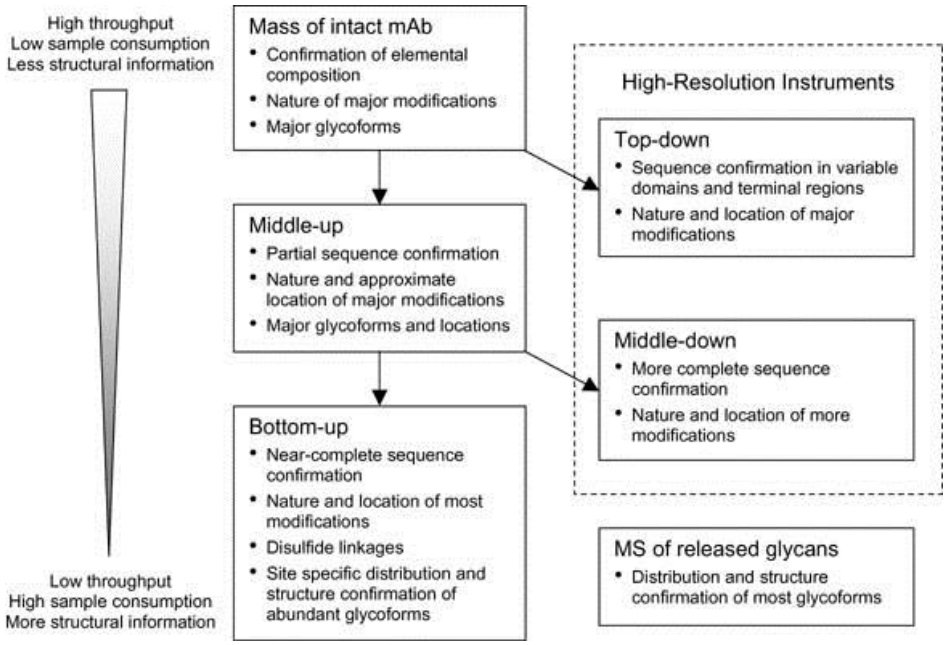


Figure 1.7. A summary of different MS-based techniques for mAb characterization. The figure is reprinted with permission from Zhongqi Zhang, Hai Pan, Xiaoyu Chen. Mass spectrometry for structural characterization of therapeutic antibodies. *Mass Spectrom Rev.*, 2008, 28 (1), 147-176. Copyright © 2008 Wiley Periodicals, Inc.

1.2.2.1 Bottom-up mAb analysis

The bottom-up approach is the most widely used method for antibody characterization. In a typical workflow, after reduction and alkylation, the antibody light chain and heavy chain are digested by one or several enzymes in-gel or in-solution, and the proteolytic peptides are separated in reverse-phase liquid chromatography (RP-HPLC) for ESI-MS analysis. MS/MS analysis of the peptides also provides sequence information including the location of PTMs on the mAb. Although protein digestion can occur in different ways, most of the examples described below use in-solution digestion. I will further discuss protein digestion methods in later sections.

Many studies employed the bottom up procedure to investigate mAb variations and PTMs. Wang et al. used trypsin and Asp-N to digest huN901 antibody and separated the peptides with a C18 column prior to ESI-TOFMS analysis. [53] They identified N-terminal pyroglutamate formation, cleavage of C-terminal lysine, glycosylation, and deamidation on the huN901 antibody. Johnson et al. performed peptide mapping of an antibody after Asp-N digestion followed by RP-HPLC/MS analysis, and revealed C-terminal α-amidation on the heavy chain.[54] Ayoub et al. analyzed trypsin and GluC digests of cetuximab, and LC-MS/MS results suggested Ala213Glu substitution, and Cys214 missing in the light chain.[55] Gahoual et al. conducted capillary electrophoresis-tandem mass spectrometry (CESI-MS/MS) analysis of Trastuzumab after trypsin in-solution digestion, and revealed the glycosylation profile.[56] The Tsybin lab developed an "extended bottom-up" method using a novel enzyme, Sap 9, and achieved high sequence coverage of the light and heavy chains, with decreased introduction of artefacts during digestion.[57] Wang et al. recently published their work on monitoring antibody PTMs by LC-MS with ultrafast tryptic digestion.[58] Du et al. used ^{18}O-water as the solvent for mAb sample preparation followed by trypsin in-solution digestion.[59] With this method they identified the deamidation artifacts introduced by sample preparation.

Taylor Zhang and coworkers identified and characterized unpaired cysteines in a recombinant antibody.[60] To locate the unpaired cysteines, they conducted trypsin digestion, and used N-ethylmaleimide (NEM) to tag free thiols. Different from iodoacetamide, NEM can alkylate thiols at neutral or slightly acidic conditions. LC-MS analysis of tryptic peptides revealed free thiols at Cys-22 and Cys-96 in the variable region of the heavy chain. Xiang et al developed a two-step alkylation method for localization and quantitation of free thiols in mAbs. Briefly, the antibody was alkylated first with ^{12}C-iodoacetic acid, and subsequently was reduced and alkylated a

second time using ^{13}C-iodoacetic acid.[61] Trypsin, Lys-C, chymotrypsin, Asp-N and Glu-C digests were analyzed by LC-MS. Peptides modified by ^{13}C-iodoacetic acid had a molecular weight 2 Da more than peptides modified by ^{12}C-iodoacetic acid, which gives information about the free thiols. Hancock's group used LC-MS with ETD to characterize mAb disulfide linkages.[62] Tryptic in-gel digests (non-reducing gel) of mAbs were fragmented by CID and ETD. ETD preferentially fragments disulfide-linked peptides into two polypeptides, and CID (MS^3) can further fragment these peptides to get more backbone cleavages. Using this method, they successfully identified the disulfide bonds scrambled under heat stress.

Bottom-up methods are also widely used in hydrogen deuterium exchange (HDX) experiments for studying the antibody higher order structure.[63] Pepsin routinely serves as the antibody protease after HDX because it is highly active under acidic conditions. Several papers summarized the protocols for using HDX.[64,65]

The bottom-up approach is the most powerful method for antibody analysis, but it still has limitations. Importantly, the procedure is usually time-consuming and labor-intensive. Artifact PTMs such as deamidation may occur during long incubations, and information about correlations between PTMs is lost because most PTMs reside on different peptides. Also, in LC-MS/MS analysis, peptides larger than 4 kDa are hard to characterize by MS/MS, and peptides with two or three amino acids as well as hydrophilic peptides are easily lost during LC separation because of their poor retention on reverse-phase columns. CE may overcome some of these limitations of RP-LC, and we expect to see more research on antibody characterization using CE-MS.[66]

1.2.2.2 Top-down approach for antibody analysis

Rapid improvements in ultrahigh resolution mass spectrometers have led to remarkable progress in the top-down approach for antibody characterization. This method determines the molecular mass of the intact protein and fragments the intact gas-phase protein ions without digestion.[67] Top-down analysis can detect correlations between multiple PTMs, and sample preparation is simple compared with the bottom-up workflow. Fragmentation of the antibody can occur through CID, HCD, ECD, ETD, and ultraviolet photodissociation (UVPD), and the high-resolution mass spectrometer is vital because the fragment ions are normally large.

Zhang and Shah performed top-down analysis of the mAb variable regions via in-source fragmentation in a LTQ-Orbitrap instrument.[68] In-source fragmentation occurs in the capillary-skimmer region and has the advantage that it can fragment all the charge states of the protein, which increases the sensitivity. They further conducted CID-MS/MS of specific fragment ions to obtain more sequence information. Bondarenko et al. conducted top-down HPLC/MS analysis of an IgG2 using an LTQ-Orbitrap.[69] Using MagTran and ProMass software for ESI mass spectra deconvolution, they achieved a mass accuracy of the intact antibody within ±2 Da (15 ppm). In-source CID of the intact IgG2 molecule showed a fragmentation pattern similar to their previous work. However, CID of an intact antibody yields limited sequence coverage. To increase the sequence coverage, Tsybin and coworkers conducted antibody analysis using a QTOF and ETD, because for large proteins ETD yields more fragmentation than CID.[70] This instrument had a stated resolution up to 50,000 over a wide *m/z* range for intact protein, and up to 30,000-40,000 over a wide *m/z* range for fragment ions. The advantages of TOF versus Orbitrap and FTICR are larger dynamic range and single-ion counting. The TOF/ETD studies gave 21% sequence coverage for Murine MOPC 21 IgG, and 15% sequence coverage for human antiRhesus D IgG.

Tsybin's group further conducted ETD of Humira on an Orbitrap Velos Pro. By averaging time-domain transients from different LC-MS experiments before FT signal processing, they obtained increased sensitivity, and higher sequence coverage (33%).[71] Allan Marshall's group recently used a 9.4 T FTICR to study an intact mAb.[72] Simultaneous ECD of all the antibody charge states (42+ to 58+) yields more fragmentation (~34% sequence coverage) than ECD of one charge state (+51, ~25% sequence coverage). The limited sequence coverage for all these top-down approaches is mainly due to the highly structured and disulfide bond-protected areas.

1.2.2.3 Middle-down approach for antibody analysis

The middle-down approach to protein characterization attempts to combine the strengths of bottom-up and top-down methods through analysis of large protein pieces (~3-20 kDa) obtained from limited digestion or reduction. Chemical or electrochemical reduction of an antibody forms free light chain and heavy chains. Some researchers use the term middle-up to describe the mass characterization of protein subunits and limited digests, while middle-down refers to MS/MS of the protein subunits or peptides from limited digestion.

Tsybin used a new enzyme, Immunoglobulin G-degrading enzyme of *Streptococcus pyogenes* (IdeS), to selectively cleave the antibody near the hinge region to give large peptides.[73] MS/MS of the large peptides is easier than fragmentation of intact antibody. Also, LC separation of the large peptides yields greater resolution than separation of intact antibodies. Bondarenko et al. analyzed the reduced and alkylated antibody light chain and heavy chain using an LTQ-Orbitrap intrument. LC-MS/MS of the Lc and Hc yielded 53/213 bond cleavages for the Lc and 42/443 bond cleavages on the Hc.[69] In-trap CID gave better sequence coverage than in-source CID

because of a reduction of noise by isolatinga 100 *m/z* range of precursor ions. In-source CID gave lower S/N, because ions from all the *m/z* range were fragmented.

Jennifer Zhang and coworkers conducted limited proteolysis of an antibody using Lys-C with disulfide bond reduction to obtain ~25 kDa-sized Fab Hc, single Fc and Lc peptides.[31] They separated the products of limited Lys-C digestion on a diphenyl HPLC column, and HCD analyses of the subunits showed 18/213 bond cleavages for deglycosylated single Fc, 25/214 for the Lc, and 29/226 of Hc Fab. Most of the fragmentation happened at the N termini of the subunits.

Wang et al. used the middle-down approach to compare two anti-CD20 antibody drug products.[74] IdeS digestion of two antibodies followed by tris (2-carboxyethyl) phosphine hydrochloride (TCEP) reduction yielded the Lc as well as the Fc and single chain Fab (Fd) on the HC. LC-MS analysis of the two digests revealed different PTM profiles on two products. They also suggested that a mass spectrometer with a resolution greater than 30,000 is suitable for the middle-down approach. Fornelli et al. described a similar protocol of using IdeS to first digest the antibody to Fc and $F(ab')_2$, followed by TCEP reduction to form Fc, Fc/2 and Fd.[73] They separated the subunits in a C4 column and fragmented them using ETD in an LTQ Orbitrap Elite instrument. Total-ion-current chromatograms showed clear separation of Fc/2, Lc and Fd. Isolation of the top 5 highly charged precursors on each subunit allowed efficient ETD. This procedure gave 67.6% sequence coverage for Fc/2, 68.5% for Lc and 58.6% for Fd. They also conducted IdeS digestion on an antibody mixture (three equimolar antibodies), and used LC to separate the subunits. Not surprisingly, the three Fc/2 subunits did not separate well because of over 90% sequence homology.

Simone Nicolardi and coworkers developed an online electrochemistry-assisted reduction of disulfide bonds workflow, and conducted fragmentation on the Lc using a 15 T FTICR.[75] The authors tried different reducing conditions, but only achieved partial reduction of the antibody to obtain Lc, Hc, Lc+Hc, and Lc+2Hc pieces. CID of the Lc revealed two intrachain disulfide bonds. Yan et al. further tested the IdeS performance on different IgG subclasses (IgG1, IgG2, and IgG4), as well as an Fc fusion protein.[76] Antibody subunits were separated on a C8 column, and interrogated with a Waters QTOF. They found that the cleavage sites were (PELL)G|G(P) for IgG1, (P.VA)G|G(P) for IgG2, (PEFL)G|G(P) for IgG4, and (PELL)G|G(P) for the Fc fusion protein. They concluded that this middle-down approach with IdeS proteolysis is convenient for IgG domain mapping.

Deyun Wang et al. performed HCD middle-down MS/MS on the Lc and Hc of an antibody variant using a Q-Exactive Orbitrap.[77] They separated the antibody and the variant by cation-exchange chromatography, reduced the two fractions and separated the resulting mAb pieces on a C4 column. HCD fragmentation energy was optimized for the light chain precursor ion, because MS/MS of protein either by CID or ECD is charge-state-dependent. They finally achieved 46% sequence coverage of the Lc, and 20.3% coverage of the Hc. They further fragmented seven impurities, and concluded they contained site-specific modifications.

Recently, Brodbelt's group demonstrated a method for analyzing antibody subunits by 193 nm UVPD.[78] They used IdeS to produce Lc, Fc/2 and Fd and analyzed these pieces using an Orbitrap Elite instrument modified with a 193 nm ArF excimer laser, which allows UVPD in the HCD collision cell. Tuning the pulse number and energy per pulse affords control of the fragmentation with the goal of maximizing sequence coverage of the three subunits. After combining data from

four independent UVPD experiments with different parameters, they obtained 80% sequence coverage for the Fc/2 and Lc, with lower than 70% coverage for the Fd peptide.

Although the bottom-up method is still very common for mAb analysis, top-down and middle-down approaches for antibody analysis have grown significantly in the past 5 years. With further improvements in MS technology, I expect to see more application of top-down and middle-down antibody analysis.

1.3 Protein-digestion methods

Protein digestion does not occur solely in the laboratory, it happens in our body every second. Protein digestion begins with pepsinolysis in the stomach. The peptic peptides move to the duodenum for further digestion into amino acids using pancreatic enzymes such as trypsin, chymotrypsin and carboxypeptidase. The body then combines these amino acids to create new specialized proteins. Similarly, in the lab protein digestion helps us understand protein sequence and structure because smaller peptides are easier to characterize with modern analytical techniques.

Catalysis of protein digestion can use enzymes or small molecules. Table 1.2 presents the commonly used proteases and chemicals for protein digestion. However, this section emphasizes enzymatic digestion.

Table 1.2. Common proteases and chemicals for catalysis of protein digestion.[a]

protease	organism	specificity	pH range	chemical	specificity	pH range
Arg-C	Clostridium histolyticum	R’	7.2–8.0[b]	CNBr	M’	acidic
Asp-N	Pseudomonas fragi	‘D	7.0–8.0[b]	HOAc	‘D’[79]	acidic
Glu-C	Staphylococcus aureus	E’[b]	4.0–7.8[b]	FA	D’	acidic
Lys-C	Lysobacter enzymogenes	K’	8.5–8.8[b]	HCl	D’[80]	2.0[80]
Lys-N	Lysobacter enzymogenes	‘K[81]	8.0[81]	NTCB	‘C[80]	9 −10[82]
Trypsin	Bos taurus	K,R’	8.0[b]	Hydroxylamine	N–G	9.0[83]
Chymotrypsin	Bos taurus	F,W,Y’	7.0–9.0[b]			
Pepsin	Sus scrofa	‘F,L,W,Y’	1.3			
		‘F,L’	2			
Thermolysin	Bacillus thermoproteolyticus	‘A,F,I,L,M,V	8.0[c]			
Papain	Carica papaya	R,K,D,H,G,Y[b]	6.0–7.0[b]			
Pronase	Streptomyces griseus	A,E,F,I,L,T,V,W,Y’	6.0–7.5[b]			

‘ refers to cleavage at the C-terminus of the amino acid, and ’ refers to cleavage at the N-terminus of the amino acid. [Acronyms: formic acid (FA), hydrochloric acid (HCl), acetic acid (HOAc), cyanogen bromide (CNBr), 2-nitro-5-thiocyanobenzoate (NTCB)] [a] All data obtained from the Expasy bioinformatics resource portal (www.expasy.org), except those noted. [b] Roche Web site (www.roche-applied-science.com). [c] Sigma-Aldrich Web site (www.sigma-aldrich.com). Reprinted with permission from Linda Switzar, Martin Giera, Wilfried M. A. Niessen. Protein Digestion: An Overview of the Available Techniques and Recent Developments. *J. Proteome Res.*, 2013, 12 (3), pp 1067–1077. Copyright (2013) American Chemical Society.

1.3.1 Enzymatic digestion

As mentioned in the previous section, the bottom-up approach, which includes extensive digestion, is the gold standard for protein characterization.[84] Trypsin is the most popular enzyme for bottom-up proteolysis because it is relatively cheap and can specifically cleave peptide bonds at the C-terminal side of Lysine (Lys) and Arginine (Arg), unless follow by Proline. Lys composes around 5.8% of the human proteome, and Arg shows a little lower abundance. The distribution of Lys and Arg in proteins results in tryptic peptides with an average length of 14 amino acids.[85] Importantly, tryptic peptides contain at least two positions for protonation, and thus usually show at least two positive charges in ESI-MS. The small sizes of tryptic peptides allow effective LC separations, high MS ionization efficiency, and extensive CID-MS/MS fragmentation. Trypsin digestion normally occurs in-solution or in-gel. For the in-solution procedure, trypsinolysis typically follows reduction to break the disulfide bonds and alkylation to protect the thiol group from reforming disulfide bonds. After proper desalting and buffer exchange, in-solution digestion occurs at 37 ℃ for overnight. This protocol is ubiquitous.

Trypsin is not always the ideal proteolytic enzyme. For instance, scientists conduct HDX experiments to study the protein higher-order structure. In this case, protein digestion has to take place in acidic conditions to prevent deuterium back exchange. Because trypsin is only active at neutral pH, it is not suitable for HDX studies. Pepsin is the most common enzyme for protein digestion in HDX because it is inexpensive and active in acidic conditions. Different from trypsin, however, pepsin is less specific and prefers to cleave proteins after hydrophobic amino acids.[86]

In addition to trypsin and pepsin, many other enzymes catalyze proteolysis. Although these enzymes are much more expensive than trypsin, their unique properties make them attractive for specific applications. Endoproteinase Lys-C, for example, is normally used in conjunction with trypsin to give complete protein digestion after Lys residues. In digestion of a yeast protein extract by trypsin alone, over 20% of the cleavage sites remain undigested, and the ratio of missed Lys to missed Arg is 5:1-6:1. Digestion using a Trypsin/Lys-C mixture gives a much lower percent of missed Lys. Lys-C can also catalyze protein digestion by itself, and the digests normally contain middle-down sized peptides. Similarly, Arg-C, Asp-N, and Glu-C can serve in middle-down approaches to protein digestion. Cong Wu et al. used outer membrane protease T (OmpT) for middle-down proteomics because this enzyme cleaves the bonds between two consecutive basic amino acids (Lys/Arg-Lys/Arg).[87] The authors digested the 20-100 kDa proteins collected from the HeLa cell proteome and identified 3,697 unique peptides with an average peptide size of 6.3 kDa. OmpT peptides were also suitable for CID and ETD analysis.

Huesgen and coworkers recently developed another new enzyme, LysargiNase, for protein digestion.[88] Interestingly, LysargiNase mirrors trypsin in specificity, and has specific cleavages at the N-termini of basic residues. CID of LysargiNase results in a series of *b*-ions. When combined with the *y*-ion series from CID of trypsin digests, the peptide identification rate greatly improves. Lys-N has a similar activity but a lower specificity (only 71 % of the Lys-N digests have an N-terminal Lys). Although trypsin still dominates protein digestion because of its cost and specificity, applications that use other enzymes are increasing.

1.3.2 Approaches for rapid protein digestion

Traditional workflows for bottom-up methods are usually time-consuming and labor-intensive. Figure 1.8 shows normal workflows for trypsin in-solution and in-gel digestion. Such complicated protocols are often bottlenecks for sample analysis. Thus, various techniques were developed to accelerate protein digestion.[89]

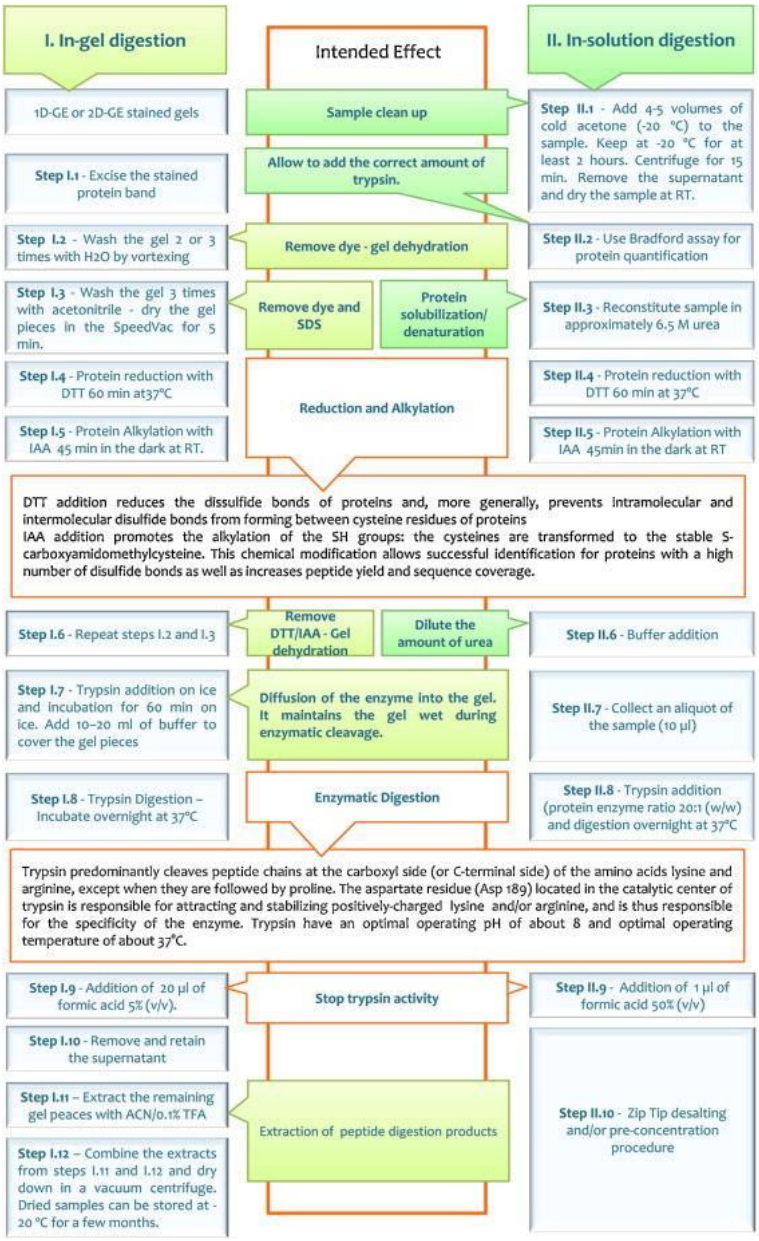


Figure 1.8. Steps and intended effects in classical workflows for protein in-gel and in-solution digestion. The figure is reprinted with permission from J.L. Capelo, R. Carreira,M. Diniz, L. Fernandes,M. Galesio, C. Lodeiro, H.M. Santos, G. Vale. Overview on modern approaches to speed up protein identification workflows relying on enzymatic cleavage and mass spectrometry-based techniques. *Anal. Chim. Acta*, 2009, 650 (2), 151-9.. Copyright © 2009 Elsevier B.V. All rights reserved.

High temperatures may decrease protein digestion time due to thermal protein denaturation.[90] Partial unfolding of the protein increases its accessibility to the proteolytic enzyme. Although enzymes have an optimal working temperature, certain modifications can increase their thermal stability. For example, reductive methylation of trypsin increases its optimized working temperature to 50-60 ℃.[91] Microwave irradiation, which can take place in a simple microwave oven, can accelerate proteolysis and decrease digestion times to several minutes.[92] Figure 1.9 shows a workflow of microwave-assisted protein.

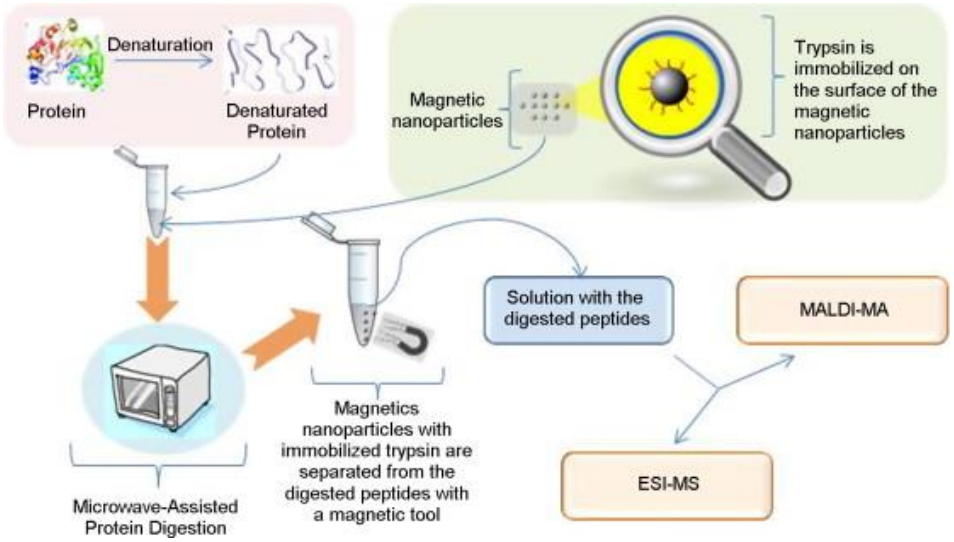


Figure 1.9. A workflow of microwave-assisted protein digestion using immobilized trypsin on magnetic nanoparticles. The figure is reprinted with permission from J.L. Capelo, R. Carreira,M. Diniz, L. Fernandes,M. Galesio, C. Lodeiro, H.M. Santos, G. Vale. Overview on modern approaches to speed up protein identification workflows relying on enzymatic cleavage and mass spectrometry-based techniques. *Anal. Chim. Acta,* 2009, 650 (2), 151-9. Copyright © 2009 Elsevier B.V. All rights reserved.

Ultrasonic energy also facilitates protein digestion. Different research groups used ultrasonic probes[93], ultrasonic baths[94] and sonoreactors[95] to decrease protein digestion times to minutes.

High pressure also increases the rate of protein digestion due to pressure-induced protein denaturation. Tryptic digestion of bovine serum albumin can occur in 60 s at 35,000 psi.[96] Gross and coworkers created an online, high pressure system for rapid digestion in HDX studies.[97] Infrared energy provides another method for enhancing protein digestion efficiency. Excitation of the vibrations of N-H, C=O and C-N bonds of proteins and enzymes increases digestion rates.[98] Table 1.3 gives a short summary of current techniques for accelerating protein digestion, including techniques that employ immobilized enzymes, which is the subject of the next section.

Table 1.3. Overview of Techniques for Accelerated Digestion.

accelerated technique	digestion time	online	compatibility	specific applications
High temperature	Minutes (~15)	Not done, but possible	Often applied to chemical digestion, not all proteases are thermostable	Wide application area
Microwave	Minutes (≤15)	Possible	Compatible with proteases and chemical cleavage reagents	Membrane proteins (increased solubility), glycoproteins (decreased sterical hindrance)
Ultrasound	Minutes (≤5)	Not feasible	Compatible with proteases and chemical cleavage reagents	Wide application area
High Pressure	Seconds (<60)	Yes	Mostly done with enzymes	HDX experiments (speed of online digestion-MS)
Infrared	Minutes (~5)	Not done	Only advantageous for enzymes due to increased interaction with protein	Wide application area
Solvent	Hours (≤5)	Not done, possible but requires stop-flow strategy due to long digestion time	Chemical digestion is often done in the presence of solvents, but some enzymes also tolerate relatively high percentages of organic solvent	Membrane proteins (increased solubility)
IMER*	Minutes (≤20)	Yes	Compatible with each protease that retains activity when immobilized	Wide application area
Magnetic particle immobilized enzyme	Seconds (~30)	Yes	Compatible with each protease that retains activity when immobilized	Wide application area
On-chip immobilized enzyme	Seconds (5)	Yes	Compatible with each protease that retains activity when immobilized	Wide application area

* IMER: Immobilized enzyme reactors. Reprinted with permission from Linda Switzar, Martin Giera, Wilfried M. A. Niessen. Protein Digestion: An Overview of the Available Techniques and Recent Developments. *J. Proteome Res.*, 2013, 12 (3), pp 1067–1077. Copyright (2013) American Chemical Society.

1.3.3 Immobilized enzyme reactors

Immobilized enzyme reactors (IMERs) are widely used for accelerating protein digestion. Immobilization of enzymes onto appropriate solid supports increases the enzyme stability under high temperature, over a broad pH range and in organic buffers.[99-102] IMERs can serve in offline protein digestion as well as in online systems that reduce sample handling and achieve high-throughput analysis. A high local enzyme to substrate ratio yields rapid protein digestion, and immobilization limits autolysis of the enzyme. Specific IMERs were developed with a range of enzymes including trypsin, pepsin, chymotrypsin, Glu-C and Lys-C.[100,103-105] Digestion in IMERs followed by MS analysis is becoming a common workflow for sample analysis.

1.3.3.1 IMER supports

Enzyme immobilization can occur on a range of solid supports including monoliths, capillaries, magnetic particles, resins, microfluidic chips, and membranes.

1.3.3.1.1 Monoliths

Monoliths are attractive IMER supports because of their low back pressure at high flow rates and fast mass transfer due to convection in small pores. Calleri et al. conducted online digestion of β-lactoglobulin A and B in an epoxy-modified silica Chromolith SpeedRood support, and coupled it to a LC-MS/MS system for rapid protein digestion and identification.[106] Digestion required less than 10 min. This group further studied the influence of the enzyme amount on the analytical performance.[107] Krenkova et al. covalently immobilized L-1-tosylamido-2-phenylethyl chloromethyl ketone-trypsin on a poly (glycidyl methacrylate-*co*-ethylene dimethacrylate) monolith in a fused silica capillary.[108] Thirty-second online digestion with LC-MS analysis gave

80% sequence coverage of cytochrome *c*, which was similar to the result with 3 h of in-solution digestion. Masaru Kato et al. fabricated a trypsin IMER by coating a trypsin-containing gel (prepared by the sol-gel technique) on a porous silica monolith.[109] Moreover, the silica monolith was developed to fit into a 96-well plate. The authors proposed to use this IMER for high-throughput protein analysis.

Hanfa Zou's group published work on coupling a monolithic capillary IMER with μRPLC-MS/MS for shotgun proteomics analysis.[110] Digestion in the microreactor for several minutes was similar to 10 h of in-solution tryptic digestion. A one-minute digestion of 590 ng of yeast protein with IMER gave 1578 unique peptides. Ota and coworkers developed trypsin-containing monolithic silica within pipette tips, which is also known as MonoTip® Trypsin.[111] Digestion occurred by pipetting the reduced and alkylated protein less than 20 times.

Regine Schoenherr et al. presented a proof-of concept CE-pepsin microreactor-CE-MS/MS platform for protein analysis.[112] CE separated intact protein in the first dimension, and after rapid protein digestion, CE peptide separation occurred in the second dimension. However, they obtained only 48% sequence coverage for cytochrome *c* and 22% for myoglobin.

Nicoli et al. developed IMERs through immobilization of trypsin onto three monolith disks (CIM® epoxy disk, CIM® CDI disk, and CIM® EDA disk), and used them for online protein digestion and peptide mass fingerprinting.[113] Results from 5 min of online digestion of five proteins with LC-ESI-MS/MS analysis was comparable with 20 h of in-solution digestion. Yukui Zhang' group developed an organic-inorganic hybrid silica monolith trypsin IMER for protein digestion.[114] The enzymatic activity of immobilized trypsin was about 6600 times greater than the activity of free trypsin in solution. A 150 s digestion of 20 μg of *E. Coli* protein using this

IMER lead to identification of 208 proteins, whereas 24 h in-solution trypsin digestion identified 176 proteins.

Krenkova and coworkers immobilized trypsin and Lys-C on a poly(glycidyl methacrylate-*co*-ethylene dimethacrylate) monolith.[115] They incorporated the IMER into an online system, and digested IgG in 6 min. The performance was similar to in-solution digestion for 24 h at 37 ℃. Sinz's group fabricated a capillary monolithic trypsin reactor and used it for online and offline trypsin digestion.[116] They prepared a poly (glycidyl methacrylate-*co*-acrylamide-*co*-ethylene glycol dimethycrylate) monolith in a fused-silica capillary, and immobilized enzyme on the monolith through glutaraldehyde chemistry. They showed a 420-fold higher enzymatic activity of the immobilized trypsin compared to trypsin in solution. Sun et al. integrated a trypsin-IMER with CE-ESI-MS system for online protein digestion.[66] The authors prepared an acrylamide-based monolith in a fused capillary, and tested the IMER performance by digestion of a seven-protein mixture and a picogram quantity of RAW 264.7 cell lysate protein. Guihua Ruan and coworkers immobilized trypsin on a polymerized high internal phase emulsion monolith.[117] Digestion of bovine serum albumin and cytochrome *c* in 10 min gave sequence coverages of 59% and 78%, respectively. The Dovichi group recently published their work on using a sulfonate-silica hybrid strong cation exchange monolith microreactor coupled to a polyacrylamide-coated capillary for online reduction, alkylation and digestion.[118] They identified 3749 peptides after IMER digestion of 50 ng of *Xenopus laevis* zygote homogenate.

1.3.3.1.2 Capillaries

Capillaries provide another attractive solid support for enzyme immobilization. Long and Wood immobilized pepsin onto a fused-silica capillary, and pulled it into a nanoESI emitter.[103] They

coated the modified capillary and the nanoESI emitter with polyaniline to provide conductivity. Digestion of myoglobin gave comparable sequence coverage with in-solution pepsin digestion. Slysz et al. packed trypsin-modified beads in fused-silica capillaries and incorporated the reactor into an online digestion system.[119] After protein separation in a capillary packed with C4 beads, the effluent was neutralized before passing through the IMER, and the resulting protein digest was acidified before injection into the mass spectrometer. Figure 1.10 presents a schematic workflow. The authors tested this integrated separation-digestion-MS system with a mixture of cytochrome *c*, myoglobin, carbonic anhydrase, and ovalbumin, and achieved sequence coverages of 74%, 100%, 53%, and 23%, respectively. The same group further developed the technology by conducting reversed-phase protein chromatography and rapid on-line tryptic digestion. Only 20 fmol of protein was needed for peptide mass fingerprinting.[120]

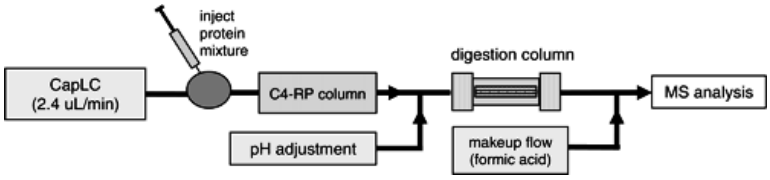


Figure 1.10. Scheme of the apparatus for real-time, on-line digestion of a protein separation. CapLc refers to a capillary LC system (Agilent 1000 Series, Waldbronn, Germany). Reprinted (with permission) from Gordon W. Slysz, David C. Schriemer. Blending Protein Separation and Peptide Analysis through Real-Time Proteolytic Digestion. *Anal. Chem.*, 2005, 77 (6), pp 1572–1579). Copyright (2005) American Chemical Society.

Yamaguchi et al. prepared trypsin and chymotrypsin-IMERs using PTFE microtubes.[121] The immobilization occurred through reaction of proteases with cross-linking agents (paraformaldehyde and glutaraldehyde). Protein digestion required 5 min without reduction and

alkylation. The sequence coverages of cytochrome *c* and BSA after digestion with the trypsin-IMER were 47% and 12%, respectively.

1.3.3.1.3 Magnetic beads

Magnetic particles are also popular substrates for enzyme immobilization. Yan Li and coworkers developed a trypsin-IMER with magnetic microspheres synthesized using a solvothermal reaction followed by coating with tetraethyl orthosilicate.[122] Reaction of microspheres with aminopropyltriethoxysilane introduced amino groups, and activation with glutaraldehyde enabled covalent linking of trypsin to the magnetic silica microspheres through reaction with trypsin primary amines. After packing the magnetic microspheres into microchannels of a chip, digestion of cytochrome *c* for 5 min gave a sequence coverage comparable to that obtained after 12 h of in-solution digestion.

Jeng et al. used trypsin-containing magnetic nanoparticles for rapid in situ protein digestion. Fe_3O_4 nanoparticles were prepared with $–NH_3^+$groups, and then trypsin was immobilized on the nanoparticles using 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) chemistry.[123] Protein digestion took place during incubation of the nanoparticles with the protein, and magnetic separation of the nanoparticles left protein digests in the solution. Digestion of lysozyme at 57 ℃ yielded 19 tryptic peptides with 98% sequence coverage. Sun and coworkers also used magnetic microspheres to prepare a trypsin-IMER. After activation of the he carboxylic acid-functionalized magnetic microspheres with N-hydroxysuccinimide (NHS) and EDC, trypsin amine groups reacted with succinimide groups on the magnetic microspheres to immobilize the enzyme. The authors tested the IMER with digestion of *E. coli* and the MCF7 cell line.[66] Two to thirty min of trypsin-IMER digestion of *E. coli* enabled identification of 1300

proteins, whereas in-solution trypsin digestion lead to identification of 1400 proteins. For MCF7 cell line digestion, in-solution and IMER digestion also give similar numbers of protein identifications.

1.3.3.1.4 Resins

Moore et al. developed a silica based IMER for online protein digestion.[124] They immobilized trypsin onto ethylene bridged hybrid silica particles and packed the modified particles into a chromatographic column. This IMER can tolerate organic solvent, and was placed after the protein separation column to compare the performance of online/offline IMER digestion with in-solution digestion. Protein/peptide elution was monitored by UV, and fractions were collected, recombined, and analyzed by a LC-MS system. With a 10-s volumetric residence time, online digestion of a yeast cell lysate identified 507 proteins, while in-solution digestion identified 490 proteins. Figure 1.11shows the detailed workflow.

Freije and coworkers immobilized acetylated trypsin on Sepharose, Agarose (Pierce beads) and Poroszyme beads.[125] They slurry-packed the beads with buffer into cartridges. Primary amino groups of Lysine and the N-terminus of trypsin were modified by acetic acid N-hydroxysuccinimide-ester to stabilize trypsin and enhance the cytochrome *c* digestion rate. Complete digestion occurred in a contact time of 4 s.

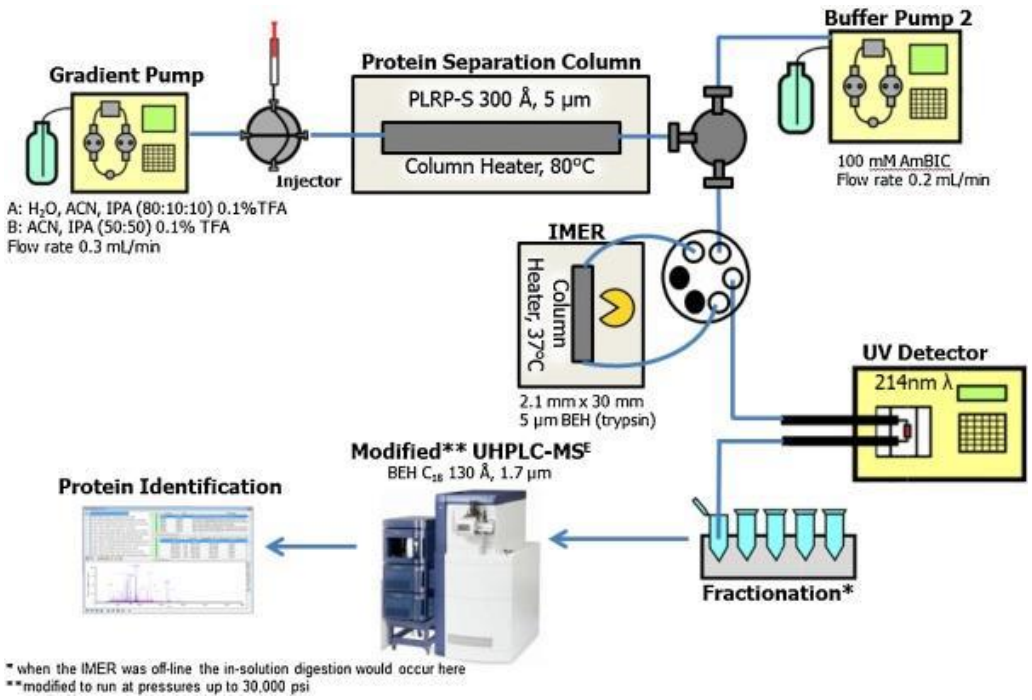


Figure 1.11. Experimental workflow for identifying proteins using an IMER. The figure is reprinted with permission from Stephanie Moore, Stephanie Hess, James Jorgenson. Characterization of an immobilized enzyme reactor for on-line protein digestion. *J. Chromatogr. A*, 2016, 1476, 1–8. . Copyright © 2016 Elsevier B.V. All rights reserved.

1.3.3.1.5 Microfluidic chips

Yun Liu and coworkers developed an IMER using layer-by-layer deposition of positively charged chitosan (CS), and negatively charged hyaluronic acid (HA) onto the surface of a poly (ethylene terephthalate) microfluidic chip.[126] Finally, they soaked a chip modified with nine CS/HA bilayers in a solution containing trypsin. The value of V_{max} per unit of trypsin (Michaelis-Menton kinetics) was ~600 mM/min µg, which is thousands of times faster than that in solution (0.2 mM/min µg). Liuni et al. made a microfluidic reactor by loading pepsin-agarose into a polymethyl methacrylate chip.[127] After digestion with a residence time <4 s, MS sequence

coverages of myoglobin, bovine ubiquitin, and reduced bovine serum albumin were 99%, 64% and 66%, respectively.

1.3.3.2 In-membrane protein digestion

Although many materials can support enzyme immobilization for protein digestion, microporous membranes present a unique platform because of their minimal thickness, large surface area-to volume ratio, low pressure drop, and convective flow through pores. Cooper et al. first used a membrane-based IMER for protein digestion. They immobilized trypsin in a poly(vinylidene difluoride) membrane using simple hydrophobic interactions and constructed a capillary fitting-based trypsin membrane reactor.[128] Flow rates of 10 nL/min gave digests that showed complete sequence coverage of cytochrome *c*, and 20 nL/min flow rates through the capillary gave full sequence coverage of ovalbumin.

The Bruening group developed several enzyme-containing membrane reactors for protein digestion. Xu and coworkers fabricated a trypsin-containing membrane by sequential deposition of polystyrene sulfonate (PSS) and trypsin in a nylon membrane.[129] Figure 1.12 presents the membrane fabrication setup.

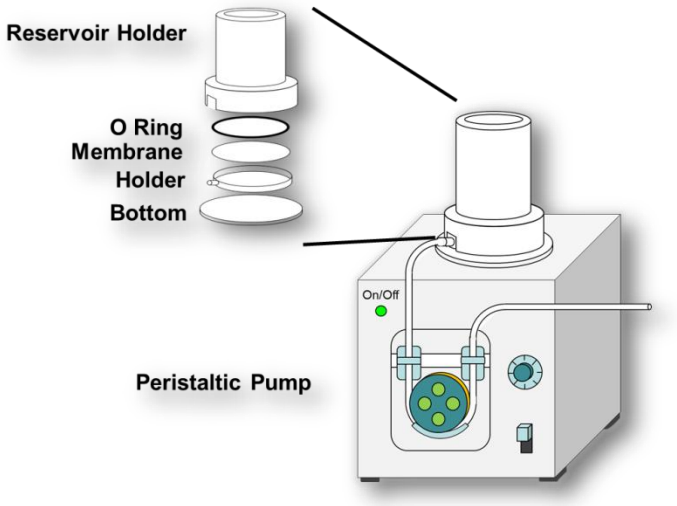


Figure 1.12. Setup for membrane reactor fabrication. (This figure was obtained from Zhefei Yang.)

During flow of a solution through the membrane, PSS adsorbs to the membrane, presumably through hydrophobic interactions with the alkyl backbone of nylon, to give a negatively charged surface, even under acidic conditions. With a pI of ~10.5, trypsin has a positive charge in acidic solution, and can electrostatically adsorb to the negatively charged PSS on the surface. Layer-by-layer deposition on inexpensive nylon membranes gives a reactor with a local concentration of 10 mg trypsin per milliliter of membrane pores, which is 450 times higher than the trypsin concentration for in-solution digestion. The short μm radial diffusion distances in the membrane pores further contribute to facile digestion. Protein solutions were pushed through the membrane reactor with a syringe pump. Figure 1.13 shows a schematic digestion workflow. With such membranes, trypsin in-membrane digestion of Bovine Serum Albumin leads to higher sequence coverage (84%) than in-solution digestion (71%).

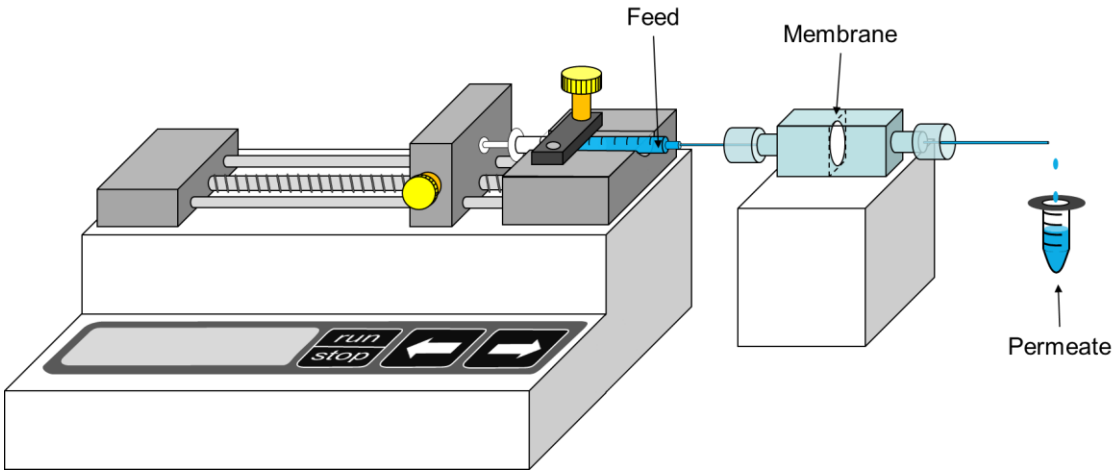


Figure 1.13. Schematic workflow for membrane-based protein digestion. (This figure was obtained from Zhefei Yang.)

Tan et al. used a similar strategy to develop a pepsin-containing membrane. Moreover, the low thickness of the membrane (100 μm) affords control over membrane residence time at the millisecond level to limit digestion.[130] Based on the work of previous group members, I developed a novel workflow for monoclonal antibody characterization using pepsin-containing membranes. By varying the residence times (from 3 ms to 3 s) of antibody solutions in the membrane, we obtained "bottom-up" (1-2 kDa) to "middle-down" (5-15 kDa) sized peptides, and these peptides cover the entire sequences of Herceptin and a Waters antibody.[131] In a recent paper, Ning described using membranes attached to pipet tips for protein digestion. Proteolysis within pipet tips was more complete than digestion for 30 min in solution. Antibody digestion at the end of a pipet tip leads to 100% peptide coverage in MS analyses.[132] From a high-throughput perspective, I further developed an enzyme-containing spin column for protein digestion. Digestion takes places during simple centrifugation of the protein solution through the spin membrane in 1 min or less, prior to analysis using direct infusion into an ultra-high resolution mass spectrometer or LC-MS/MS analysis. One-min peptic or tryptic spin-membrane digestion

yields nearly 100% sequence coverages of apomyoglobin and four commercialized monoclonal antibodies (Herceptin, Avastin, Rituxan and Vectibix).

1.4 Outline of the dissertation

This dissertation contains four subsequent chapters. Chapter 2 describes antibody characterization using pepsin-containing membranes and MS analysis. Controlled in-membrane digestion generates bottom-up to middle-down sized peptides, and direct infusion MS and MS/MS analysis gives nearly 100% sequence coverage and identifies antibody PTMs. Chapter 3 presents a novel protein-digestion device, the enzyme-containing spin membrane. By inserting functionalized membranes into commercial spin columns, this new device provides protein digestion in less than 30 sec. Subsequently, chapter 4 describes the application of in-membrane digestion to determining the sequence differences between two antibodies. Finally, chapter 5 summarizes the aforementioned work and provides future research directions, such as, protease-containing membrane for protein higher order structure analysis, digestion of polyclonal antibodies, *de novo* antibody sequencing and fabrication of glycosidase-containing membranes. Specifically, the next four chapter titles are:

Chapter 2: Pepsin-containing membranes for monoclonal antibody analysis

Chapter 3: Enzyme-containing spin membranes for rapid protein digestion

Chapter 4: Membrane-base proteolytic digestion for antibody sequence comparisons

Chapter 5: Summary and future work

REFERENCES

REFERENCES

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

(2) Boughton, B. A.; Thinagaran, D.; Sarabia, D.; Bacic, A.; Roessner, U. Mass spectrometry imaging for plant biology: a review. *Phytochem. Rev.* **2016,** *15*, 445.

(3) Baghdady, Y. Z.; Schug, K. A. Review of in situ derivatization techniques for enhanced bioanalysis using liquid chromatography with mass spectrometry. *J. Sep. Sci.* **2016,** *39* (1), 102.

(4) Senyuva, H. Z.; Gokmen, V.; Sarikaya, E. A. Future perspectives in Orbitrap-high-resolution mass spectrometry in food analysis: a review. *Food Addit. Contam. Part A Chem. Anal. Control Expo. Risk Assess.* **2015,** *32* (10), 1568.

(5) Domon, B.; Aebersold, R. Mass spectrometry and protein analysis. *Science* **2006,** *312* (5771), 212.

(6) Larsen, M. R.; Trelle, M. B.; Thingholm, T. E.; Jensen, O. N. Analysis of posttranslational modifications of proteins by tandem mass spectrometry. *Biotechniques* **2006,** *40* (6), 790.

(7) Yates, J. R.; Ruse, C. I.; Nakorchevsky, A. Proteomics by mass spectrometry: approaches, advances, and applications. *Annu. Rev. Biomed. Eng.* **2009,** *11*, 49.

(8) El-Aneed, A.; Cohen, A.; Banoub, J. Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. *Appl. Spectrosc. Rev.* **2009,** *44* (3), 210.

(9) Karas, M.; Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* **1988,** *60* (20), 2299.

(10) Zhang, G.; Annan, R. S.; Carr, S. A.; Neubert, T. A. Overview of peptide and protein analysis by mass spectrometry. *Curr. Protoc. Protein Sci.* **2010,** *Chapter 16*, Unit16 1.

(11) Zenobi, R.; Knochenmuss, R. Ion formation in MALDI mass spectrometry. *Mass Spectrom. Rev.* **1998,** *17* (5), 337.

(12) Karas, M.; Gluckmann, M.; Schafer, J. Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors. *J. Mass Spectrom.* **2000,** *35* (1), 1.

(13) Signor, L.; Boeri Erba, E. Matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometric analysis of intact proteins larger than 100 kDa. *J. Vis. Exp.* **2013**, *79*, 50635.

(14) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989,** *246* (4926), 64.

(15) Meng, C. K.; Mann, M.; Fenn, J. B. Of Protons or Proteins - a Beams a Beam for a That (Burns,O.S.). *Z. Phys. D Atom Mol. Cl.* **1988,** *10* (2-3), 361.

(16) Kebarle, P. A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. *J. Mass Spectrom.* **2000,** *35* (7), 804.

(17) Wilm, M. Principles of electrospray ionization. *Mol. Cell. Proteomics.* **2011**, DOI:10.1074/mcp.R111.009407 10.1074/mcp.R111.009407.

(18) Wilm, M.; Mann, M. Analytical properties of the nanoelectrospray ion source. *Anal. Chem.* **1996,** *68* (1), 1.

(19) Hoffmann, E. d.; Stroobant, V. *Mass spectrometry : principles and applications*; 3rd ed ed.; J. Wiley: Chichester, West Sussex, England ; Hoboken, NJ, 2007.

(20) Londry, F. A.; Hager, J. W. Mass selective axial ion ejection from a linear quadrupole ion trap. *J. Am. Soc. Mass. Spectrom.* **2003,** *14* (10), 1130.

(21) Schwartz, J. C.; Senko, M. W.; Syka, J. E. A two-dimensional quadrupole ion trap mass spectrometer. *J. Am. Soc. Mass. Spectrom.* **2002,** *13* (6), 659.

(22) Griffin, T. J.; Xie, H.; Bandhakavi, S.; Popko, J.; Mohan, A.; Carlis, J. V.; Higgins, L. iTRAQ reagent-based quantitative proteomic analysis on a linear ion trap mass spectrometer. *J. Proteome. Res.* **2007,** *6* (11), 4200.

(23) Makarov, A. Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal. Chem.* **2000,** *72* (6), 1156.

(24) Olsen, J. V.; de Godoy, L. M.; Li, G.; Macek, B.; Mortensen, P.; Pesch, R.; Makarov, A.; Lange, O.; Horning, S.; Mann, M. Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* **2005,** *4* (12), 2010.

(25) Perry, R. H.; Cooks, R. G.; Noll, R. J. Orbitrap mass spectrometry: instrumentation, ion motion and applications. *Mass Spectrom. Rev.* **2008,** *27* (6), 661.

(26) Olsen, J. V.; Schwartz, J. C.; Griep-Raming, J.; Nielsen, M. L.; Damoc, E.; Denisov, E.; Lange, O.; Remes, P.; Taylor, D.; Splendore, M.et al. A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed. *Mol. Cel.l Proteomics* **2009,** *8* (12), 2759.

(27) Watson, J. T.; Sparkman, O. D. *Introduction to mass spectrometry : instrumentation, applications and strategies for data interpretation*; 4th ed ed.; John Wiley & Sons: Chichester, England ; Hoboken, NJ, 2007.

(28) Wells, J. M.; McLuckey, S. A. Collision-induced dissociation (CID) of peptides and proteins. *Methods Enzymol.* **2005,** *402*, 148.

(29) Yang, Y. H.; Lee, K.; Jang, K. S.; Kim, Y. G.; Park, S. H.; Lee, C. S.; Kim, B. G. Low mass cutoff evasion with q(z) value optimization in ion trap. *Anal. Biochem.* **2009,** *387* (1), 133.

(30) Olsen, J. V.; Macek, B.; Lange, O.; Makarov, A.; Horning, S.; Mann, M. Higher-energy C-trap dissociation for peptide modification analysis. *Nat. Methods* **2007,** *4* (9), 709.

(31) Zhang, J.; Liu, H.; Katta, V. Structural characterization of intact antibodies by high-resolution LTQ Orbitrap mass spectrometry. *J. Mass Spectrom.* **2010,** *45* (1), 112.

(32) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. Electron capture dissociation of multiply charged protein cations. A nonergodic process. *J. Am. Chem. Soc.* **1998,** *120* (13), 3265.

(33) Zubarev, R. A.; Haselmann, K. F.; Budnik, B.; Kjeldsen, F.; Jensen, F. Towards an understanding of the mechanism of electron-capture dissociation: a historical perspective and modern ideas. *Eur. J. Mass Spectrom.* **2002,** *8* (5), 337.

(34) Syka, J. E.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** *101* (26), 9528.

(35) Chames, P.; Van Regenmortel, M.; Weiss, E.; Baty, D. Therapeutic antibodies: successes, limitations and hopes for the future. *Br. J. Pharmacol.* **2009,** *157* (2), 220.

(36) Llewelyn, M. B.; Hawkins, R. E.; Russell, S. J. Discovery of antibodies. *BMJ* **1992,** *305* (6864), 1269.

(37) Kohler, G.; Milstein, C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **1975,** *256* (5517), 495.

(38) Insel, R. A. In vivo production of human hybridoma antibody to the Haemophilus influenzae B capsule in athymic nude mice. *J. Infect. Dis.* **1984,** *150* (6), 959.

(39) Pardue, R. L.; Brady, R. C.; Perry, G. W.; Dedman, J. R. Production of monoclonal antibodies against calmodulin by in vitro immunization of spleen cells. *J. Cell Biol.* **1983,** *96* (4), 1149.

(40) Breedveld, F. C. Therapeutic monoclonal antibodies. *Lancet* **2000,** *355* (9205), 735.

(41) Carter, P. Improving the efficacy of antibody-based cancer therapies. *Nat. Rev. Cancer* **2001,** *1* (2), 118.

(42) Leavy, O. Therapeutic antibodies: past, present and future. *Nat. Rev. Immunol.* **2010,** *10* (5), 297.

(43) Reichert, J. M. Monoclonal Antibodies as Innovative Therapeutics. *Curr. Pharm. Biotechno.* **2008,** *9* (6), 423.

(44) Ecker, D. M.; Jones, S. D.; Levine, H. L. The therapeutic monoclonal antibody market. *MAbs* **2015,** *7* (1), 9.

(45) Li, F.; Vijayasankaran, N.; Shen, A. Y.; Kiss, R.; Amanullah, A. Cell culture processes for monoclonal antibody production. *MAbs* **2010,** *2* (5), 466.

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

(47) Zhang, A.; Hu, P.; MacGregor, P.; Xue, Y.; Fan, H.; Suchecki, P.; Olszewski, L.; Liu, A. Understanding the conformational impact of chemical modifications on monoclonal antibodies with diverse sequence variation using hydrogen/deuterium exchange mass spectrometry and structural modeling. *Anal. Chem.* **2014,** *86* (7), 3468.

(48) Du, Y.; Walsh, A.; Ehrick, R.; Xu, W.; May, K.; Liu, H. Chromatographic analysis of the acidic and basic species of recombinant monoclonal antibodies. *MAbs* **2012,** *4* (5), 578.

(49) Beck, A.; Sanglier-Cianferani, S.; Van Dorsselaer, A. Biosimilar, biobetter, and next generation antibody characterization by mass spectrometry. *Anal. Chem.* **2012,** *84* (11), 4637.

(50) Beck, A.; Wagner-Rousset, E.; Ayoub, D.; Van Dorsselaer, A.; Sanglier-Cianferani, S. Characterization of therapeutic antibodies and related products. *Anal. Chem.* **2013,** *85* (2), 715.

(51) Zhang, Z.; Pan, H.; Chen, X. Mass spectrometry for structural characterization of therapeutic antibodies. *Mass Spectrom. Rev.* **2009,** *28* (1), 147.

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

(53) Wang, L.; Amphlett, G.; Lambert, J. M.; Blattler, W.; Zhang, W. Structural characterization of a recombinant monoclonal antibody by electrospray time-of-flight mass spectrometry. *Pharm. Res.* **2005,** *22* (8), 1338.

(54) Johnson, K. A.; Paisley-Flango, K.; Tangarone, B. S.; Porter, T. J.; Rouse, J. C. Cation exchange-HPLC and mass spectrometry reveal C-terminal amidation of an IgG1 heavy chain. *Anal. Biochem.* **2007,** *360* (1), 75.

(55) Ayoub, D.; Jabs, W.; Resemann, A.; Evers, W.; Evans, C.; Main, L.; Baessmann, C.; Wagner-Rousset, E.; Suckau, D.; Beck, A. Correct primary structure assessment and extensive glyco-profiling of cetuximab by a combination of intact, middle-up, middle-down and bottom-up ESI and MALDI mass spectrometry techniques. *MAbs* **2013,** *5* (5), 699.

(56) Gahoual, R.; Busnel, J. M.; Beck, A.; Francois, Y. N.; Leize-Wagner, E. Full antibody primary structure and microvariant characterization in a single injection using transient isotachophoresis and sheathless capillary electrophoresis-tandem mass spectrometry. *Anal. Chem.* **2014,** *86* (18), 9074.

(57) Srzentic, K.; Fornelli, L.; Laskay, U. A.; Monod, M.; Beck, A.; Ayoub, D.; Tsybin, Y. O. Advantages of extended bottom-up proteomics using Sap9 for analysis of monoclonal antibodies. *Anal. Chem.* **2014,** *86* (19), 9945.

(58) Wang, Y.; Li, X.; Liu, Y. H.; Richardson, D.; Li, H.; Shameem, M.; Yang, X. Simultaneous monitoring of oxidation, deamidation, isomerization, and glycosylation of monoclonal antibodies by liquid chromatography-mass spectrometry method with ultrafast tryptic digestion. *MAbs* **2016,** *8* (8), 1477.

(59) Du, Y.; Wang, F.; May, K.; Xu, W.; Liu, H. Determination of deamidation artifacts introduced by sample preparation using 18O-labeling and tandem mass spectrometry analysis. *Anal. Chem.* **2012,** *84* (15), 6355.

(60) Zhang, T.; Zhang, J.; Hewitt, D.; Tran, B.; Gao, X.; Qiu, Z. J.; Tejada, M.; Gazzano-Santoro, H.; Kao, Y. H. Identification and characterization of buried unpaired cysteines in a recombinant monoclonal IgG1 antibody. *Anal. Chem.* **2012,** *84* (16), 7112.

(61) Xiang, T.; Chumsae, C.; Liu, H. Localization and quantitation of free sulfhydryl in recombinant monoclonal antibodies by differential labeling with 12C and 13C iodoacetic acid and LC-MS analysis. *Anal. Chem.* **2009,** *81* (19), 8101.

(62) Wang, Y.; Lu, Q.; Wu, S. L.; Karger, B. L.; Hancock, W. S. Characterization and comparison of disulfide linkages and scrambling patterns in therapeutic monoclonal antibodies: using LC-MS with electron transfer dissociation. *Anal. Chem.* **2011,** *83* (8), 3133.

(63) Pan, J.; Zhang, S.; Borchers, C. H. Comparative higher-order structure analysis of antibody biosimilars using combined bottom-up and top-down hydrogen-deuterium exchange mass spectrometry. *Biochim. Biophys. Acta* **2016,** *1864* (12), 1801.

(64) Li, J.; Rodnin, M. V.; Ladokhin, A. S.; Gross, M. L. Hydrogen-deuterium exchange and mass spectrometry reveal the pH-dependent conformational changes of diphtheria toxin T domain. *Biochemistry* **2014,** *53* (43), 6849.

(65) 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 Discov. Today* **2014,** *19* (1), 95.

(66) Sun, L.; Zhu, G.; Dovichi, N. J. Integrated capillary zone electrophoresis-electrospray ionization tandem mass spectrometry system with an immobilized trypsin microreactor for online digestion and analysis of picogram amounts of RAW 264.7 cell lysate. *Anal. Chem.* **2013,** *85* (8), 4187.

(67) Zhou, H.; Ning, Z.; Starr, A. E.; Abu-Farha, M.; Figeys, D. Advancements in top-down proteomics. *Anal. Chem.* **2012,** *84* (2), 720.

(68) Zhang, Z.; Shah, B. Characterization of variable regions of monoclonal antibodies by top-down mass spectrometry. *Anal. Chem.* **2007,** *79* (15), 5723.

(69) Bondarenko, P. V.; Second, T. P.; Zabrouskov, V.; Makarov, A. A.; Zhang, Z. Mass measurement and top-down HPLC/MS analysis of intact monoclonal antibodies on a hybrid linear quadrupole ion trap-Orbitrap mass spectrometer. *J. Am. Soc. Mass. Spectrom.* **2009,** *20* (8), 1415.

(70) Tsybin, Y. O.; Fornelli, L.; Stoermer, C.; Luebeck, M.; Parra, J.; Nallet, S.; Wurm, F. M.; Hartmer, R. Structural analysis of intact monoclonal antibodies by electron transfer dissociation mass spectrometry. *Anal. Chem.* **2011,** *83* (23), 8919.

(71) Fornelli, L.; Damoc, E.; Thomas, P. M.; Kelleher, N. L.; Aizikov, K.; Denisov, E.; Makarov, A.; Tsybin, Y. O. Analysis of intact monoclonal antibody IgG1 by electron transfer dissociation Orbitrap FTMS. *Mol. Cell. Proteomics* **2012,** *11* (12), 1758.

(72) Mao, Y.; Valeja, S. G.; Rouse, J. C.; Hendrickson, C. L.; Marshall, A. G. Top-down structural analysis of an intact monoclonal antibody by electron capture dissociation-Fourier transform ion cyclotron resonance-mass spectrometry. *Anal. Chem.* **2013,** *85* (9), 4239.

(73) Fornelli, L.; Ayoub, D.; Aizikov, K.; Beck, A.; Tsybin, Y. O. Middle-down analysis of monoclonal antibodies with electron transfer dissociation orbitrap fourier transform mass spectrometry. *Anal. Chem.* **2014,** *86* (6), 3005.

(74) Wang, B.; Gucinski, A. C.; Keire, D. A.; Buhse, L. F.; Boyne, M. T., 2nd. Structural comparison of two anti-CD20 monoclonal antibody drug products using middle-down mass spectrometry. *Analyst* **2013,** *138* (10), 3058.

(75) Nicolardi, S.; Deelder, A. M.; Palmblad, M.; van der Burgt, Y. E. Structural analysis of an intact monoclonal antibody by online electrochemical reduction of disulfide bonds and Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **2014,** *86* (11), 5376.

(76) An, Y.; Zhang, Y.; Mueller, H. M.; Shameem, M.; Chen, X. A new tool for monoclonal antibody analysis: application of IdeS proteolysis in IgG domain-specific characterization. *MAbs* **2014,** *6* (4), 879.

(77) Wang, D.; Wynne, C.; Gu, F.; Becker, C.; Zhao, J.; Mueller, H. M.; Li, H.; Shameem, M.; Liu, Y. H. Characterization of drug-product-related impurities and variants of a therapeutic monoclonal antibody by higher energy C-trap dissociation mass spectrometry. *Anal. Chem.* **2015,** *87* (2), 914.

(78) Cotham, V. C.; Brodbelt, J. S. Characterization of Therapeutic Monoclonal Antibodies at the Subunit-Level using Middle-Down 193 nm Ultraviolet Photodissociation. *Anal. Chem.* **2016,** *88* (7), 4004.

(79) Swatkoski, S.; Gutierrez, P.; Ginter, J.; Petrov, A.; Dinman, J. D.; Edwards, N.; Fenselau, C. Integration of residue-specific acid cleavage into proteomic workflows. *Journal of Proteome Research* **2007,** *6* (11), 4525.

(80) Smith, B. J. Chemical cleavage of polypeptides. *Methods Mol. Biol.* **2003,** *211*, 63.

(81) Raijmakers, R.; Neerincx, P.; Mohammed, S.; Heck, A. J. Cleavage specificities of the brother and sister proteases Lys-C and Lys-N. *Chem. Commun.* **2010,** *46* (46), 8827.

(82) Tang, H. Y.; Speicher, D. W. Identification of alternative products and optimization of 2-nitro-5-thiocyanatobenzoic acid cyanylation and cleavage at cysteine residues. *Anal. Biochem.* **2004,** *334* (1), 48.

(83) Crimmins, D. L.; Mische, S. M.; Denslow, N. D. Chemical cleavage of proteins in solution. *Curr. Protoc. Protein Sci.* **2005,** *Chapter 11*, Unit 11 4.

(84) Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M. C.; Yates, J. R., 3rd. Protein analysis by shotgun/bottom-up proteomics. *Chem. Rev.* **2013,** *113* (4), 2343.

(85) Switzar, L.; Giera, M.; Niessen, W. M. Protein digestion: an overview of the available techniques and recent developments. *J. Proteome. Res.* **2013,** *12* (3), 1067.

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

(87) Wu, C.; Tran, J. C.; Zamdborg, L.; Durbin, K. R.; Li, M.; Ahlf, D. R.; Early, B. P.; Thomas, P. M.; Sweedler, J. V.; Kelleher, N. L. A protease for 'middle-down' proteomics. *Nat. Methods* **2012,** *9* (8), 822.

(88) Huesgen, P. F.; Lange, P. F.; Rogers, L. D.; Solis, N.; Eckhard, U.; Kleifeld, O.; Goulas, T.; Gomis-Ruth, F. X.; Overall, C. M. LysargiNase mirrors trypsin for protein C-terminal and methylation-site identification. *Nat. Methods* **2015,** *12* (1), 55.

(89) Capelo, J. L.; Carreira, R.; Diniz, M.; Fernandes, L.; Galesio, M.; Lodeiro, C.; Santos, H. M.; Vale, G. Overview on modern approaches to speed up protein identification workflows relying on enzymatic cleavage and mass spectrometry-based techniques. *Anal. Chim. Acta* **2009,** *650* (2), 151.

(90) Park, Z. Y.; Russell, D. H. Thermal denaturation: a useful technique in peptide mass mapping. *Anal. Chem.* **2000,** *72* (11), 2667.

(91) Havlis, J.; Thomas, H.; Sebela, M.; Shevchenko, A. Fast-response proteomics by accelerated in-gel digestion of proteins. *Anal. Chem.* **2003,** *75* (6), 1300.

(92) Pramanik, B. N.; Mirza, U. A.; Ing, Y. H.; Liu, Y. H.; Bartner, P. L.; Weber, P. C.; Bose, A. K. Microwave-enhanced enzyme reaction for protein mapping by mass spectrometry: a new approach to protein digestion in minutes. *Protein Sci.* **2002,** *11* (11), 2676.

(93) Lopez-Ferrer, D.; Capelo, J. L.; Vazquez, J. Ultra fast trypsin digestion of proteins by high intensity focused ultrasound. *J. Proteome. Res.* **2005,** *4* (5), 1569.

(94) Shin, S.; Yang, H. J.; Kim, J.; Kim, J. Effects of temperature on ultrasound-assisted tryptic protein digestion. *Anal. Biochem.* **2011,** *414* (1), 125.

(95) Rial-Otero, R.; Carreira, R. J.; Cordeiro, F. M.; Moro, A. J.; Santos, H. M.; Vale, G.; Moura, I.; Capelo, J. L. Ultrasonic assisted protein enzymatic digestion for fast protein identification by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Sonoreactor versus ultrasonic probe. *J. Chromatogr. A* **2007,** *1166* (1-2), 101.

(96) Lopez-Ferrer, D.; Petritis, K.; Hixson, K. K.; Heibeck, T. H.; Moore, R. J.; Belov, M. E.; Camp, D. G., 2nd; Smith, R. D. Application of pressurized solvents for ultrafast trypsin hydrolysis in proteomics: proteomics on the fly. *J. Proteome. Res.* **2008,** *7* (8), 3276.

(97) Jones, L. M.; Zhang, H.; Vidavsky, I.; Gross, M. L. Online, high-pressure digestion system for protein characterization by hydrogen/deuterium exchange and mass spectrometry. *Anal. Chem.* **2010,** *82* (4), 1171.

(98) Wang, S.; Liu, T.; Zhang, L.; Chen, G. Efficient chymotryptic proteolysis enhanced by infrared radiation for peptide mapping. *J. Proteome. Res.* **2008,** *7* (11), 5049.

(99) Ma, J. F.; Zhang, L. H.; Liang, Z.; Shan, Y. C.; Zhang, Y. K. Immobilized enzyme reactors in proteomics. *Trac-Trend Anal. Chem.* **2011,** *30* (5), 691.

(100) Regnier, F. E.; Kim, J. Accelerating trypsin digestion: the immobilized enzyme reactor. *Bioanalysis* **2014,** *6* (19), 2685.

(101) Monzo, A.; Sperling, E.; Guttman, A. Proteolytic enzyme-immobilization techniques for MS-based protein analysis. *Trac-Trend Anal. Chem.* **2009,** *28* (7), 854.

(102) Ma, J. F.; Zhang, L. H.; Liang, Z.; Zhang, W. B.; Zhang, Y. K. Recent advances in immobilized enzymatic reactors and their applications in proteome analysis. *Anal. Chim. Acta* **2009,** *632* (1), 1.

(103) Long, Y.; Wood, T. D. Immobilized pepsin microreactor for rapid peptide mapping with nanoelectrospray ionization mass spectrometry. *J. Am. Soc. Mass. Spectrom.* **2015,** *26* (1), 194.

(104) Temporini, C.; Calleri, E.; Campese, D.; Cabrera, K.; Felix, G.; Massolini, G. Chymotrypsin immobilization on epoxy monolithic silica columns: development and characterization of a bioreactor for protein digestion. *J. Sep. Sci.* **2007,** *30* (17), 3069.

(105) Prikryl, P.; Ticha, M.; Kucerova, Z. Immobilized endoproteinase Glu-C to magnetic bead cellulose as a tool in proteomic analysis. *J. Sep. Sci.* **2013,** *36* (12), 2043.

(106) Calleri, E.; Temporini, C.; Perani, E.; De Palma, A.; Lubda, D.; Mellerio, G.; Sala, A.; Galliano, M.; Caccialanza, G.; Massolini, G. Trypsin-based monolithic bioreactor coupled on-line with LC/MS/MS system for protein digestion and variant identification in standard solutions and serum samples. *J. Proteome. Res.* **2005,** *4* (2), 481.

(107) Temporini, C.; Perani, E.; Mancini, F.; Bartolini, M.; Calleri, E.; Lubda, D.; Felix, G.; Andrisano, V.; Massolini, G. Optimization of a trypsin-bioreactor coupled with high-performance liquid chromatography-electrospray ionization tandem mass spectrometry for quality control of biotechnological drugs. *J. Chromatogr. A* **2006,** *1120* (1-2), 121.

(108) Krenkova, J.; Bilkova, Z.; Foret, F. Characterization of a monolithic immobilized trypsin microreactor with on-line coupling to ESI-MS. *J. Sep. Sci.* **2005,** *28* (14), 1675.

(109) Kato, M.; Inuzuka, K.; Sakai-Kato, K.; Toyo'oka, T. Monolithic bioreactor immobilizing trypsin for high-throughput analysis. *Anal. Chem.* **2005,** *77* (6), 1813.

(110) Feng, S.; Ye, M.; Jiang, X.; Jin, W.; Zou, H. Coupling the immobilized trypsin microreactor of monolithic capillary with muRPLC-MS/MS for shotgun proteome analysis. *J. Proteome. Res.* **2006,** *5* (2), 422.

(111) Ota, S.; Miyazaki, S.; Matsuoka, H.; Morisato, K.; Shintani, Y.; Nakanishi, K. High-throughput protein digestion by trypsin-immobilized monolithic silica with pipette-tip formula. *J. Biochem. Biophys. Methods* **2007,** *70* (1), 57.

(112) Schoenherr, R. M.; Ye, M.; Vannatta, M.; Dovichi, N. J. CE-microreactor-CE-MS/MS for protein analysis. *Anal. Chem.* **2007,** *79* (6), 2230.

(113) Nicoli, R.; Gaud, N.; Stella, C.; Rudaz, S.; Veuthey, J. L. Trypsin immobilization on three monolithic disks for on-line protein digestion. *J. Pharm. Biomed. Anal.* **2008,** *48* (2), 398.

(114) Ma, J.; Liang, Z.; Qiao, X.; Deng, Q.; Tao, D.; Zhang, L.; Zhang, Y. Organic-inorganic hybrid silica monolith based immobilized trypsin reactor with high enzymatic activity. *Anal. Chem.* **2008,** *80* (8), 2949.

(115) Krenkova, J.; Lacher, N. A.; Svec, F. Highly efficient enzyme reactors containing trypsin and endoproteinase LysC immobilized on porous polymer monolith coupled to MS suitable for analysis of antibodies. *Anal. Chem.* **2009,** *81* (5), 2004.

(116) Spross, J.; Sinz, A. A capillary monolithic trypsin reactor for efficient protein digestion in online and offline coupling to ESI and MALDI mass spectrometry. *Anal. Chem.* **2010,** *82* (4), 1434.

(117) Ruan, G.; Wu, Z.; Huang, Y.; Wei, M.; Su, R.; Du, F. An easily regenerable enzyme reactor prepared from polymerized high internal phase emulsions. *Biochem. Biophys. Res. Commun.* **2016,** *473* (1), 54.

(118) Zhang, Z.; Sun, L.; Zhu, G.; Cox, O. F.; Huber, P. W.; Dovichi, N. J. Nearly 1000 Protein Identifications from 50 ng of Xenopus laevis Zygote Homogenate Using Online Sample Preparation on a Strong Cation Exchange Monolith Based Microreactor Coupled with Capillary Zone Electrophoresis. *Anal. Chem.* **2016,** *88* (1), 877.

(119) Slysz, G. W.; Schriemer, D. C. Blending protein separation and peptide analysis through real-time proteolytic digestion. *Anal. Chem.* **2005,** *77* (6), 1572.

(120) Slysz, G. W.; Lewis, D. F.; Schriemer, D. C. Detection and identification of sub-nanogram levels of protein in a nanoLC-trypsin-MS system. *J. Proteome. Res.* **2006,** *5* (8), 1959.

(121) Yamaguchi, H.; Miyazaki, M.; Honda, T.; Briones-Nagata, M. P.; Arima, K.; Maeda, H. Rapid and efficient proteolysis for proteomic analysis by protease-immobilized microreactor. *Electrophoresis* **2009,** *30* (18), 3257.

(122) Li, Y.; Yan, B.; Deng, C.; Yu, W.; Xu, X.; Yang, P.; Zhang, X. Efficient on-chip proteolysis system based on functionalized magnetic silica microspheres. *Proteomics* **2007,** *7* (14), 2330.

(123) Jeng, J.; Lin, M. F.; Cheng, F. Y.; Yeh, C. S.; Shiea, J. Using high-concentration trypsin-immobilized magnetic nanoparticles for rapid in situ protein digestion at elevated temperature. *Rapid Commun. Mass Spectrom.* **2007,** *21* (18), 3060.

(124) Moore, S.; Hess, S.; Jorgenson, J. Characterization of an immobilized enzyme reactor for on-line protein digestion. *J. Chromatogr. A* **2016,** *1476*, 1.

(125) Freije, J. R.; Mulder, P. P.; Werkman, W.; Rieux, L.; Niederlander, H. A.; Verpoorte, E.; Bischoff, R. Chemically modified, immobilized trypsin reactor with improved digestion efficiency. *J. Proteome. Res.* **2005,** *4* (5), 1805.

(126) Liu, Y.; Lu, H.; Zhong, W.; Song, P.; Kong, J.; Yang, P.; Girault, H. H.; Liu, B. Multilayer-assembled microchip for enzyme immobilization as reactor toward low-level protein identification. *Anal. Chem.* **2006,** *78* (3), 801.

(127) Liuni, P.; Rob, T.; Wilson, D. J. A microfluidic reactor for rapid, low-pressure proteolysis with on-chip electrospray ionization. *Rapid Commun. Mass Spectrom.* **2010,** *24* (3), 315.

(128) Cooper, J. W.; Chen, J.; Li, Y.; Lee, C. S. Membrane-based nanoscale proteolytic reactor enabling protein digestion, peptide separation, and protein identification using mass spectrometry. *Anal. Chem.* **2003,** *75* (5), 1067.

(129) Xu, F.; Wang, W. H.; Tan, Y. J.; Bruening, M. L. Facile trypsin immobilization in polymeric membranes for rapid, efficient protein digestion. *Anal. Chem.* **2010,** *82* (24), 10045.

(130) Tan, Y. J.; Wang, W. H.; Zheng, Y.; Dong, J.; Stefano, G.; Brandizzi, F.; Garavito, R. M.; Reid, G. E.; Bruening, M. L. Limited proteolysis via millisecond digestions in protease-modified membranes. *Anal. Chem.* **2012,** *84* (19), 8357.

(131) Pang, Y.; Wang, W. H.; Reid, G. E.; Hunt, D. F.; Bruening, M. L. Pepsin-Containing Membranes for Controlled Monoclonal Antibody Digestion Prior to Mass Spectrometry Analysis. *Anal. Chem.* **2015,** *87* (21), 10942.

(132) Ning, W.; Bruening, M. L. Rapid Protein Digestion and Purification with Membranes Attached to Pipet Tips. *Anal. Chem.* **2015,** *87* (24), 11984.

Chapter 2 . Pepsin-Containing Membranes for Controlled Monoclonal Antibody Digestion Prior to Mass Spectrometry Analysis

(Part of this chapter was originally published in *Analytical Chemistry*. Reprinted with permission from Pang, Y., Wang, W. H., Reid, G. E., Hunt, D. F., Bruening, M. L. *Anal. Chem.* 2015, 87 (21), 10942-9. Copyright (2015) American Chemical Society. Work on "top-down" analysis was performed by Weihan Wang and Donald F. Hunt at the University of Virginia.)

Monoclonal antibodies (mAbs) are the fastest growing class of therapeutic drugs because of their high specificities to target cells. Facile analysis of therapeutic mAbs and their post-translational modifications (PTMs) is essential for quality control, and mass spectrometry (MS) is the most powerful tool for antibody characterization. This study uses pepsin-containing nylon membranes as controlled proteolysis reactors for mAb digestion prior to ultra-high resolution Orbitrap MS analysis. Variation of the residence times (3 ms to 3 s) of antibody solutions in the membranes yields "bottom-up" (1-2 kDa) to "middle-down" (5-15 kDa) peptide sizes in less than 10 min. These peptides cover the entire sequences of Trastuzumab and a Waters^{TM} antibody, and a proteolytic peptide comprised of 140 amino acids from the Waters^{TM} antibody contains all three complementarity determining regions on the light chain. This work compares the performance of "bottom-up" (in-solution tryptic digestion), "top-down" (intact protein fragmentation) and "middle-down" (in-membrane digestion) analysis of an antibody light chain. Data from tandem MS show 99%, 55%, and 99% bond cleavage for "bottom-up", "top-down", and "middle-down" analyses, respectively. In-membrane digestion also facilitates detection of PTMs such as

oxidation, deamidation, N-terminal pyroglutamic acid formation and glycosylation. Compared to "bottom-up" and "top-down" approaches for antibody characterization, in-membrane digestion uses minimal sample preparation time, and this technique also yields high peptide and sequence coverage for identification of PTMs.

2.1 Introduction

Monoclonal antibodies (mAbs) have emerged as an important class of biotherapeutic drugs with high selectivity and specificity,[1-3] and the U.S. Food and Drug Administration (FDA) has approved more than 35 antibodies[4] for treatment of diseases such as breast cancer,[5] non-Hodgkin lymphoma,[6] and colorectal cancer.[7] According to the FDA's "quality by design" policy, biotherapeutic materials such as mAbs must adhere to a consistent, predefined quality during manufacturing.[8] Facile mAb characterization, especially in the complementarity determining regions (CDRs), is vital for quality control, not only because these proteins are vulnerable to chemical modifications during expression, purification and long-term storage, but also because they have natural heterogeneities.[9-11] Common PTMs on mAbs include methionine oxidation, asparagine deamidation, asparagine glycosylation in the heavy-chain constant region 2 (CH2), and heavy chain C-terminal processing.[12-21]

MS is the most powerful tool for antibody characterization because of its high resolution and mass accuracy within a wide dynamic range. Current MS-based strategies for antibody characterization employ "top-down",[22-25] "bottom-up"[26-35] and "middle-down"[36-40] approaches with ultra-high resolution time-of-flight (TOF),[24,27,29,30,33-35,39] Fourier transform ion cyclotron resonance[25,41] and Orbitrap[22,23,26,36-38,40] mass spectrometers. "Top-down" methods introduce the intact antibody into the mass spectrometer through liquid chromatography (LC) or direct infusion.

Determination of the intact protein mass and subsequent gas-phase fragmentation via collision-induced dissociation (CID),[22,23] higher-energy collision dissociation (HCD), electron capture dissociation (ECD),[25] or/and electron transfer dissociation (ETD)[24] give an overview of the major PTMs with minimal sample manipulation time. Unfortunately PTMs with small mass changes, e.g. deamidation (+1 Da), cannot be detected, and sequence coverage for "top-down" analysis typically reaches only ~35%.[25,42] The incomplete fragmentation likely results from highly structured and disulfide bond-protected areas.[41]

The "bottom-up" method uses enzymatic antibody digestion followed by LC and tandem mass spectrometry (MS/MS) analysis to provide accurate mass values and product ions that imply the sequences of individual peptides. However, protein digestion typically requires several time-consuming steps during which antibody modification may occur.[43,44] Trypsin digestion, for instance, generally includes antibody denaturation and reduction followed by alkylation of thiol groups to prevent reforming of disulfide bonds. Moreover, proteolysis usually takes place at 37 ℃ overnight, and basic digestion conditions favor deamidation of asparagine.[45] Peptide coverage is frequently incomplete because of weak ionization efficiencies for some peptides along with loss of a few peptides during LC.[46] Nevertheless, analysis of several digests catalyzed by different proteases often yields 100% sequence coverage.[47] Recently, Srzentić et al. reported that digestion using the enzyme Sap 9 yields relatively large peptides (compared to tryptic digestion) and enables extended "bottom-up" LC-MS/MS analysis with nearly 100% peptide coverage for both light chain and heavy chains.[26] Importantly, this enzyme functions in acidic conditions that limit deamidation and avoid the need for alkylation of cysteine. However, Sap9 must be recombinantly expressed and purified.

Analysis of larger peptides (3-20 kDa) obtained from limited digestion is usually termed a "middle-down" approach.[48] Ultra-high resolution mass spectrometers can resolve the isotopic distributions of these peptides, and their relatively large size enhances the total peptide coverage (relative to tryptic digestion), which increases the sequence coverage when fragmentation is complete. Additionally, compared to "bottom-up" methods, the large size of "middle-down" peptides increases the probability that two or more PTMs will occur on the same peptide to enable correlation of these PTMs.[49] In one "middle-down" strategy that yields peptides with masses around 25 kDa, papain cleaves antibodies at the hinge region and forms subunits such as Fab or $F(ab')_2$, but the digestion is sometimes difficult to control.[11] Fornelli and coworkers employed the Immunoglobulin-degrading enzyme of Streptococcus (IdeS) to fragment Adalimumab into $F(ab')_2$ and Fc portions at the G-G bond below the hinge region.[38] After reduction and denaturation, they obtained Fd, Lc and Fc/2 fragments, and analyzed these large peptides with LC-ETD MS/MS. Sequence coverage reached 70%. The large size of these peptides may still makes detection of deamidation difficult, and incomplete fragmentation limits sequence coverage.

In addition to varying the digestion enzyme, limiting the digestion time may yield the large peptides required for "middle-down" protein characterization. Recently, Tan and coworkers adsorbed pepsin in the pores of nylon membranes and found that the lengths of proteolytic peptides from myoglobin and bovine serum albumin vary with the residence time of the protein solution in the membrane, i.e. shorter residence times generate longer peptides.[50] This method can generate both "bottom-up" (long residence times) and "middle-down" (short residence times) peptides using a single enzyme, so control over digestion yields peptides with overlapping sequences. Small peptides give detailed sequence information, whereas larger peptides lead to

higher coverage. The large peptides may also contain many basic residues that lead to higher charge states, which is beneficial for ETD fragmentation.[51]

This research employs antibody proteolysis in pepsin-containing porous membranes to decrease the time and cost of digestion, increase both the peptide and sequence coverages in MS analysis, and limit antibody modification during digestion. Figure 2.1 shows the work flow for digestion and analysis. Porous nylon membranes and pepsin are inexpensive and readily available, and the high enzyme concentration in membrane pores allows digestion in a few minutes.

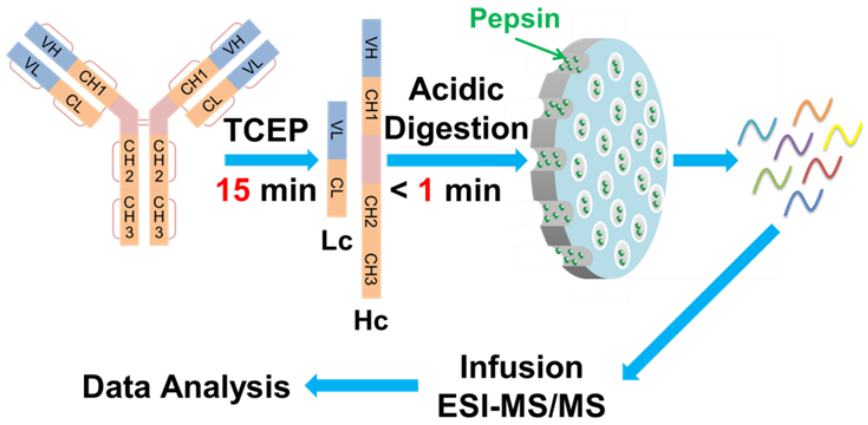


Figure 2.1. Workflow for controlled digestion and analysis of antibodies. Acronyms: TCEP-tris(2-carboxyethyl) phosphine, VH-variable region of the heavy chain; CH1, CH2, and CH3-different constant regions of the heavy chain; CL-constant region of the light chain; VL-variable region of the light chain; Lc-light chain; Hc- heavy chain.

Moreover, acidic digestion conditions limit deamidation and do not require protection of the thiol groups of cysteine. Other recent substrates for pepsin immobilization include aldehyde-modified polymethacrylate monoliths[52] and fused-silica capillaries,[53] but such supports do not readily afford the ms residence times available with in-membrane digestion. This study investigates in-membrane digestion of the entire antibody without separation of the light and heavy chains as

well as digestion of separated chains. Remarkably, digestion of 35 pmol of a reduced Waters™ antibody (WIgG1) occurs in less than 1 min with 100% peptide coverage of the light and heavy chains. We further demonstrate the benefits of this digestion strategy by comparing MS analyses of a monoclonal antibody light chain after digestion in a pepsin-containing membrane, after traditional in-solution digestion and using a "top-down" method.

2.2 Experimental

2.2.1 Materials

A monoclonal immunoglobulin G was purchased from Waters (WIgG1, Intact mAb Mass Check Standard, 186006552), and Trastuzumab (Herceptin, Genentech) was dissolved in a phosphate buffer (KH_2PO_4 144 mg/L, NaCl 9000 mg/L, and $Na_2HPO_4 \cdot 7H_2O$ 795 mg/L, pH 7.4, Thermo Fisher) at a concentration of 21 mg/mL. Nylon membranes (LoProdyne LP, pore size 1.2 μm, 110 μm thickness) were acquired from Pall Corporation. The holder for membrane digestion (flangeless fitting system, Upchurch Scientific, A-424) was connected to 1/16 inch OD tubing via ferrules.[50] Pepsin from porcine gastric mucosa (lyophilized powder, 3200-4500 units/mg protein), iodoacetamide (IAM, ⩾99%), polystyrene sulfonate (PSS, average molecular weight ~70,000), and acetonitrile (ACN, HPLC grade, ⩾99.9%) were obtained from Sigma Aldrich. Isopropyl alcohol (IPA, MACRON), sequencing grade modified trypsin (Promega), and trifluoroacetic acid (TFA, purchased from EMD) were used as received. Important chemicals for reduction and digestion include tris(2-carboxyethyl) phosphine hydrochloride (TCEP-HCl, >98%, Fluka), acetic acid (HOAc, Mallinckrodt, ACS), formic acid (>96%, Spectrum), and ammonium bicarbonate (Columbus Chemical).

2.2.2 Modification of Membranes with Pepsin

We previously described pepsin-containing nylon membranes,[50] but this work uses membranes with a nominal pore size of 1.2 μm instead of 0.45 μm. The modification procedure includes sequential adsorption of PSS and pepsin, and the amount of pepsin adsorbed to the nylon membrane was estimated by determining the pepsin concentration in the loading solution before and after circulation through the membrane. A Nanodrop UV-Vis spectrometer (NanoDrop 2000, Thermo) measured the pepsin UV absorbance at 280 nm.

2.2.3 mAb Reduction and Characterization

Antibodies were dissolved (WIgG1) or diluted (Trastuzumab) in deionized water to prepare 1 mg/mL stock solutions. The solution was stored at 4 ºC until use. For antibody reduction, 1 μL of 0.1 M HOAc and 1 μL of 0.1 M TCEP-HCl were added to 10 μL of antibody stock solution, and this reaction mixture was incubated at 75 ºC for 15 min and finally diluted with 88 μL of 5% FA. Alkylation of antibodies after reduction was conducted only prior to in-solution, tryptic digestion. In that case, 20 μg of WIgG1 was dried and reconstituted in 7 μL of 2 mM TCEP-HCl solution prepared in 0.1 % HOAc containing 8 M urea. This mixture was incubated at 50 ℃ for 10 min, and 7 μL of 20 mM IAM in a 2 M NH_4HCO_3 solution containing 8 M urea was added. After incubation in the dark for 30 min, 6 μL of 30 mM dithiothreitol in 100 mM NH_4HCO_3 solution containing 8 M urea was added. The reaction was incubated in the dark for 20 min to quench the IAM. Ultra-performance liquid chromatography coupled with electrospray ionization quadrupole time-of-flight mass spectrometry (UPLC-ESI-QTOF-MS) confirmed separation of the light and heavy chains. Reduced antibody light and heavy chains were separated with an ACQUITY UPLC Protein BEH C4 column (1.7 μm diameter, 300Å pore size, 1 mm I.D. × 100 mm). UPLC

(Waters Open Architecture UPLC system) was performed at 40 ℃ using a flow rate of 0.2 mL/min. In the gradient elution, solution A contained 0.1% FA in H_2O, whereas solution B was 100% ACN. The applied gradient was 5 to 25% solution B in 2 min, 25 to 55% solution B in 12 min, and 55 to 90% solution B in 25 min. UPLC-ESI-QTOF-MS analysis of reduced antibody was conducted on a Waters Xevo G2-S QTOF mass spectrometer in positive-ion, sensitivity mode. The mass spectrometry parameters included: capillary voltage = 2.50 kV, sample cone voltage = 30.0 V, source offset = 12 V, source temperature = 100 ℃, desolvation temperature = 350 ℃, cone gas flow = 10 L/h, and desolvation gas flow = 600 L/h. The light-chain and heavy-chain signals were deconvoluted with the MaxEnt1 function in the MassLynx software. Parameters were set as follows: m/z input range = 400–3000; output resolution = 0.2 Da/channel; output mass range = 22000-27000 Da for the light chain and 48000–52000 Da for the heavy chain; uniform Gaussian width at half height of 0.5 Da for both the light and heavy chain; minimum intensity ratios of 50% for left and right; and a maximum of 10 iterations. Offline HPLC isolation of light and heavy chains (prior to digestion of these chains) was performed with a Shimadzu LC-20AB instrument equipped with a SPD-20AV ultraviolet detector. After reduction, the light and heavy chain mixture was injected onto a reversed phase TSKgel Protein C4-300 column (3.0 μm diameter, 300 Å pore size, 4.6 mm I.D. × 150 mm). LC was performed at 40 ℃ using a flow rate of 1 mL/min. In the gradient elutions, solution A contained 0.05% TFA, 10% ACN, and 89.95% H_2O, whereas solution B consisted of 0.045% TFA, 70% ACN, 9.955% IPA, and 20% H_2O. Two different gradients were used to separate nonalkylated and alkylated chains. For nonalkylated antibody, the applied gradient was 0% to 30% solution B in 5 min, 30% to 60% solution B in 30 min, and 60% to 100% solution B in 5 min. For alkylated antibody, the gradient was 0% to 30% solution B in 5 min, 30% to 65% solution B in 35 min, and 65% to 100%

solution B in 10 min. UV absorbances at 215 nm and 280 nm were used to monitor the elution of antibody chains. Effluent was collected offline, dried with a SpeedVac, and analyzed by SDS-PAGE with Coomassie Blue staining.

2.2.4 In-Membrane Digestion of Intact Antibody

After reduction by TCEP, the nonalkylated mixture was passed through a pepsin-containing membrane in an Upchurch holder at flow rates of 0.13 or 130 mL/h using a syringe pump. The residence time was estimated assuming a nylon membrane porosity of 50% and an exposed membrane area of 0.02 cm^2 (see equation (2-1) below). Hence, residence times were 3 s and 3 ms for 0.13 and 130 mL/h flow rates, respectively, and 100 μL of effluent was collected for direct infusion MS analysis.

2.2.5 Digestion of the mAb Light and Heavy Chains

2.2.5.1 In-membrane digestion

Ten μg of nonalkylated light chain (with the assumption that all of the light chain was recovered from the separation) was dissolved in 100 μL of 2 mM TCEP in 5% FA solution. The mixture was heated at 80 ℃ for 10 min, allowed to cool, and the light chain solution was passed through the membrane at flow rates of 0.13, 13 and 130 mL/h. Effluent was collected, dried with a SpeedVac and saved for MS analysis. The heavy chain was digested similarly.

2.2.5.2 In-solution digestion

Four μg of alkylated light chain (with assumption that all of the light chain was recovered from the separation) was dissolved in 10 μL of 2 mM TCEP solution containing 10 mM NH_4HCO_3.

The mixture was heated at 80 ℃ for 10 min, and cooled to room temperature. Two μL of 0.1 μg/μL trypsin solution was added to the mixture prior to incubation at 37 ℃ for 16 h. The reaction was quenched by addition of 5 μL of acetic acid, immediately frozen with liquid nitrogen, and dried with a SpeedVac before reconstitution for MS.

2.2.6 "Top-down" Analysis of a mAb Light Chain

An Agilent Technologies (Palo Alto, CA) 1100 Series binary HPLC system was interfaced with a Thermo Fisher Scientific Orbitrap Elite^{TM} Hybrid Ion Trap-Orbitrap Mass Spectrometer (San Jose, CA) for online separation of TCEP-reduced WIgG1. About 100 fmol of reduced WIgG1 was pressure-loaded onto a fused silica capillary column (75 μm I.D. × 360 μm O.D.) packed with 10 cm of Agilent POROSHELL 300SB-C18 particles (5 μm diameter, 300 Å pore size). The back end of the column was equipped with a laser-pulled nanoelectrospray emitter tip,[54] and the column was initially rinsed for 10 min with 0.3% formic acid in water to remove salts. Protein sample was eluted at 60 ℃ at a flow rate of 100 nL/min using the following gradient: 0-30% B for 5 min, 30-50% B for 20 min, 50-100% B for 5 min. Solution A contained 0.3% FA in H_2O, whereas solution B consisted of 0.3% FA, 72% ACN, 18% IPA and 9.7% H_2O.

2.2.7 Mass Spectrometry and Data Analysis

After drying with a Speedvac, in-membrane and in-solution digests were reconstituted in 1% acetic acid, 49% H_2O, and 50% methanol, loaded into a Whatman multichem 96-well plate (Sigma Aldrich) and sealed with Teflon Ultrathin Sealing Tape (Analytical Sales and Services, Prompton Plains, NJ). An Advion Triversa Nanomate nanoelectrospray ionization (nESI) source (Advion, Ithaca, NY) was used to introduce the sample into a high-resolution accurate mass Thermo Fisher Scientific LTQ Orbitrap VelosTM mass spectrometer (San Jose, CA) equipped

with a dual pressure ion trap, HCD cell, and ETD. The spray voltage and gas pressure were set to 1.4 kV and 1.0 psi, respectively. The ion source interface had an inlet temperature of 200 ℃ with an S-Lens value of 65%. High-resolution mass spectra were acquired in positive ionization mode across the m/z range of 300-2000 using the FT analyzer operating at 100 000 mass resolving power. Spectra were the average of 100 scans. Mass spectra were deconvoluted using the Xtract function of the XCalibur software. Proteolytic peptide identification and CID/HCD/ETD MS/MS data analysis were performed manually (for isotopic distributions with signal to noise >5) by matching MS and MS/MS product ions with data generated in silico using ProteinProspector (v 5.14.1 University of California, San Francisco). Mass tolerance was set for 5 ppm.

"Top-down" MS analyses of the WIgG1 light chain include a full MS scan at m/z 300-2000 in the Orbitrap at 240,000 mass resolving power, and three MS/MS scans (5-ms ETD, 15-ms ETD, and CID targeted on the +28 charge-state Lc ion at m/z 865.2 with a 3 m/z isolation window) in the Orbitrap at 120,000 mass resolving power (5 microscans per MS/MS scan). The Lc-targeted ETD (13 scans) or CID (9 scans) MS/MS spectra were merged and extracted from the raw file using Xcalibur™ 2.1 (Thermo Scientific). Each extracted ETD or CID spectrum was then searched against the sequence of WIgG1 (provided by Waters) using ProSightPC 3.0. Search parameters included: 5 Da precursor tolerance (monoisotopic), 15 ppm fragment tolerance (monoisotopic), Δm mode on, and disulfide off. The c-, z-, b- and y-type fragment ions assigned by ProSightPC 3.0 were manually verified before acceptance.

2.3 Results and discussion

2.3.1 Protease-containing Membranes

Figure 2.1 shows the workflow we employ to analyze mAbs using in-membrane digestion. The procedure exploits enzyme-containing membranes that we prepare using sequential adsorption of PSS and pepsin in nylon membranes at pH 2.3. PSS adsorption provides a negatively charged surface that captures pepsin, which is positively charged at low pH. Similar to other membrane modifications through electrostatic adsorption, protease immobilization should occur throughout the membrane.[55] These reactors are very active because of the high local enzyme concentration in membrane pores.[50] The extent of digestion varies with the solution residence time in the membrane, t_{res}, which is a function of the membrane thickness, l, the volumetric flow rate, Q, the exposed area at the faces of the membrane, A, and the membrane porosity, ε (equation (2-1)).

$$t_{res}=lA\varepsilon/Q \tag{2-1}$$

Our prior study used a membrane with nominal 0.45 μm pores and a thickness of 170 μm,[50] whereas this work employs both a larger pore size (1.2 μm) and a lower thickness (110 μm) to further limit digestion and provide longer peptides. The lower thickness decreases the residence time for a given flow rate, and the larger pore size should give longer radial diffusion distances to immobilized enzymes. Analysis of the pepsin loading solution before and after circulating through the membrane suggests an immobilized pepsin concentration of ~70 mg per mL of membrane.

2.3.2 mAb Reduction and Characterization

Effective antibody digestion requires reduction of disulfide bonds to give the protease access to cleavage sites. We employ TCEP as a reducing agent because it can function under acidic conditions that both prevent reformation of disulfide bonds and partially denature the antibody. Thus, acidic conditions avoid the need for urea denaturation and alkylation. Because pepsin is enzymatically active at pH 2-3, peptic digests are compatible with ESI-MS without further purification. UPLC-ESI-QTOF-MS analysis verified the separation of light and heavy chains after TCEP reduction (Figure 2.2). Based on deconvolution using MaxEnt1 software, the average mass of the antibody light chain is 24198.0, which agrees with the theoretical mass of 24197.7 provided by the manufacturer. The deconvoluted heavy chain mass spectrum shows three glycoforms with mass differences of 162 Da.

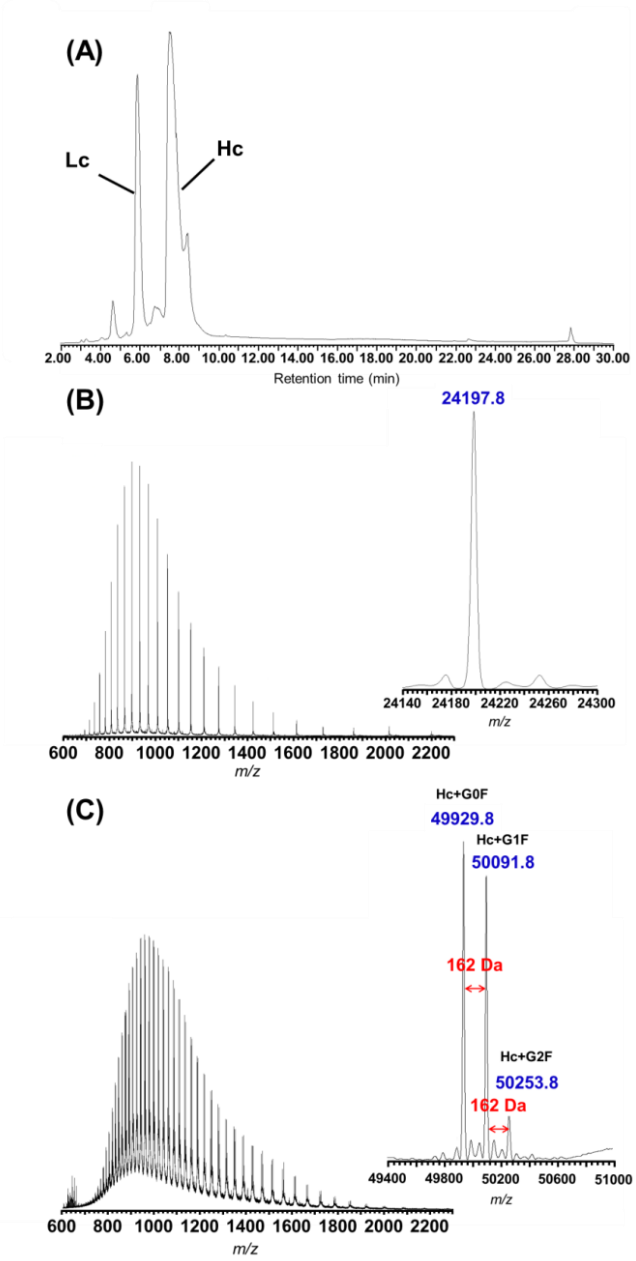


Figure 2.2. mAb Reduction and Characterization. (A) Chromatogram showing the UPLC separation of the WIgG1 light and heavy chains. (B) Mass spectrum at a retention time of 5.85 min showing the light chain isotopic envelope. The inset is the deconvoluted mass spectrum. (C) Mass spectrum at a retention time of 7.51 min showing the heavy chain isotopic envelope. The inset is the deconvoluted mass spectrum. See Figure 2.9 for glycoform assignments.

2.3.3 mAb Digestion in Membranes

In-membrane digestion occurs during passage of a reduced-antibody solution through a pepsin-containing membrane. At a flow rate of 130 mL/h, which corresponds to a 3-ms residence time, digestion of a 100-µL solution requires less than one minute. As Figure 2.3 shows, the infusion mass spectrum of digested WIgG1 contains signals from two large peptides, L1-140 (light-chain amino acids 1-140) and L141-219, which cover the entire light-chain sequence. L1-140 shows consecutive charge distribution isotopic envelopes from +8 to +14, and this large peptide (15 kDa) covers the CDR-L1, CDR-L2 and CDR-L3 regions of the antibody.

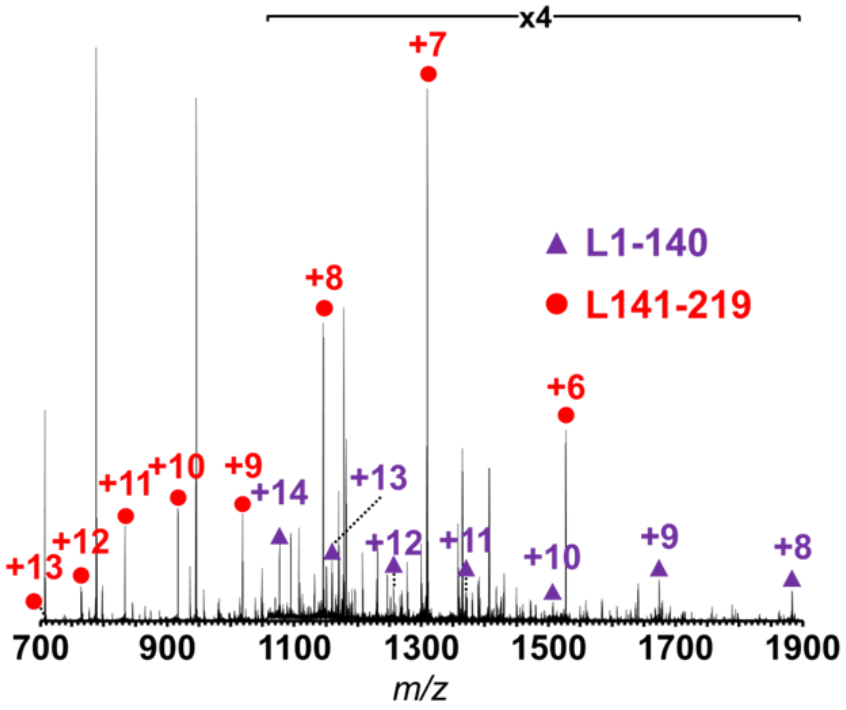


Figure 2.3. Part of the mass spectrum of an in-membrane digest (3-ms residence time) of reduced WIgG1 antibody. The labeled signals show the charge-state distributions of peptides containing the amino acids 1-140 (purple), and 141-219 (red) of the light chain. These two large peptides cover the entire light chain sequence.

Similarly, the other large peptide with light-chain amino acids 141-219 shows multiple charge states from +6 to +13. These charge states correspond well with the number of basic residues in the peptides. L1-140, for instance, includes 13 basic residues, 6 K, 5 R, and 2 H, which explains the highest charge state of +14, considering that the N-terminus also can capture one proton. L141-219 contains 12 basic residues, 7 K, 3 R, and 2 H, and the highest charge state is +13. Some of the other abundant signals in the mass spectrum of reduced antibody (see Figure 2.4) result from peptides whose sequences overlap with these two large peptides, such as amino acids L1-51, L1-75, L1-90, L76-140, L91-140, L141-165, and L166-219. Combinations of these peptides also cover the entire light chain sequence.

The mass spectrum of the reduced-antibody digest (Figure 2.4) also shows many peptides from the heavy chain, such as H1-104, H105-113, H114-179, H180-235, H236-273, H274-363, H364-399, and H400-441. H1-104 covers CDR-H1 and CDR-H2 regions, and shows N-terminal pyroglutamic acid formation (-17 Da compared to the mass of the original sequence). The heavy-chain peptide coverage (percentage of amino acids comprised by the detected peptides) for the 3-ms digestion is 100%.

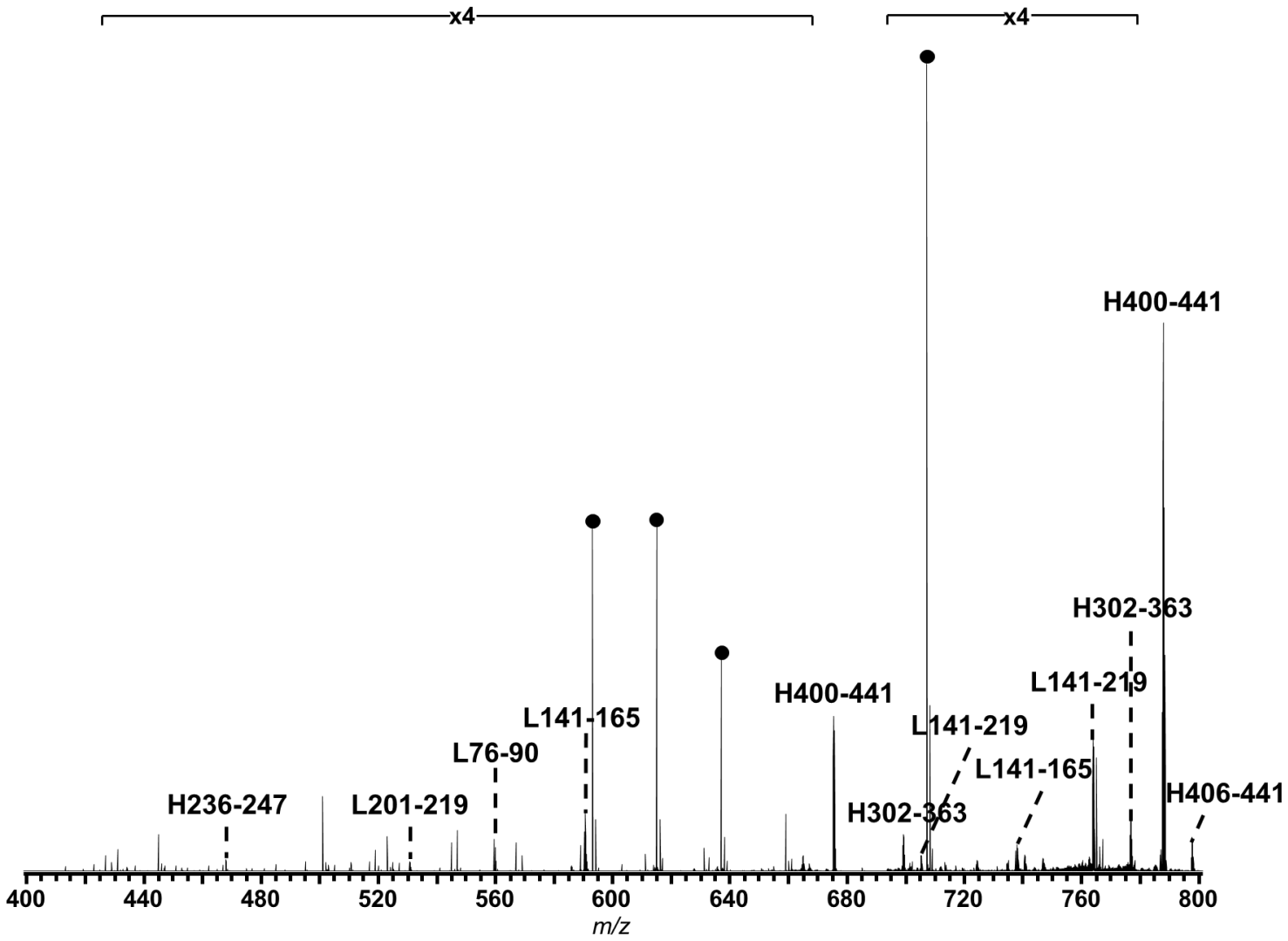


Figure 2.4. Part of the mass spectrum of a 3-ms, in-membrane digest of WIgG1. Labels show the amino acids on the Hc and Lc. Black circles denote unidentified impurities with +1 charge states.

Figure 2.4 (cont'd)

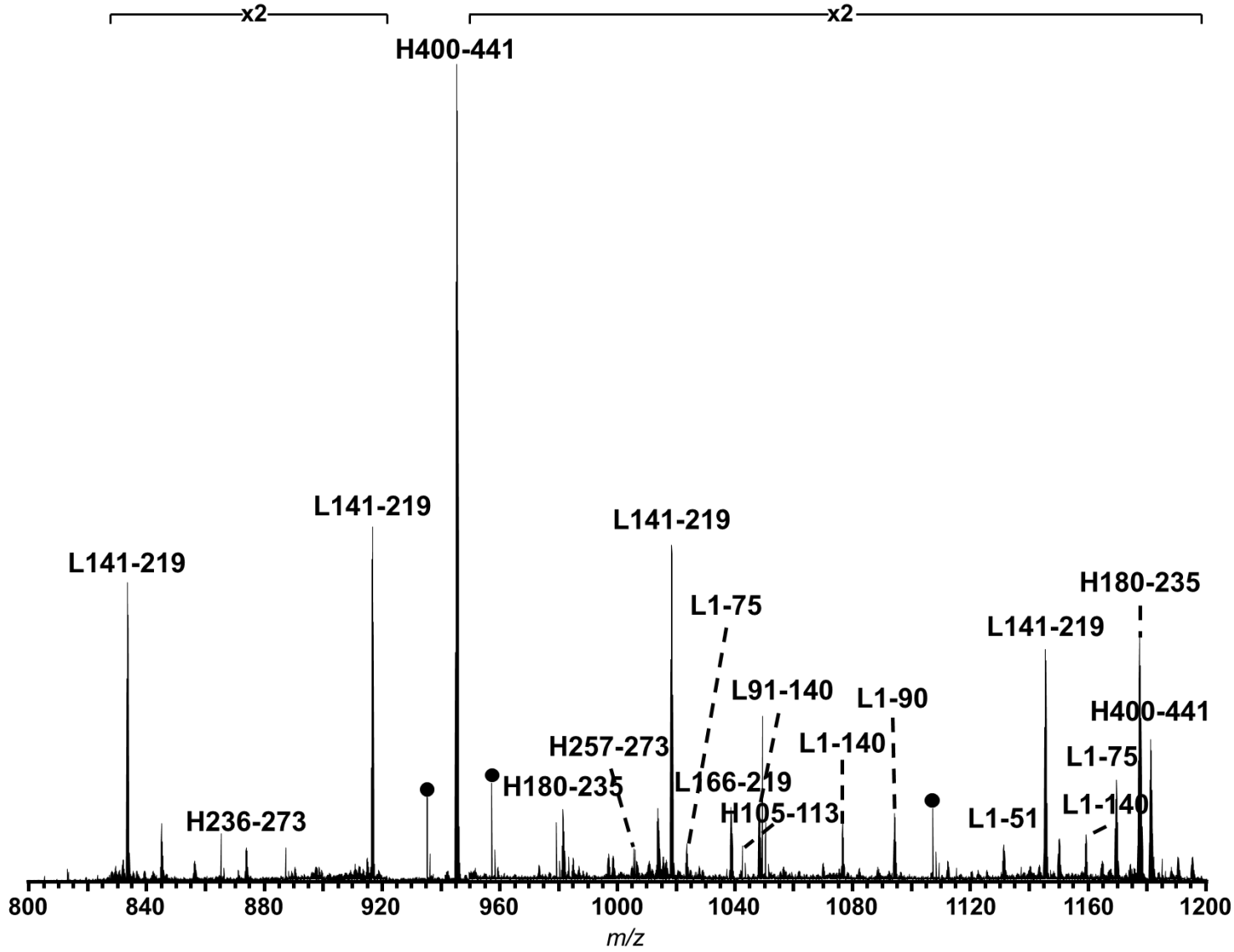


Figure 2.4 (cont'd)

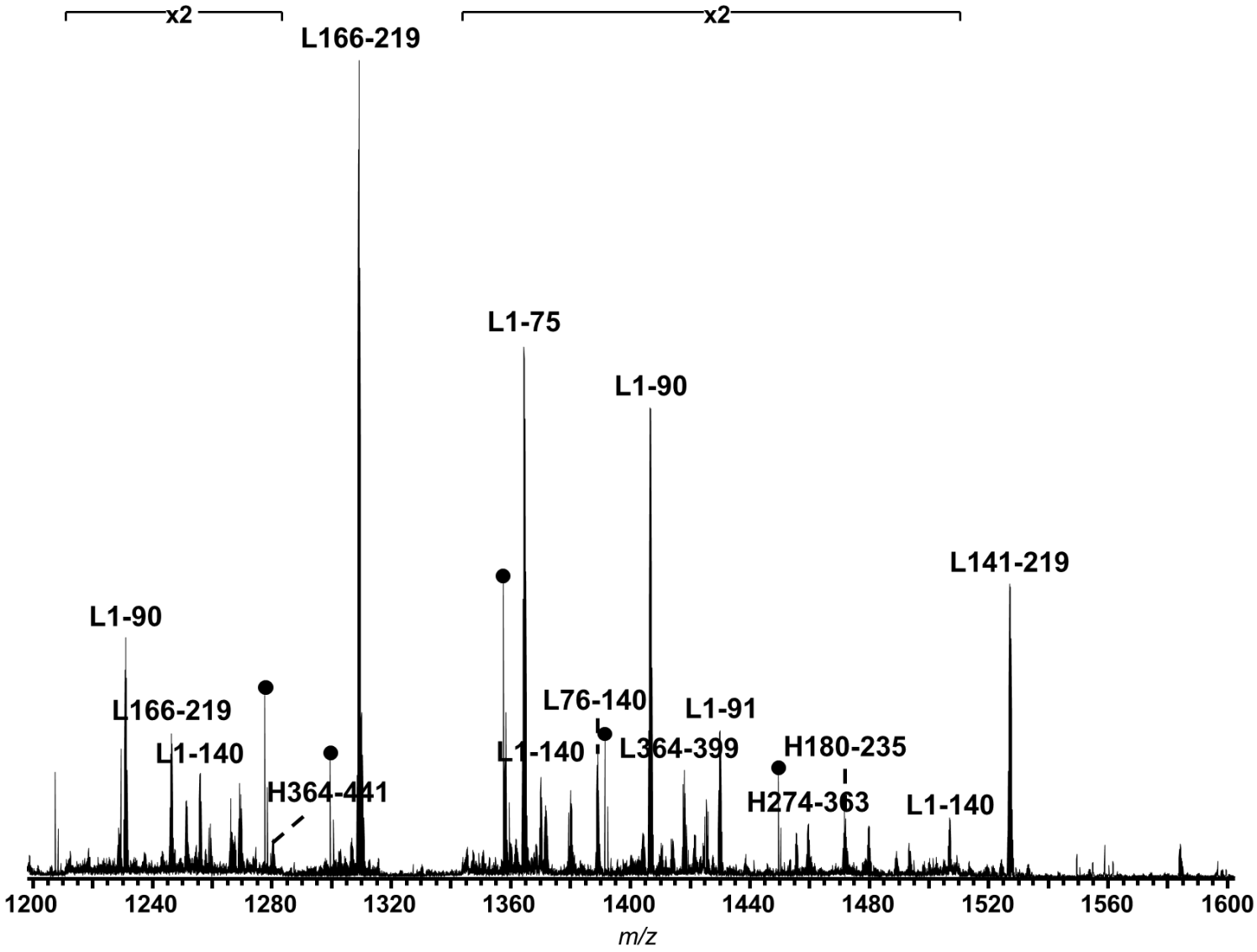
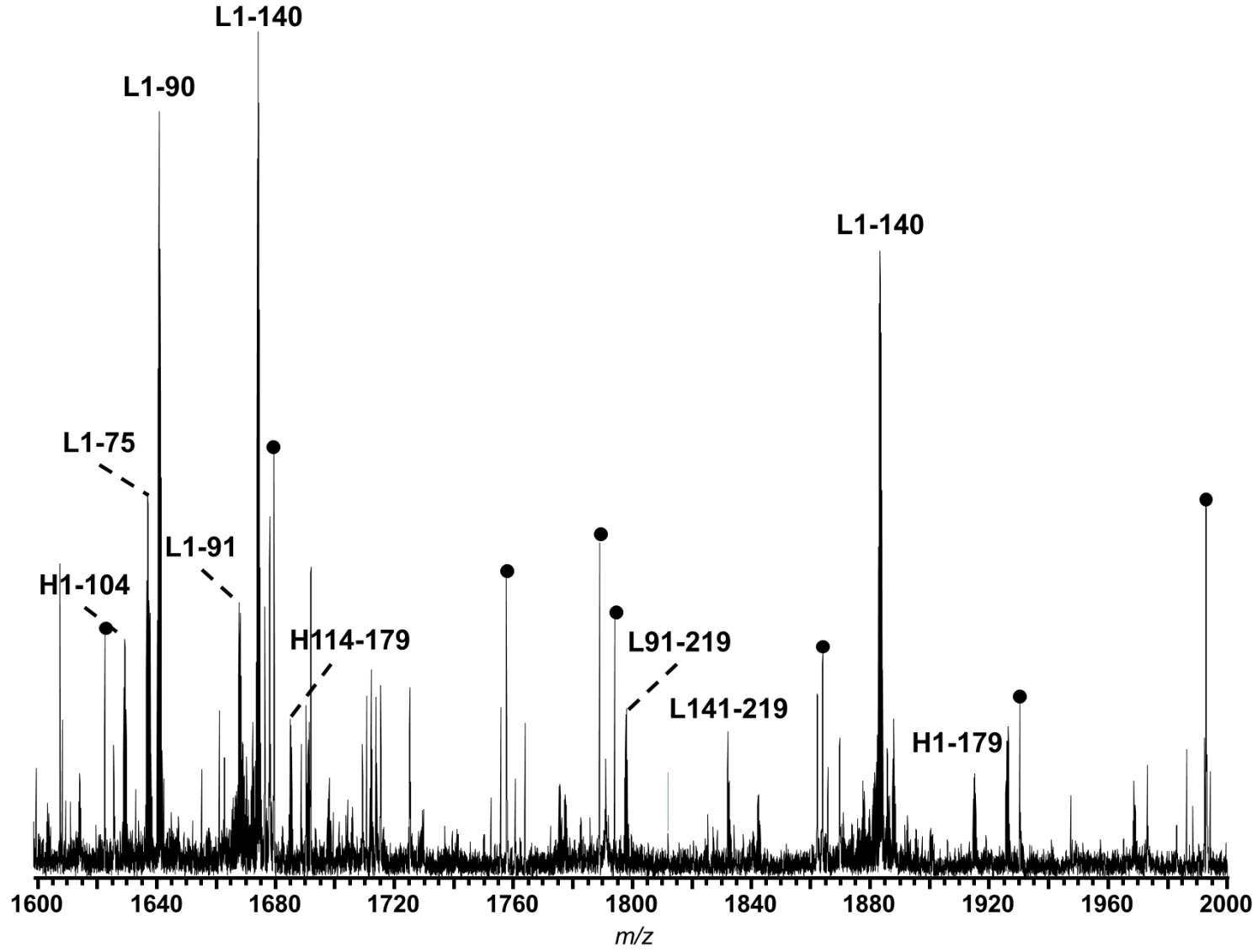


Figure 2.4 (cont'd)



Replicate 3-ms, in-membrane digestions with 3 different membrane pieces show similar signal intensities (Figure 2.5). In contrast to ms digestion, 3-s in-membrane digestion (slower flow through the membrane) shows shorter peptides from the Lc and Hc (see Tables 2.1 and 2.2 for lists of peptides).

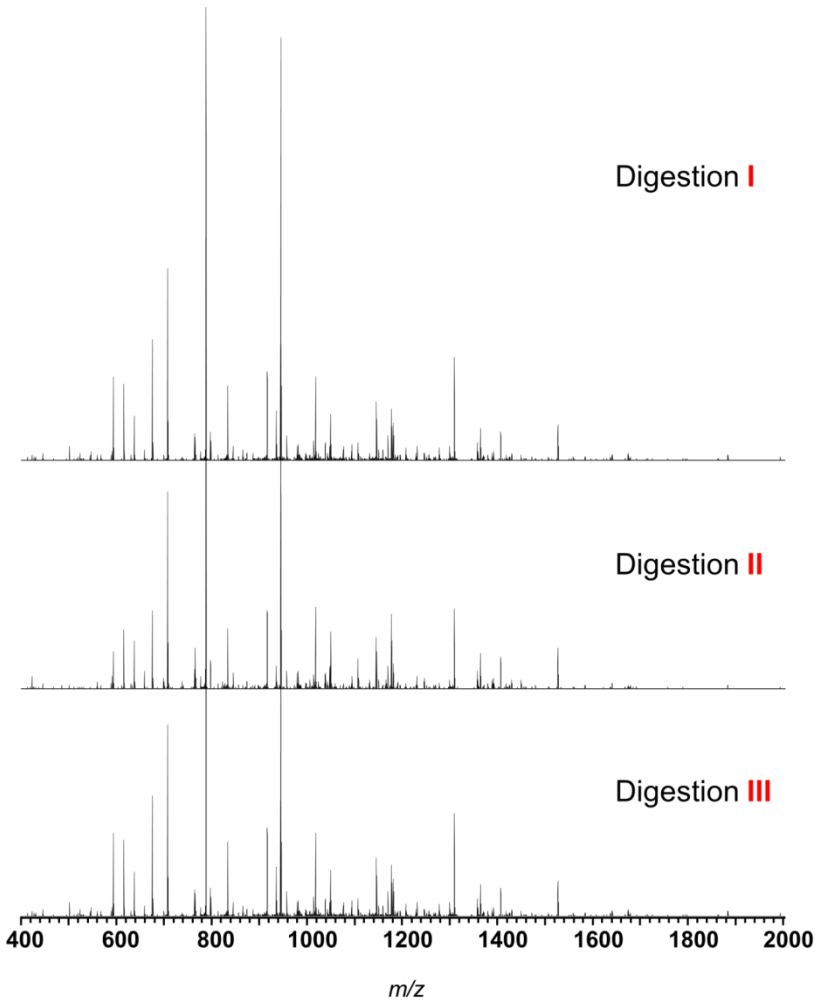


Figure 2.5. Mass spectra of 3 different 3-ms, in-membrane digests of WIgG1. Each digestion employed a different piece of membrane cut from a larger membrane with a diameter of 25 mm. For 18 out of the 20 peptides that give the highest signal intensities, standard deviations of signals (relative to the most intense peak in the spectrum) from the three digests are <3% of the average signal. All of the relative standard deviations are <10%.

Table 2.1. Light- and heavy-chain peptides identified from a 3-ms in-membrane digest of WIgG1.

m/z of $[M+H]^+$	Peptide Sequence	Amino Acids
5649.8334	(-)DVLMTQTPLSLPVSLGDQASISCRSSQYIVHSNGNTYLEWY LQKPGQSPKL(L)	L1-51
8177.0736	(-)DVLMTQTPLSLPVSLGDQASISCRSSQYIVHSNGNTYLEWY LQKPGQSPKLLIYKVSNRFSGVPDRFSGSGSGTD(F)	L1-75
8324.1442	(-)DVLMTQTPLSLPVSLGDQASISCRSSQYIVHSNGNTYLEWY LQKPGQSPKLLIYKVSNRFSGVPDRFSGSGSGTDF(T)	L1-76
9834.9742	(-)DVLMTQTPLSLPVSLGDQASISCRSSQYIVHSNGNTYLEWY LQKPGQSPKLLIYKVSNRFSGVPDRFSGSGSGTDFTLKISRVEA EDLGV(Y)	L1-90
9998.0410	(-)DVLMTQTPLSLPVSLGDQASISCRSSQYIVHSNGNTYLEWY LQKPGQSPKLLIYKVSNRFSGVPDRFSGSGSGTDFTLKISRVEA EDLGVY(Y)	L1-91
15050.5984	(-)DVLMTQTPLSLPVSLGDQASISCRSSQYIVHSNGNTYLEWY LQKPGQSPKLLIYKVSNRFSGVPDRFSGSGSGTDFTLKISRVEA EDLGVYYCFQGSHVPLTFGAGTKLEIKRADAAPTVSIFPPSSEQ LTSGGASVVCF(L)	L1-140
2546.2594	(L)LIYKVSNRFSGVPDRFSGSGSGTD(F)	L52-75
1676.9197	(D)FTLKISRVEAEDLGV(Y)	L76-90
6892.5108	(D)FTLKISRVEAEDLGVYYCFQGSHVPLTFGAGTKLEIKRADA APTVSIFPPSSEQLTSGGASVVCF(L)	L76-140
5234.5978	(V)YYCFQGSHVPLTFGAGTKLEIKRADAAPTVSIFPPSSEQLTS GGASVVCF(L)	L91-140
2947.5393	(F)LNNFYPKDINVKWKIDGSERQNGVL(N)	L141-165
9152.2863	(F)LNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSM SSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNEC(-)	L141-219
6223.7777	(L)NSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTS TSPIVKSFNRNEC(-)	L166-219
2120.0518	(E)ATHKTSTSPIVKSFNRNEC(-)	L201-219
11392.6549	(-)QVQLKESGPGLVAPSQSLSITCTVSGFSLLGYGVNWVRQP PGQGLEWLMGIWGDGSTDYNSALKSRISITKDNSKSQVFLKM NSLQTDDTAKYYCTRAPYGKQY(F)	H1-104
19131.4846	(-)QVQLKESGPGLVAPSQSLSITCTVSGFSLLGYGVNWVRQP PGQGLEWLMGIWGDGSTDYNSALKSRISITKDNSKSQVFLKM NSLQTDDTAKYYCTRAPYGKQYFAYWGQGTLVTVSAAKTTPP SVYPLAPGSAAQTDSMVTLGCLVKGYFPEPVTVTWNSGSLSSG VHTFPAVLQSDL(Y)	H1-179
2958.4110	(F)LKMNSLQTDDTAKYYCTRAPYGKQY(F)	H80-104
1042.4996	(Y)FAYWGQGTL(V)	H105-113
6734.4204	(L)VTVSAAKTTPPSVYPLAPGSAAQTDSMVTLGCLVKGYFPEP VTVTWNSGSLSSGVHTFPAVLQSDL(Y)	H114-179
4275.1546	(L)YTLSSSVTVPSSTWPSETVTCNVAHPASSTKVDKKIVPRD(C)	H180-219
5880.8620	(L)YTLSSSVTVPSSTWPSETVTCNVAHPASSTKVDKKIVPRDCG CKPCICTVPEVSSV(F)	H180-235
1624.7200	(D)CGCKPCICTVPEVSSV(F)	H220-235

Table 2.1 (cont'd)

1401.8122	(V)FIFPPKPKDVLT(I)	H236-247
2358.3482	(V)FIFPPKPKDVLTITLTPKVTC(V)	H236-256
4349.3218	(V)FIFPPKPKDVLTITLTPKVTCVVVDISKDDPEVQFSWF(V)	H236-273
2009.9866	(C)VVVDISKDDPEVQFSWF(V)	H257-273
11662.6778*	(F)VDDVEVHTAHTQPREEQFNSTFRSVSELPIMHQDWLNGKEF KCRVNSAAFPAPIEKTISKTKGRPKAPQVYTIPPPKEQMAKDKV SLTCM(I)	H274-363
11824.7464**	(F)VDDVEVHTAHTQPREEQFNSTFRSVSELPIMHQDWLNGKEF KCRVNSAAFPAPIEKTISKTKGRPKAPQVYTIPPPKEQMAKDKV SLTCM(I)	H274-363
6979.6366	(L)PIMHQDWLNGKEFKCRVNSAAFPAPIEKTISKTKGRPKAPQ VYTIPPPKEQMAKDKVSLTCM(I)	H302-363
4249.8985	(M)ITDFFPEDITVEWQWNGQPAENYKNTQPIMDTDGSY(F)	H364-399
4720.2686	(Y)FVYSKLNVQKSNWEAGNTFTCSVLHEGLHNHHTEKSLSHS PG(-)	H400-441
3982.8622	(L)NVQKSNWEAGNTFTCSVLHEGLHNHHTEKSLSHSPG(-)	H406-441
2507.1853	(F)TCSVLHEGLHNHHTEKSLSHSPG(-)	H419-441

* 11662.6778 is the monoisotopic mass for H274-363 with G0F glycosylation.
** 11824.7464 is the monoisotopic mass for H274-363 with G1F glycosylation.

Table 2.2. Light- and heavy-chain peptides identified from a 3-s in-membrane digest of WIgG1.

m/z of $[M+H]^+$	Peptide Sequence	Amino Acids
1217.6434	(-)DVLMTQTPLSL(P)	L1-11
5649.8334	(-)DVLMTQTPLSLPVSLGDQASISCRSSQYIVHSNGNTYLEW YLQKPGQSPKL(L)	L1-51
8177.0736	(-)DVLMTQTPLSLPVSLGDQASISCRSSQYIVHSNGNTYLEW YLQKPGQSPKLLIYKVSNRFSGVPDRFSGSGSGTD(F)	L1-75
8324.1442	(-)DVLMTQTPLSLPVSLGDQASISCRSSQYIVHSNGNTYLEW YLQKPGQSPKLLIYKVSNRFSGVPDRFSGSGSGTDF(T)	L1-76
9834.9742	(-)DVLMTQTPLSLPVSLGDQASISCRSSQYIVHSNGNTYLEW YLQKPGQSPKLLIYKVSNRFSGVPDRFSGSGSGTDFTLKISRVE AEDLGV(Y)	L1-90
9998.0410	(-)DVLMTQTPLSLPVSLGDQASISCRSSQYIVHSNGNTYLEW YLQKPGQSPKLLIYKVSNRFSGVPDRFSGSGSGTDFTLKISRVE AEDLGVY(Y)	L1-91
5322.6528	(L)MTQTPLSLPVSLGDQASISCRSSQYIVHSNGNTYLEWYLQK PGQSPKL(L)	L4-51
7849.8865	(L)MTQTPLSLPVSLGDQASISCRSSQYIVHSNGNTYLEWYLQK PGQSPKLLIYKVSNRFSGVPDRFSGSGSGTD(F)	L4-75
7996.9674	(L)MTQTPLSLPVSLGDQASISCRSSQYIVHSNGNTYLEWYLQK PGQSPKLLIYKVSNRFSGVPDRFSGSGSGTDF(T)	L4-76
1634.7802	(L)PVSLGDQASISCRSSQ(Y)	L12-27
4451.2054	(L)PVSLGDQASISCRSSQYIVHSNGNTYLEWYLQKPGQSPKL(L)	L12-51

Table 2.2 (cont'd)

6978.4512	(L)PVSLGDQASISCRSSQYIVHSNGNTYLEWYLQKPGQSPKLLIYKVSNRFSGVPDRFSGSGSGTD(F)	L12-75
2835.4426	(Q)YIVHSNGNTYLEWYLQKPGQSPKL(L)	L28-51
5362.6873	(Q)YIVHSNGNTYLEWYLQKPGQSPKLLIYKVSNRFSGVPDRFSGSGSGTD(F)	L28-75
1258.7134	(W)YLQKPGQSPKL(L)	L41-51
3785.9593	(W)YLQKPGQSPKLLIYKVSNRFSGVPDRFSGSGSGTD(F)	L41-75
2546.2606	(L)LIYKVSNRFSGVPDRFSGSGSGTD(F)	L52-75
2693.3298	(L)LIYKVSNRFSGVPDRFSGSGSGTDF(T)	L52-76
4204.1648	(L)LIYKVSNRFSGVPDRFSGSGSGTDFTLKISRVEAEDLGV(Y)	L52-90
1676.9197	(D)FTLKISRVEAEDLGV(Y)	L76-90
1839.9848	(D)FTLKISRVEAEDLGVY(Y)	L76-91
1529.8510	(F)TLKISRVEAEDLGV(Y)	L77-90
1692.9145	(F)TLKISRVEAEDLGVY(Y)	L77-91
2089.0210	(V)YYCFQGSHVPLTFGAGTKL(E)	L91-109
5234.5978	(V)YYCFQGSHVPLTFGAGTKLEIKRADAAPTVSIFPPSSEQLTSGGASVVCF(L)	L91-140
1925.9566	(Y)YCFQGSHVPLTFGAGTKL(E)	L92-109
5071.5334	(Y)YCFQGSHVPLTFGAGTKLEIKRADAAPTVSIFPPSSEQLTSGGASVVCF(L)	L92-140
3164.5924	(L)EIKRADAAPTVSIFPPSSEQLTSGGASVVCF(L)	L110-140
2947.5393	(F)LNNFYPKDINVKWKIDGSERQNGVL(N)	L141-165
7051.2610	(F)LNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCE(A)	L141-200
9152.2863	(F)LNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNEC(-)	L141-219
4122.7438	(L)NSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCE(A)	L166-200
6223.7705	(L)NSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNEC(-)	L166-219
1988.8645	(L)TLTKDEYERHNSYTCE(A)	L185-200
4089.8950	(L)TLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNEC(-)	L185-219
2120.0518	(E)ATHKTSTSPIVKSFNRNEC(-)	L201-219
3405.6730	(L)MGIWGDGSTDYNSALKSRISITKDNSKSQVF(L)	H49-79
6345.0612	(L)MGIWGDGSTDYNSALKSRISITKDNSKSQVFLKMNSLQTDDTAKYYCTRAPYGKQY(F)	H49-104
3217.6082	(G)IWGDGSTDYNSALKSRISITKDNSKSQVF(L)	H51-79
2223.2054	(Y)NSALKSRISITKDNSKSQVF(L)	H60-79
1838.0098	(L)KSRISITKDNSKSQVF(L)	H64-79
2210.2288	(L)KSRISITKDNSKSQVFLKM(N)	H64-82
4777.4020	(L)KSRISITKDNSKSQVFLKMNSLQTDDTAKYYCTRAPYGKQY(F)	H64-104
1353.6991	(I)SITKDNSKSQVF(L)	H68-79
1725.9178	(I)SITKDNSKSQVFLKM(N)	H68-82
4293.0918	(I)SITKDNSKSQVFLKMNSLQTDDTAKYYCTRAPYGKQY(F)	H68-104
1627.7971	(F)LKMNSLQTDDTAKY(Y)	H80-93
2958.4106	(F)LKMNSLQTDDTAKYYCTRAPYGKQY(F)	H80-104

Table 2.2 (cont'd)

2586.1906	(M)NSLQTDDTAKYYCTRAPYGKQY(F)	H83-104
1349.6284	(Y)YCTRAPYGKQY(F)	H94-104
1042.4996	(Y)FAYWGQGTL(V)	H105-113
2646.3451	(L)VTVSAAKTTPPSVYPLAPGSAAQTNSM(V)	H114-140
3233.6629	(L)VTVSAAKTTPPSVYPLAPGSAAQTNSMVTLGCL(V)	H114-146
6290.2188	(L)VTVSAAKTTPPSVYPLAPGSAAQTNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVL(Q)	H114-175
3662.8918	(M)VTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVL(Q)	H141-175
3076.5775	(L)VKGYFPEPVTVTWNSGSLSSGVHTFPAVL(Q)	H147-175
3519.7810	(L)VKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSDL(Y)	H147-179
4275.1566	(L)YTLSSSVTVPSSTWPSETVTCNVAHPASSTKVDKKIVPRD(C)	H180-219
5880.8620	(L)YTLSSSVTVPSSTWPSETVTCNVAHPASSTKVDKKIVPRDCGCKPCICTVPEVSSV(F)	H180-235
1624.7196	(D)CGCKPCICTVPEVSSV(F)	H220-235
1300.7644	(V)FIFPPKPKDVL(T)	H236-246
1401.8122	(V)FIFPPKPKDVLT(I)	H236-247
2358.3478	(V)FIFPPKPKDVLTITLTPKVTC(V)	H236-256
1141.6604	(I)FPPKPKDVLT(I)	H238-247
1076.6008	(L)TITLTPKVTC(V)	H247-256
975.5532	(T)ITLTPKVTC(V)	H248-256
989.5138	(C)VVVDISKDD(P)	H257-265
1442.7366	(C)VVVDISKDDPEVQ(F)	H257-269
2009.986	(C)VVVDISKDDPEVQFSWF(V)	H257-273
4702.0793*	(F)VDDVEVHTAHTQPREEQFNSTFRSVSEL(P)	H274-301
4864.1330**	(F)VDDVEVHTAHTQPREEQFNSTFRSVSEL(P)	H274-301
5026.1942***	(F)VDDVEVHTAHTQPREEQFNSTFRSVSEL(P)	H274-301
1467.7027	(L)PIMHQDWLNGKE(F)	H302-313
4718.5089	(L)PIMHQDWLNGKEFKCRVNSAAFPAPIEKTISKTKGRPKAPQV(Y)	H302-343
6979.6366	(L)PIMHQDWLNGKEFKCRVNSAAFPAPIEKTISKTKGRPKAPQVYTIPPPKEQMAKDKVSLTCM(I)	H302-363
5530.9434	(E)FKCRVNSAAFPAPIEKTISKTKGRPKAPQVYTIPPPKEQMAKDKVSLTCM(I)	H314-363
2280.1402	(V)YTIPPPKEQMAKDKVSLTCM(I)	H344-363
1425.6776	(M)ITDFFPEDITVE(W)	H364-375
2843.2366	(E)WQWNGQPAENYKNTQPIMDTDGSY(F)	H376-399
756.4278	(Y)FVYSKL(N)	H400-405
4720.2686	(Y)FVYSKLNVQKSNWEAGNTFTCSVLHEGLHNHHTEKSLSHSPG(-)	H400-441
3982.8598	(L)NVQKSNWEAGNTFTCSVLHEGLHNHHTEKSLSHSPG(-)	H406-441
2507.1853	(F)TCSVLHEGLHNHHTEKSLSHSPG(-)	H419-441
2406.1318	(T)CSVLHEGLHNHHTEKSLSHSPG(-)	H420-441
2303.1238	(C)SVLHEGLHNHHTEKSLSHSPG(-)	H421-441
2216.0888	(S)VLHEGLHNHHTEKSLSHSPG(-)	H422-441
2003.9402	(L)HEGLHNHHTEKSLSHSPG(-)	H424-441

* 4702.0793 show monoisotopic mass of H274-301 with G0F glycosylation.
** 4864.1330 show monoisotopic mass of H274-301 with G1F glycosylation.
*** 5026.1942 show monoisotopic mass of H274-301 with G2F glycosylation.

As a second example, we digested Trastuzumab (Herceptin), a commercial humanized mAb for breast cancer treatment. Detailed characterization of Trastuzumab is crucial for quality control, and several groups recently presented MS-based interrogation of this antibody.[3,37,56-60] Remarkably, a 3-ms digestion of reduced Trastuzumab followed by MS analysis yields 100% peptide coverage for both the light and heavy chains. (See Figure 2.6 for the mass spectrum.) A large peptide, L1-83, covers the CDR-L1 and CDR-L2 regions. Moreover, a peptide containing amino acids 1-115 on the heavy chain contains all the CDR-H regions. These data demonstrate that rapid digestion (less than 1-min total digestion time) with a pepsin-containing membrane can yield essentially complete peptide coverage. Moreover, the simple acidic reduction procedure requires only 15 min, and the MS data collection takes 5 min when using direct infusion.

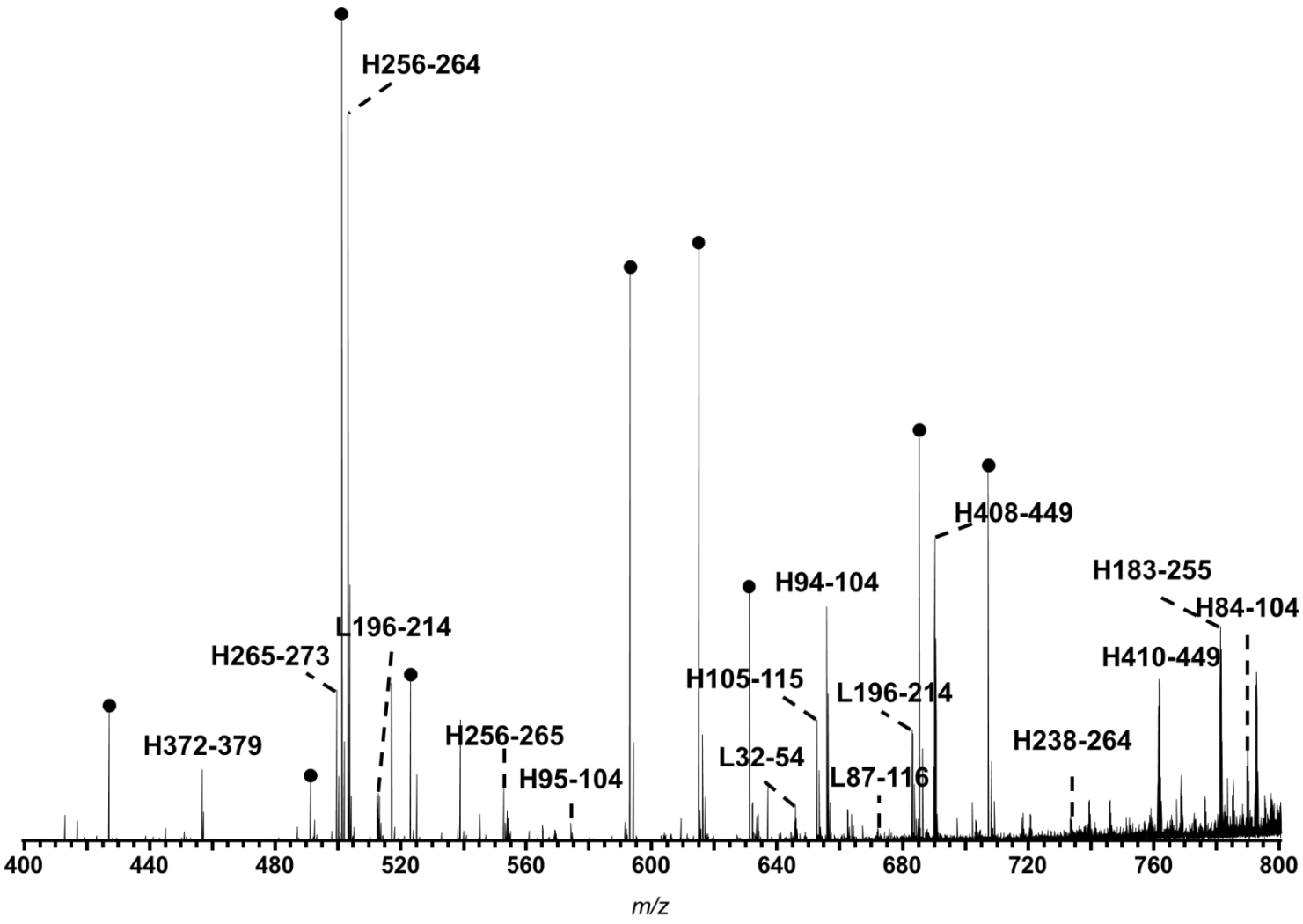


Figure 2.6. Part of the mass spectrum of a 3-ms, in-membrane digest of Trastuzumab. Labels show the amino acids on the Hc and Lc. Black circles denote unidentified impurities with +1 charge states.

Figure 2.6 (cont'd)

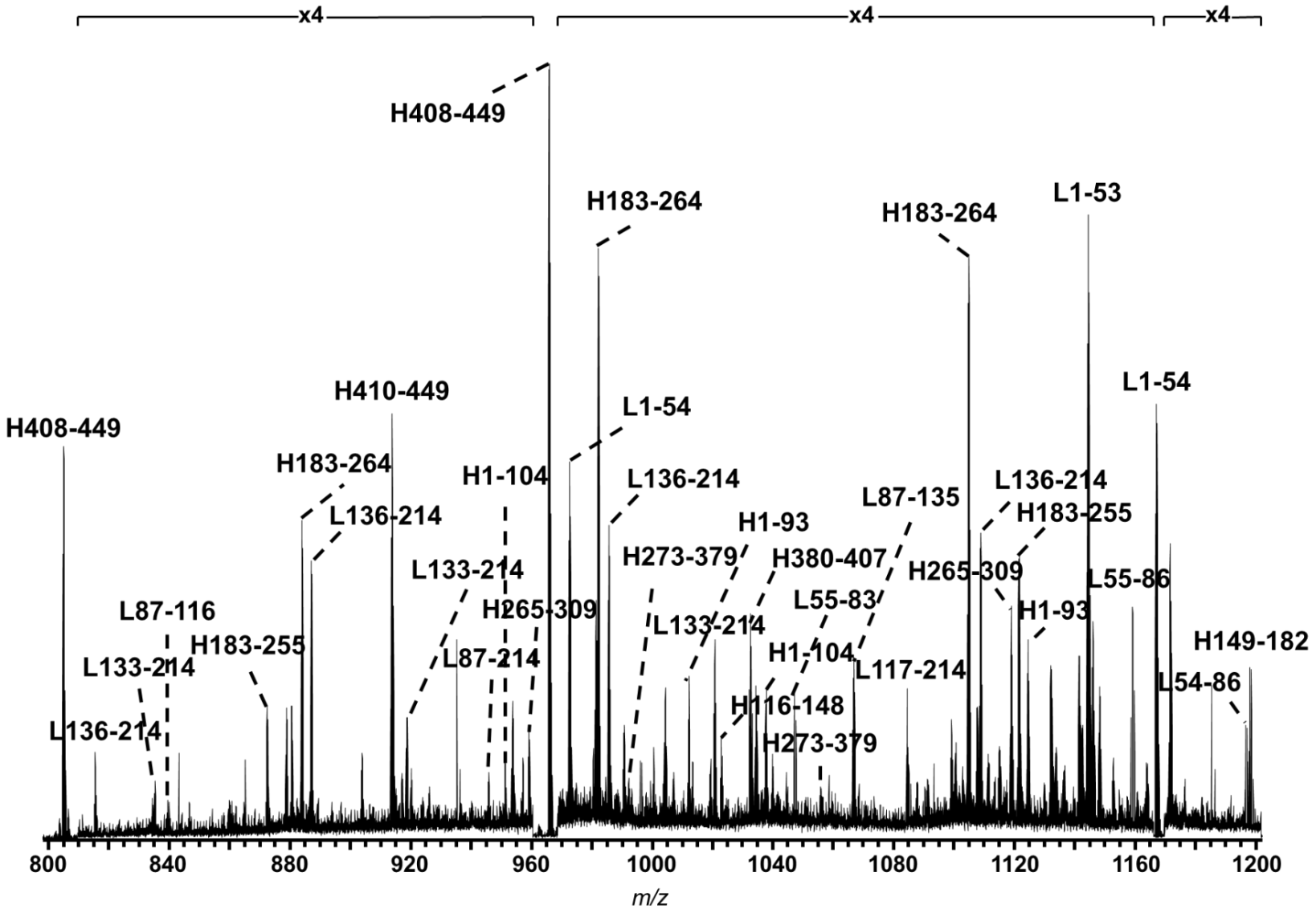


Figure 2.6 (cont'd)

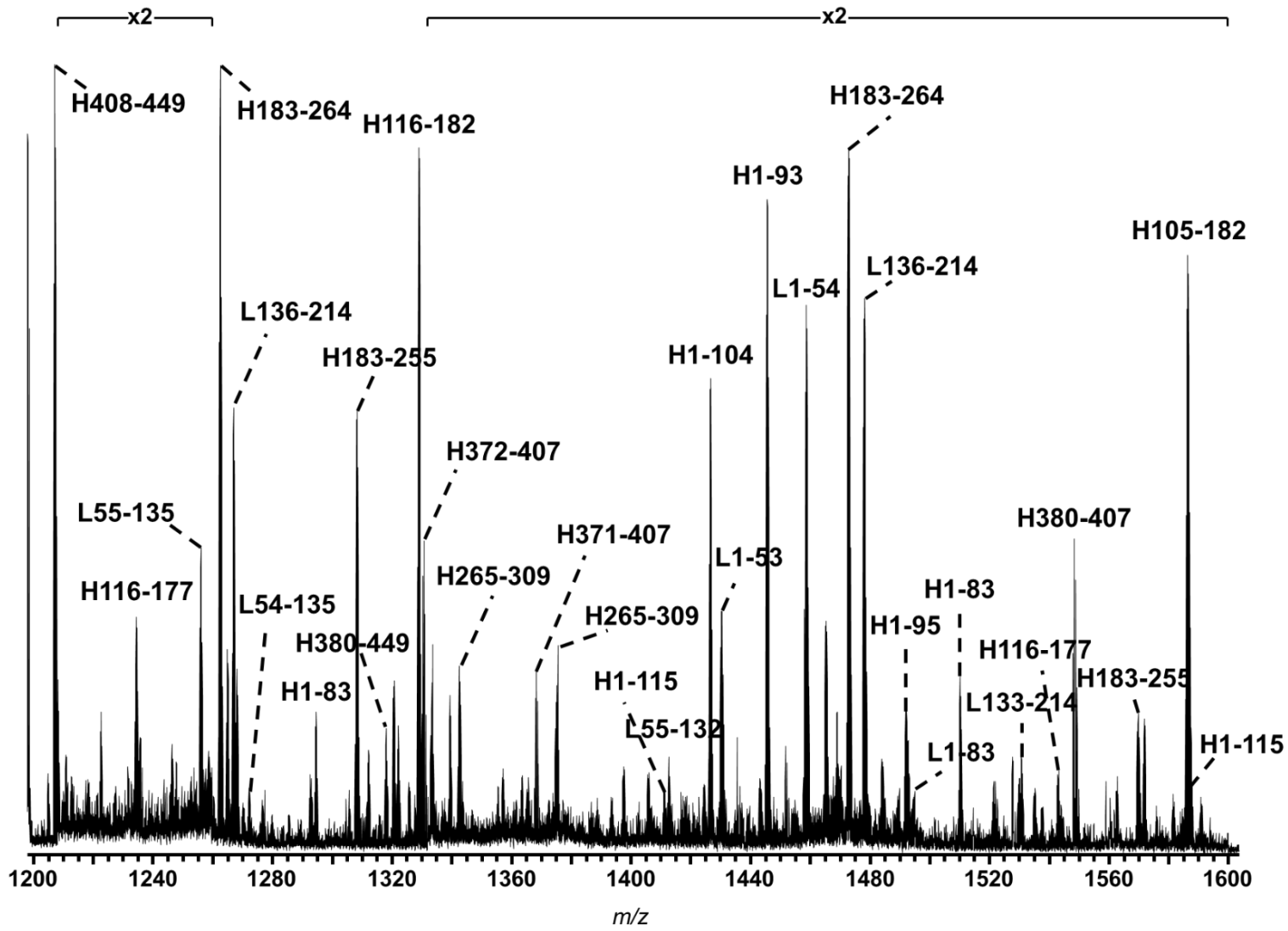
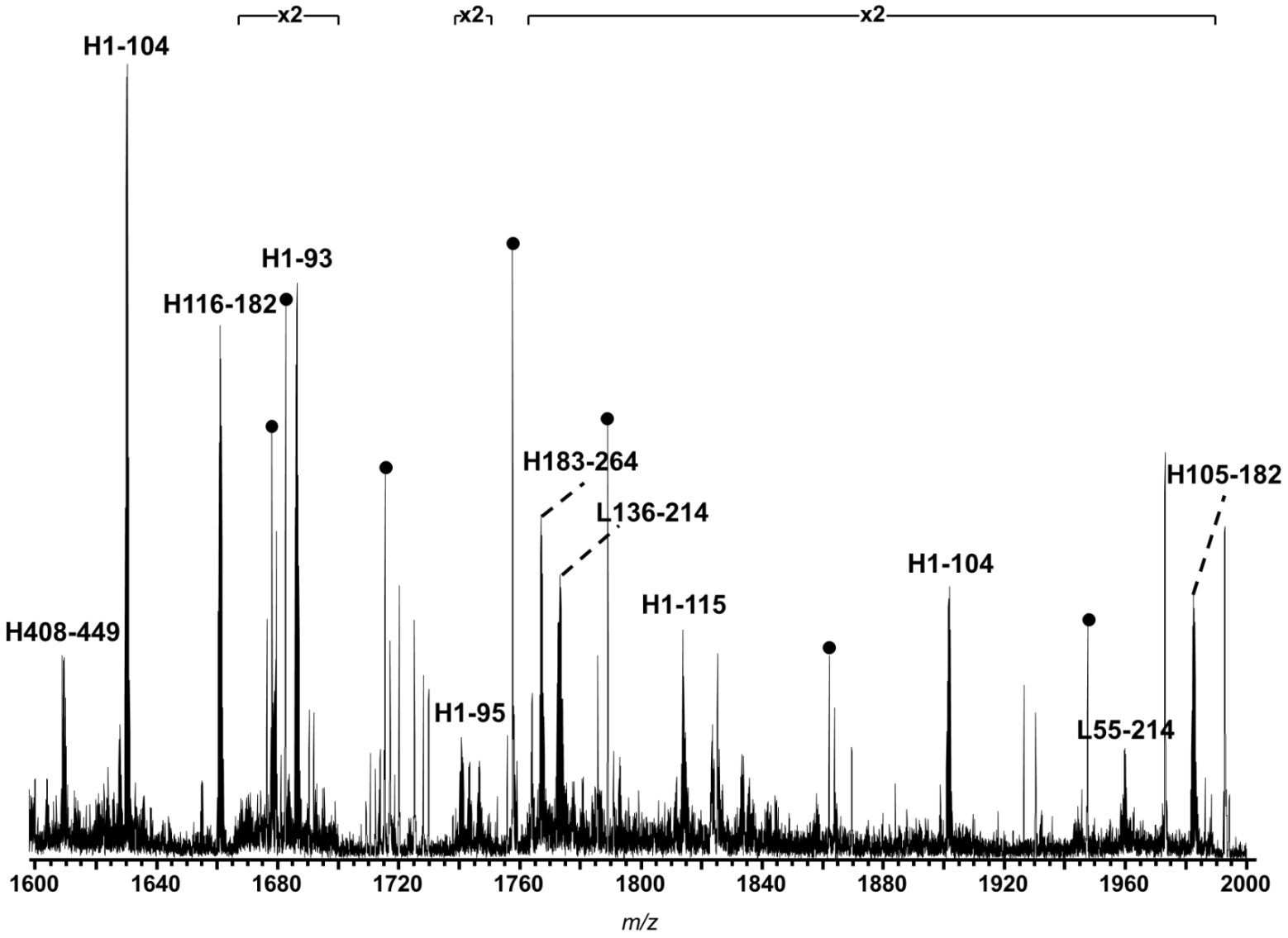


Figure 2.6 (cont'd)



2.3.4 mAb Light-Chain Analysis using In-membrane or Tryptic In-solution Digestion

To further investigate antibody sequence information and compare "middle-down" and "bottom-up" methods, we analyzed the WIgG1 light chain after its isolation using HPLC. SDS-PAGE showed the separation of light and heavy chains, and HPLC retention times guided offline collection of the desired fractions. We reconstituted the antibody light chain in a low-concentration TCEP solution to prevent reformation of disulfide bonds.

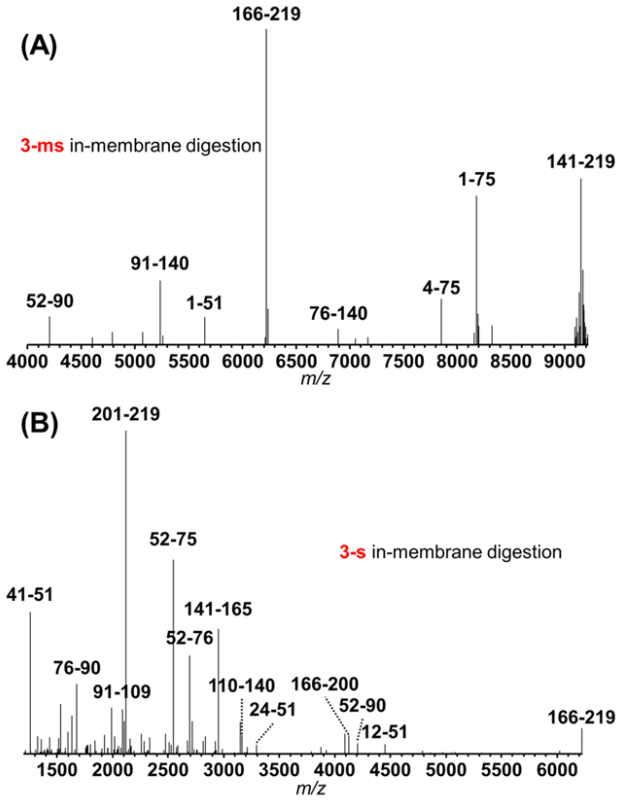


Figure 2.7. Deconvoluted ESI-Orbitrap mass spectra of the WIgG1 light chain digested with 3-ms (A) and 3-s (B) residence times in a pepsin-containing membrane. Other small peptides appear at lower m/z values. Deconvoluted peaks (generated by Xtract) represent m/z values of peptides with a +1 charge.

Figure 2.7 shows the deconvoluted mass spectra of the WIgG1 light chain digested with 3-ms and 3-s residence times in the pepsin-containing membrane. Upon increasing the digestion residence time, large peptides undergo additional cleavages to give smaller peptides. Dominant signals from the 3-ms digestion stem from large peptides, such as L52-90, L91-140, L1-51, L166-219, L76-140, L4-75, L1-75 and L141-219. In contrast, 3-s digestion yields smaller peptides such as L41-51, L52-75, and L76-90, and in several cases combinations of the small peptides give the larger peptides from the 3-ms digestion. The 3-ms digestion enables 100% peptide coverage of the light chain, whereas the 3-s digestion shows only 95% coverage due to the absence of amino acids 1-11. The lack of basic amino acids in the first 11 residues (see Figure 2.9 below for the light-chain sequence) may explain the absence of peptide coverage in this region.

We also performed an in-solution, tryptic digestion of the alkylated light chain for comparison. Eighteen tryptic peptides cover the entire Lc sequence. These peptides show an average length of 12 amino acids, which agrees with the theoretical tryptic peptide length, 8-25 residues. Table 2.3 gives details of the tryptic peptide m/z values.

2.3.5 Comparison of Light-chain Sequence Coverage Using "Middle-down", "Bottom-up" and "Top-down" methods

Specific amino acid sequence information requires MS/MS data with extensive fragmentation. While larger peptides may contain multiple CDRs and give high peptide coverage, small peptides are easier to fragment in some MS/MS methods and may provide higher sequence coverage. CID is widely used in "bottom up" LC-MS/MS proteomics and gives the peptide sequence information via a series of b- and y- ions. One of the peptides, L52-75, from a 30-ms,

in-membrane digestion of the WIgG1 light chain covers the entire CDR-L2 region. CID-MS/MS of this peptide shows 100% sequence coverage, i.e. cleavage of all amide bonds (Figure 2.8).

Table 2.3. WIgG1 peptides identified from an in-solution tryptic digest of the alkylated light chain.

m/z of $[M+H]^+$	Peptide Sequence	Amino Acids
2588.2982	(-)DVLMTQTPLSLPVSLGDQASISC(Carbamidomethyl)R(S)	L1-24
3024.4741	(R)SSQYIVHSNGNTYLEWYLQKPGQSPK(L)	L25-50
649.4264	(K)LLIYK(V)	L51-55
475.2607	(K)VSNR(F)	L56-59
777.3866	(R)FSGVPDR(F)	L60-66
1303.6118	(R)FSGSGSGTDFTLK(I)	L67-79
375.2338	(K)ISR(V)	L80-82
2845.3390	(R)VEAEDLGVYYC(Carbamidomethyl)FQGSHVPLTFGAGTK(L)	L83-108
502.3219	(K)LEIK(R)	L109-112
658.4225	(K)LEIKR(A)	L109-113
3727.8322	(K)RADAAPTVSIFPPSSEQLTSGGASVVC(Carbamidomethyl)FLNNFYPK(D)	L113-147
588.3333	(K)DINVK(W)	L148-152
990.4974	(K)WKIDGSER(Q)	L153-160
1591.7294	(R)QNGVLNSWTDQDSK(D)	L161-174
1534.7250	(K)DSTYSMSSTLTLTK(D)	L175-188
711.2921	(K)DEYER(H)	L189-193
1347.5758	(R)HNSYTC(Carbamidomethyl)EATHK(T)	L194-204
832.4749	(K)TSTSPIVK(S)	L205-212
926.3756	(K)SFNRNEC(Carbamidomethyl)(-)	L213-219

However, not all peptides show such complete fragmentation. One limitation of quadrupole ion-trap CID is the loss of fragment ions in the low m/z range due to the low-mass cutoff determined by the radio frequency amplitude. In contrast, HCD in the LTQ Orbitrap VelosTM mass spectrometer does not have this limitation and facilitates identification of N and C terminal fragment ions.[61] MS/MS analyses of 12 proteolytic peptides L1-51, L4-51, L12-51, L24-51, L52-75, L76-90, L91-136, L110-136, L91-140, L141-165, L166-200, and L201-219 were conducted using both CID and HCD. Combining CID and HCD gives higher sequence coverage than CID

alone. For instance, CID of L141-165 does not break the Pro-Lys bond, while HCD cleaves this bond and gives 100% sequence coverage of this peptide

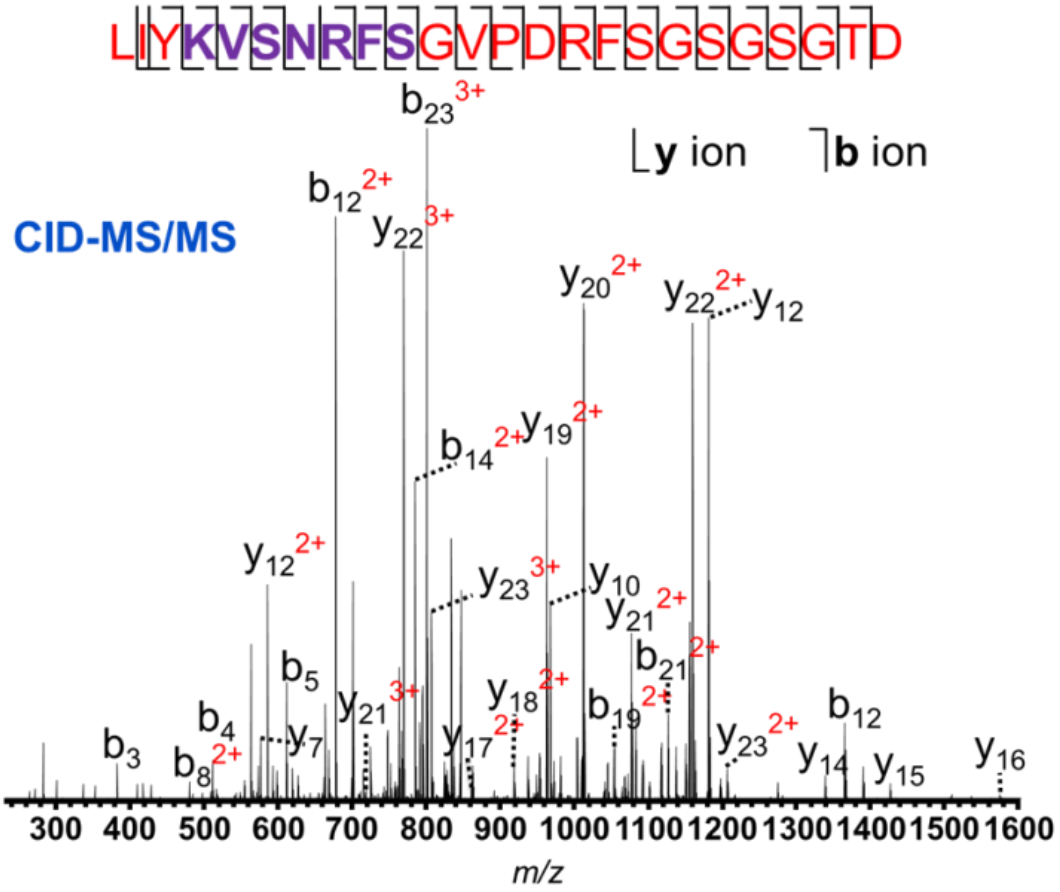


Figure 2.8. CID-MS/MS spectrum of the WIgG1 light-chain peptide L52-75, which covers the entire CDR-L2 region. The sequence at the top of the figure denotes the formation of b and y ions (only some of the b and y ions are labeled in the spectrum).

CID and HCD of L52-75 and L76-90 give 100% sequence coverage of L52-90. Relatively large peptides produced by 3-ms light-chain digestion, L1-51 and L91-140, were fragmented by ETD. In contrast to CID, which cleaves the labile bonds on the peptide chain, ETD induces fragmentation in a sequence-independent manner. A low residue/charge ratio results in effective

ETD fragmentation, so we chose the highest charge states of each peptide for ETD analysis, (+5 for L1-51 and for L91-140). The c- and z- ions produced by ETD (of L1-51 and L91-140) along with the b- and y- ions from CID and HCD of the 12 peptides mentioned previously give 99% bond cleavages in the light chain. Only three amide bond linkages did not dissociate, 25S-26S, 100P-101L, and 113R-114A. Figure 2.9 gives a summary of the cleavage sites.

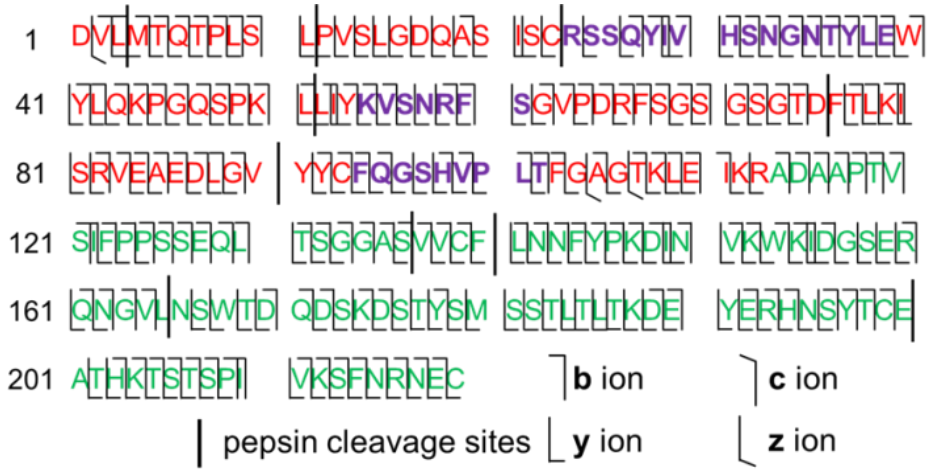


Figure 2.9 Summary of bond cleavage sites from CID, HCD and ETD-MS/MS of WIgG1 light-chain peptides obtained from a 30-ms digestion in pepsin-containing membranes. Red and purple letters cover the light-chain variable region (VL), with purple letters denoting complementarity determining regions (CDRs). Green letters represent the light-chain constant region (CL). The figure does not show redundant cleavages sites from c and z ions.

For comparison, we fragmented the 18 alkylated light-chain tryptic peptides (mentioned in the last section), using CID. The MS/MS spectra reveal cleavage of 205 out of 218 bonds via either enzyme cleavage or CID-MS/MS, which gives 94% sequence coverage. We also conducted HCD on these peptides to obtain nine more fragmentation sites. The combination of CID, HCD, and tryptic cleavage sites yields 99% sequence coverage (Figure 2.10). Notably, with in-solution

tryptic digestion, the CDR-L2 region spans three peptides, 51-55, 56-59, and 60-66 making direct characterization of this region difficult.

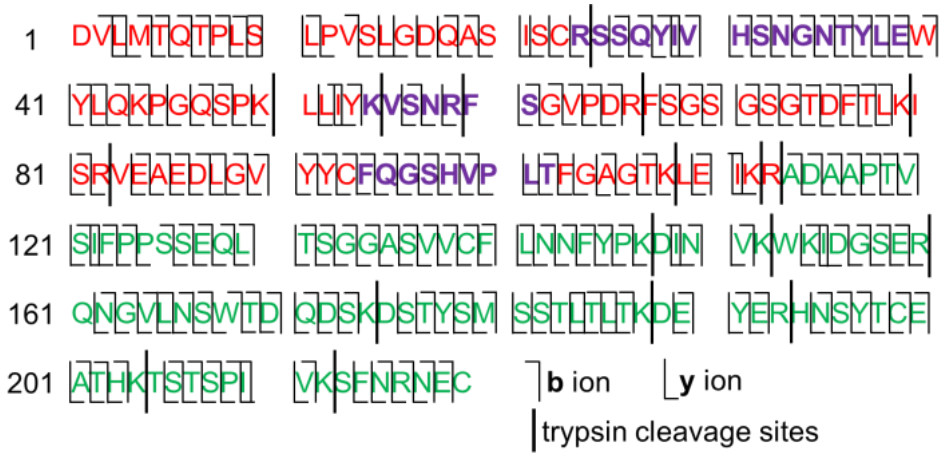


Figure 2.10. Summary of the bond cleavage sites from CID and HCD-MS/MS of peptides obtained from tryptic, in-solution diges-tion of the WIgG1 Lc.

We also performed ETD and CID for the entire reduced WIgG1 light chain. Figure 2.11 shows Orbitrap MS/MS spectra from online LC-MS analysis with 5-ms ETD, 15-ms ETD, and CID with default energy, respectively. Figure 2.12 gives an example of "top-down" ETD MS/MS showing partial sequence coverage. Direct ETD (combining the results from 5-ms and 15-ms ETD) of the light chain generated 53 and 50 detectable c- and z-type ions, respectively; and CID yielded 9 and 25 detectable b- and y-type ions, respectively. These c-, z-, b-, and y-type fragment ions collectively give 120 unique gas-phase backbone cleavages (Figure 2.13), which corresponds to 55% sequence coverage of the WIgG1 light chain.

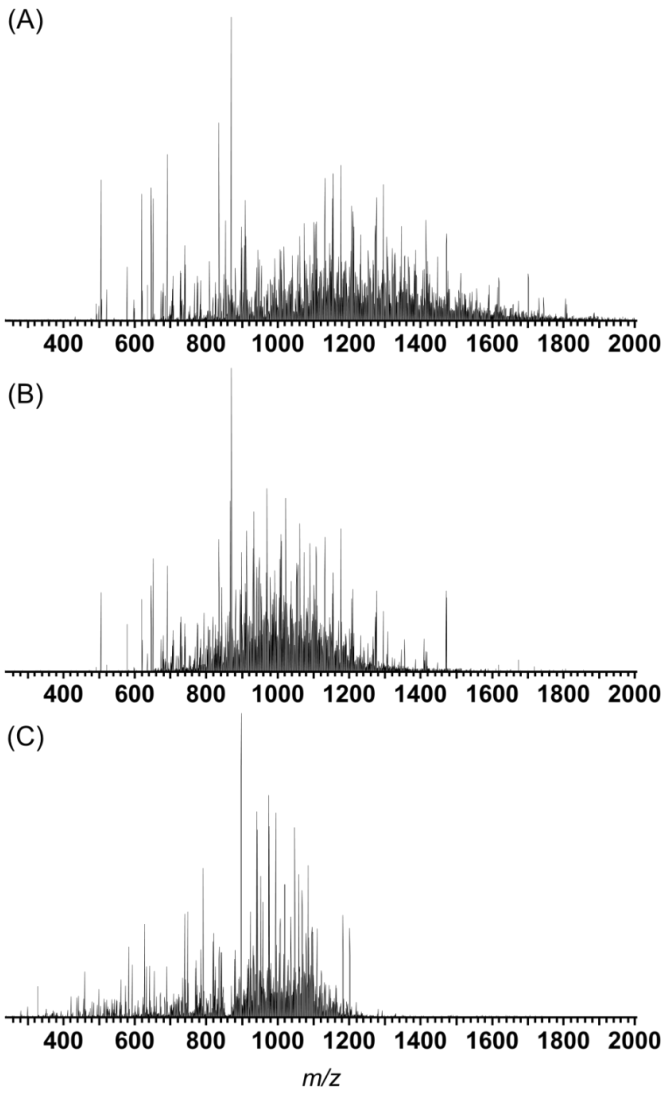


Figure 2.11 "Top-Down" HPLC MS/MS spectra of the WIgG1 light chain. The MS/MS parameters were 15-ms ETD (A, 13 scans merged), 5-ms ETD (B, 13 scans merged), and CID (C, 9 scans merged), respectively. The precursor ion isolation window was 865.2 ±1.5 m/z. The FTMS MSn automatic gain control (AGC) target was set at 3 x 10^5 (A and B) or 1 x 10^5 (C), and the ETD reagent AGC target was set at 3 x 10^5. Each MS/MS scan included 5 microscans.

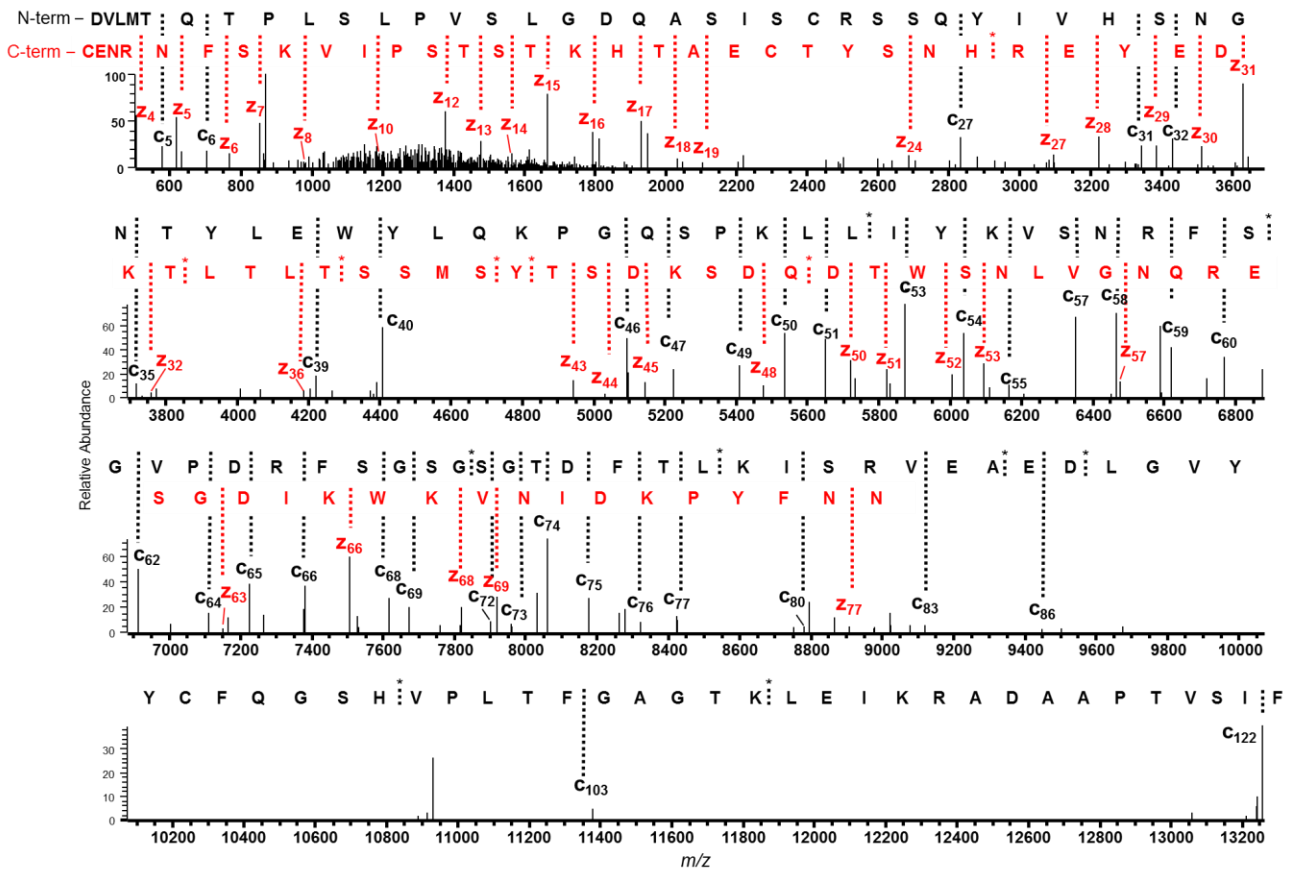


Figure 2.12. Orbitrap FT MS/MS spectrum after Xtract deconvolution of the original MS/MS spectrum (Figure 2.11A) resulting from 15-ms ETD of the WIgG1 Lc. Fragment c- and z- ions are labeled with black and red colors, respectively. The N-terminal (black) and C-terminal (red) sequences of the WIgG1 light chain are placed above the corresponding c- and z- ions, with the fragmentation sites indicated by dashed lines. The dashed lines labeled with * stem from c- or z-ions found from the original MS/MS spectrum, these ions are not displayed in this figure due to incomplete Xtract deconvolution.

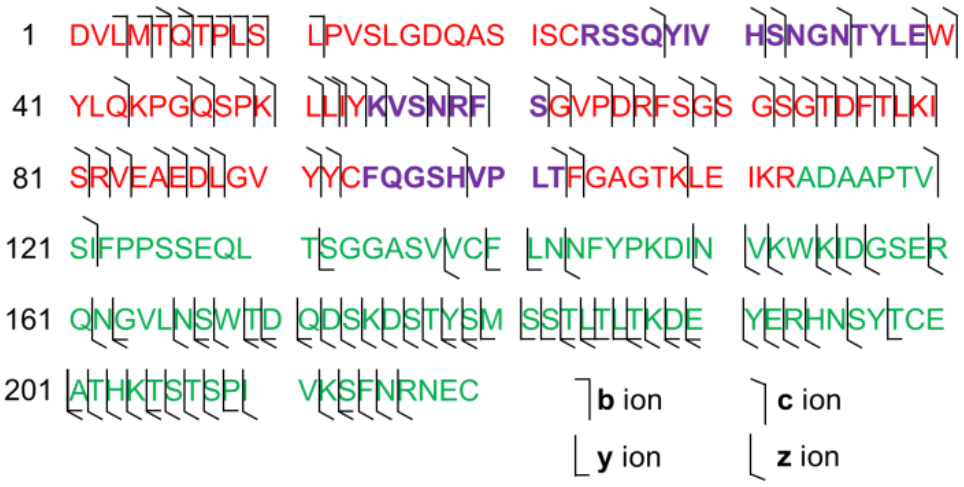


Figure 2.13. Cleavage sites in "top-down" analysis of the antibody light chain.

The sequence coverage is 55% after combining CID, 5-ms ETD and 15-ms ETD.

2.3.6 Detecting PTMs on the light and heavy chains

PTMs affect antibody activities in a variety of ways and are commonly introduced to the sequence during manufacture, purification and storage. Hu et al. showed that oxidation on the light chain may induce a structural change and destabilize the protein.[62] We compared two batches of WIgG1 by conducting 30-ms digestions of their light chains and performing MS analysis. Figure 2.14 compares the signals for peptide L166-219 (+5 charge state) for the two batches of antibody. For the second batch of antibody, the spectrum shows two strong isotopic envelopes whose deconvoluted mass difference is 15.9980, suggesting a Met oxidation. Because the spectra for the two batches of antibody were obtained under the same analysis conditions, this result suggests that oxidation does not occur significantly during ESI. Further MS/MS analysis of these two peptides from the second batch of antibody confirms that oxidation occurs at M 180. (See Figure 2.15 for MS/MS spectra). A similar strategy revealed M 393 oxidation (Figure 2.16) and N 138 deamidation on the heavy chain (Figure 2.17).

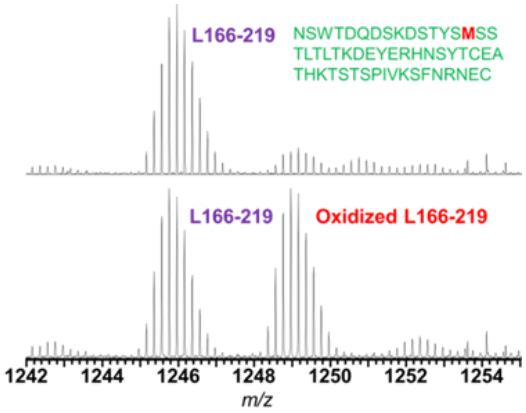


Figure 2.14. Part of the ESI-Orbitrap mass spectra of reduced WIgG1 light chains after digestion for 30-ms in pepsin-containing membranes. The top and bottom spectra come from two batches of antibody, and the largest isotopic envelopes represent the +5 charge state of the peptide containing amino acids L166-219.

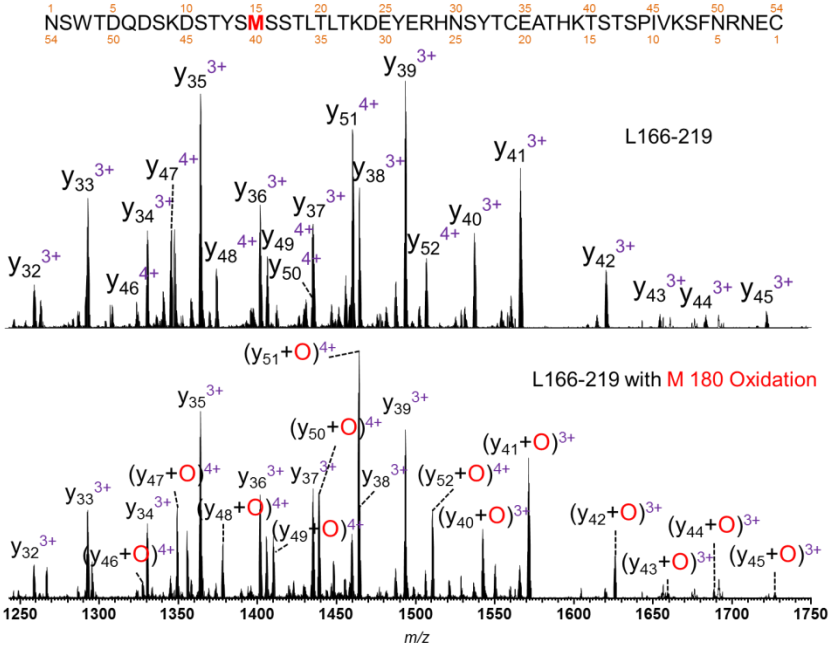


Figure 2.15. Part of the CID-MS/MS spectra of light-chain amino acids L166-219 (top) and oxidized L166-219 (bottom), demonstrating the oxidation at methionine 180. Peptides were obtained from 30-ms, in-membrane digestion of the light chain.

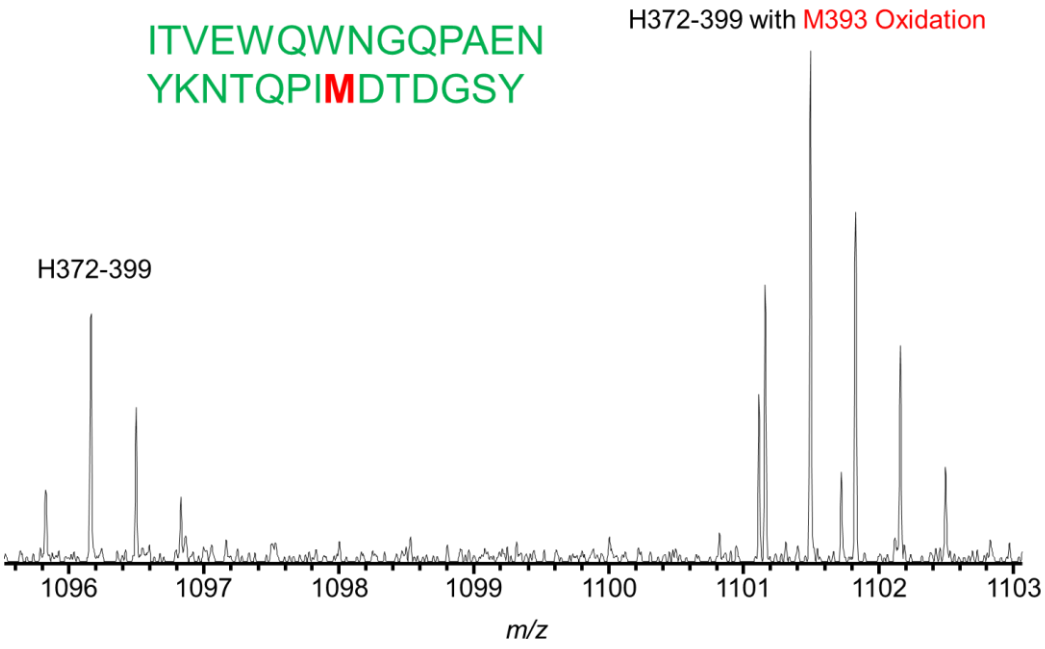


Figure 2.16. Part of the ESI-Orbitrap mass spectrum of a reduced WIgG1 heavy chain after in-membrane digestion with a residence time of 3 sec. The labeled isotope envelopes stem from the peptide H372-399 with (right) and without (left) oxidation.

Another important PTM, glycosylation, plays an important role for binding to the Fc receptor and, thus, affects antibody-dependent cell-mediated cytotoxicity. A 3-s digestion of the heavy chain produced peptides with amino acids H274-301, and these peptides show a clear mass distribution characteristic of three glycoforms (Figure 2.18). This result matches with characterization of the entire heavy chain by LC-ESI-Q-TOF MS (Figure 2.2). Overall, these data show that rapid, in-membrane digestion enables facile characterization of mAb PTMs.

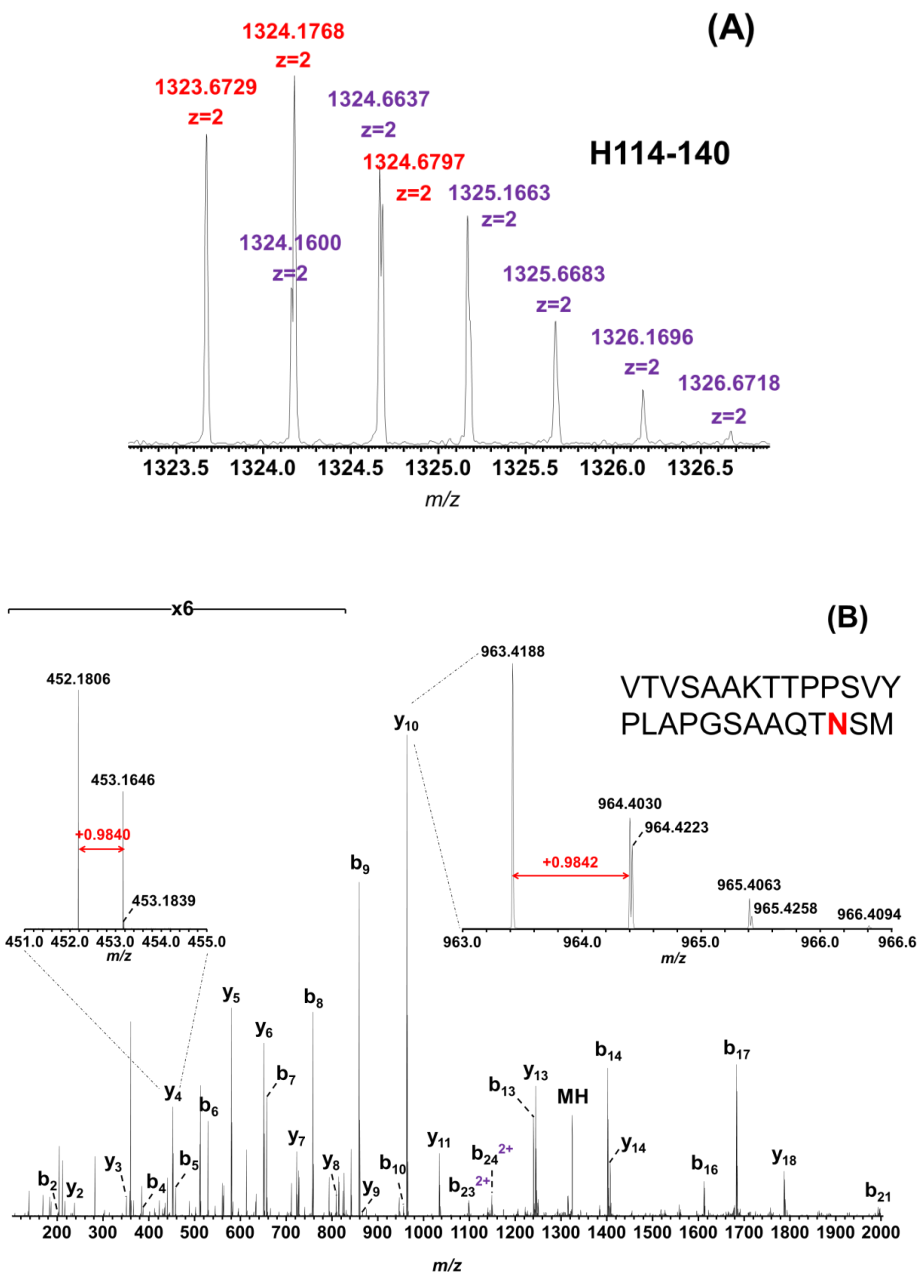


Figure 2.17. MS and CID-MS/MS spectra of the heavy-chain peptide H114-140 from a 3-s, in-membrane digestion. The spectra reveal deamidation on N138. (A) Part of the ESI-Orbitrap mass spectrum. The signals stem from the peptide H114-140 (red labels), and the same peptide with N138 deamidation (purple labels). (B) HCD-MS/MS reveals N138 deamidation. Insets show expanded regions of the spectrum.

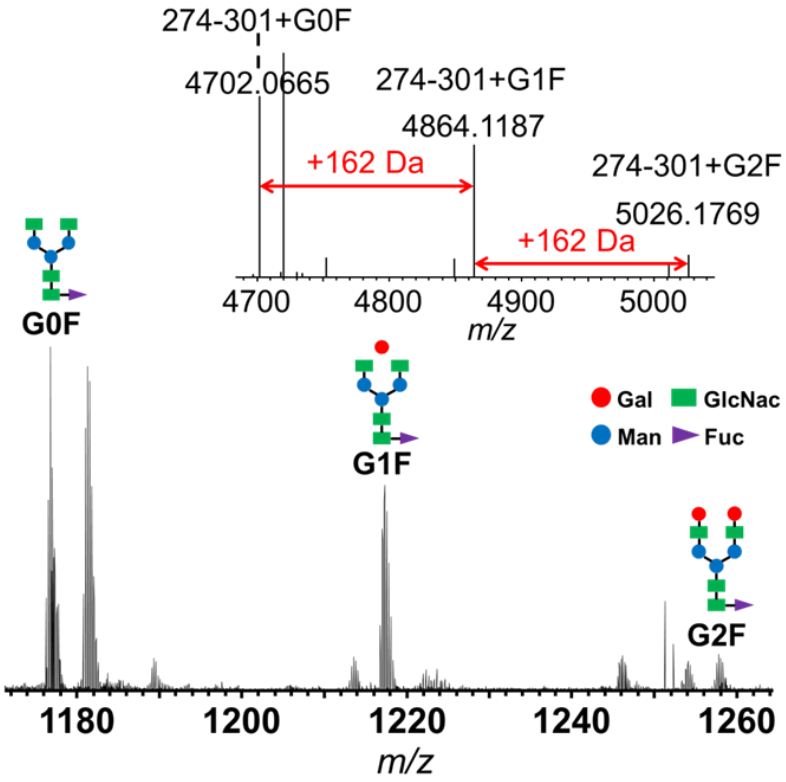


Figure 2.18. Part of the ESI-Orbitrap mass spectrum of a reduced WIgG1 heavy chain after digestion with a 3-s residence time in a pepsin-containing membrane. The labeled isotopic envelopes result from the peptide H274-301 (+4 charge state), which contains 3 N292 glycoforms separated by 162 Da intervals due to galactose units. The inset shows the spectrum that results from deconvolution of the mass range in the spectrum. Unlabeled peaks stem from other peptides. Drawings represent the different glycoforms with Gal=galactose, GlcNac=N-acetylglucosamine, Man=mannose, and Fuc=fucose.

2.4 Conclusion

This work used pepsin-modified membranes as controlled reactors for antibody proteolysis prior to MS analysis. Pepsin is an inexpensive protease that enables membrane digestion in acidic conditions, which avoids the need for antibody alkylation and minimizes oxidation during

digestion. Moreover, the high local enzyme concentration in membrane pores affords digestion of 100 μL of antibody solution in less than one minute. Variation of the residence times of reduced antibody solutions in the membranes yields "bottom-up" (1-2 kDa) to "middle-down" sized peptides (5-15 kDa) for the light and heavy chains, and these peptides cover the entire antibody sequence. As needed, digestion with different flow rates can enhance sequence coverage. Analysis of the light-chain and heavy-chain proteolytic peptides reveals sites for oxidation, deamidation and N-terminal pyroglutamic acid formation as well as glycosylation patterns. Furthermore, 30-ms in-membranes proteolysis of the light-chain followed by CID, HCD and ETD MS/MS of peptic peptides cleaves 99% of the amino acid bonds in the light chain. Traditional in-solution tryptic digestion of the light chain combined with CID and HCD-MS/MS also gives 99% sequence coverage, whereas "top-down" analysis of the entire light chain by CID, 5-ms, and 15-ms ETD shows a sequence coverage of 55%. With minimal sample preparation time, membrane digestion leads to high peptide and sequence coverages for identification of PTMs by MS.

2.5 Acknowledgement

We gratefully acknowledge the US National Science Foundation (CHE-1152762 and CHE-1506315) for funding this work. We thank Dr. Mohammad Muhsin Chisti from Michigan State Uni-versity for providing Trastuzumab. We also thank the Michigan State University Mass Spectrometry Facility and Dr. Todd Lydic from the Molecular Metabolism and Disease Collaborative Mass Spectrometry Core for helping analyze the samples.

REFERENCES

REFERENCES

(1) Weiner, G. J. Building better monoclonal antibody-based therapeutics. *Nat. Rev. Cancer* **2015,** *15* (6), 361.

(2) Weiner, L. M.; Surana, R.; Wang, S. Monoclonal antibodies: versatile platforms for cancer immunotherapy. *Nat. Rev. Immunol.* **2010,** *10* (5), 317.

(3) Beck, A.; Sanglier-Cianferani, S.; Van Dorsselaer, A. Biosimilar, biobetter, and next generation antibody characterization by mass spectrometry. *Anal. Chem.* **2012,** *84* (11), 4637.

(4) Reichert, J. M. Antibodies to watch in 2014. *MAbs* **2014,** *6* (1), 5.

(5) Duong, M. N.; Cleret, A.; Matera, E. L.; Chettab, K.; Mathe, D.; Valsesia-Wittmann, S.; Clemenceau, B.; Dumontet, C. Adipose cells promote resistance of breast cancer cells to trastuzumab-mediated antibody-dependent cellular cytotoxicity. *Breast Cancer Res.* **2015,** *17*, 57.

(6) Motta, G.; Cea, M.; Moran, E.; Carbone, F.; Augusti, V.; Patrone, F.; Nencioni, A. Monoclonal antibodies for non-Hodgkin's lymphoma: state of the art and perspectives. *Clin. Dev. Immunol.* **2010,** *2010*, 428253.

(7) Saltz, L.; Easley, C.; Kirkpatrick, P. Panitumumab. *Nat. Rev. Drug Discov.* **2006,** *5* (12), 987.

(8) Rathore, A. S.; Winkle, H. Quality by design for biopharmaceuticals. *Nat. Biotechnol.* **2009,** *27* (1), 26.

(9) Beck, A.; Wagner-Rousset, E.; Ayoub, D.; Van Dorsselaer, A.; Sanglier-Cianferani, S. Characterization of therapeutic antibodies and related products. *Anal. Chem.* **2013,** *85* (2), 715.

(10) Rosati, S.; van den Bremer, E. T.; Schuurman, J.; Parren, P. W.; Kamerling, J. P.; Heck, A. J. In-depth qualitative and quantitative analysis of composite glycosylation profiles and other micro-heterogeneity on intact monoclonal antibodies by high-resolution native mass spectrometry using a modified Orbitrap. *MAbs* **2013,** *5* (6), 917.

(11) Yan, B.; Valliere-Douglass, J.; Brady, L.; Steen, S.; Han, M.; Pace, D.; Elliott, S.; Yates, Z.; Han, Y.; Balland, A.et al. Analysis of post-translational modifications in recombinant monoclonal antibody IgG1 by reversed-phase liquid chromatography/mass spectrometry. *J. Chromatogr. A* **2007,** *1164* (1-2), 153.

(12) Zhang, Z.; Pan, H.; Chen, X. Mass spectrometry for structural characterization of therapeutic antibodies. *Mass Spectrom. Rev.* **2009,** *28* (1), 147.

(13) Rosati, S.; Yang, Y.; Barendregt, A.; Heck, A. J. Detailed mass analysis of structural heterogeneity in monoclonal antibodies using native mass spectrometry. *Nat. Protoc.* **2014,** *9* (4), 967.

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

(15) Song, T.; Ozcan, S.; Becker, A.; Lebrilla, C. B. In-depth method for the characterization of glycosylation in manufactured recombinant monoclonal antibody drugs. *Anal. Chem.* **2014,** *86* (12), 5661.

(16) Cai, B.; Pan, H.; Flynn, G. C. C-terminal lysine processing of human immunoglobulin G2 heavy chain in vivo. *Biotechnol. Bioeng.* **2011,** *108* (2), 404.

(17) Liu, Y. D.; van Enk, J. Z.; Flynn, G. C. Human antibody Fc deamidation in vivo. *Biologicals* **2009,** *37* (5), 313.

(18) Montesino, R.; Calvo, L.; Vallin, A.; Rudd, P. M.; Harvey, D. J.; Cremata, J. A. Structural characterization of N-linked oligosaccharides on monoclonal antibody Nimotuzumab through process development. *Biologicals* **2012,** *40* (4), 288.

(19) Maeda, E.; Kita, S.; Kinoshita, M.; Urakami, K.; Hayakawa, T.; Kakehi, K. Analysis of nonhuman N-glycans as the minor constituents in recombinant monoclonal antibody pharmaceuticals. *Anal. Chem.* **2012,** *84* (5), 2373.

(20) Huang, L.; Lu, J.; Wroblewski, V. J.; Beals, J. M.; Riggin, R. M. In vivo deamidation characterization of monoclonal antibody by LC/MS/MS. *Anal. Chem.* **2005,** *77* (5), 1432.

(21) Bailey, M. J.; Hooker, A. D.; Adams, C. S.; Zhang, S.; James, D. C. A platform for high-throughput molecular characterization of recombinant monoclonal antibodies. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2005,** *826* (1-2), 177.

(22) Zhang, Z.; Shah, B. Characterization of variable regions of monoclonal antibodies by top-down mass spectrometry. *Anal. Chem.* **2007,** *79* (15), 5723.

(23) Bondarenko, P. V.; Second, T. P.; Zabrouskov, V.; Makarov, A. A.; Zhang, Z. Mass measurement and top-down HPLC/MS analysis of intact monoclonal antibodies on a hybrid linear quadrupole ion trap-Orbitrap mass spectrometer. *J. Am. Soc. Mass. Spectrom.* **2009,** *20* (8), 1415.

(24) Tsybin, Y. O.; Fornelli, L.; Stoermer, C.; Luebeck, M.; Parra, J.; Nallet, S.; Wurm, F. M.; Hartmer, R. Structural analysis of intact monoclonal antibodies by electron transfer dissociation mass spectrometry. *Anal. Chem.* **2011,** *83* (23), 8919.

(25) Mao, Y.; Valeja, S. G.; Rouse, J. C.; Hendrickson, C. L.; Marshall, A. G. Top-down structural analysis of an intact monoclonal antibody by electron capture dissociation-

Fourier transform ion cyclotron resonance-mass spectrometry. *Anal. Chem.* **2013,** *85* (9), 4239.

(26) Srzentic, K.; Fornelli, L.; Laskay, U. A.; Monod, M.; Beck, A.; Ayoub, D.; Tsybin, Y. O. Advantages of extended bottom-up proteomics using Sap9 for analysis of monoclonal antibodies. *Anal. Chem.* **2014,** *86* (19), 9945.

(27) Ayoub, D.; Bertaccini, D.; Diemer, H.; Wagner-Rousset, E.; Colas, O.; Cianferani, S.; Van Dorsselaer, A.; Beck, A.; Schaeffer-Reiss, C. Characterization of the N-Terminal Heterogeneities of Monoclonal Antibodies Using In-Gel Charge Derivatization of alpha-Amines and LC-MS/MS. *Anal. Chem.* **2015,** *87* (7), 3784.

(28) Li, H.; Ortiz, R.; Tran, L. T.; Salimi-Moosavi, H.; Malella, J.; James, C. A.; Lee, J. W. Simultaneous analysis of multiple monoclonal antibody biotherapeutics by LC-MS/MS method in rat plasma following cassette-dosing. *AAPS J.* **2013,** *15* (2), 337.

(29) Zhang, T.; Zhang, J.; Hewitt, D.; Tran, B.; Gao, X.; Qiu, Z. J.; Tejada, M.; Gazzano-Santoro, H.; Kao, Y. H. Identification and characterization of buried unpaired cysteines in a recombinant monoclonal IgG1 antibody. *Anal. Chem.* **2012,** *84* (16), 7112.

(30) Du, Y.; Wang, F.; May, K.; Xu, W.; Liu, H. LC-MS analysis of glycopeptides of recombinant monoclonal antibodies by a rapid digestion procedure. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2012,** *907*, 87.

(31) Wang, Y.; Lu, Q.; Wu, S. L.; Karger, B. L.; Hancock, W. S. Characterization and comparison of disulfide linkages and scrambling patterns in therapeutic monoclonal antibodies: using LC-MS with electron transfer dissociation. *Anal. Chem.* **2011,** *83* (8), 3133.

(32) Lesur, A.; Varesio, E.; Hopfgartner, G. Accelerated tryptic digestion for the analysis of biopharmaceutical monoclonal antibodies in plasma by liquid chromatography with tandem mass spectrometric detection. *J. Chromatogr. A* **2010,** *1217* (1), 57.

(33) Xiang, T.; Chumsae, C.; Liu, H. Localization and quantitation of free sulfhydryl in recombinant monoclonal antibodies by differential labeling with 12C and 13C iodoacetic acid and LC-MS analysis. *Anal. Chem.* **2009,** *81* (19), 8101.

(34) Rehder, D. S.; Dillon, T. M.; Pipes, G. D.; Bondarenko, P. V. Reversed-phase liquid chromatography/mass spectrometry analysis of reduced monoclonal antibodies in pharmaceutics. *J. Chromatogr. A* **2006,** *1102* (1-2), 164.

(35) Wang, L.; Amphlett, G.; Lambert, J. M.; Blattler, W.; Zhang, W. Structural characterization of a recombinant monoclonal antibody by electrospray time-of-flight mass spectrometry. *Pharm. Res.* **2005,** *22* (8), 1338.

(36) Zhang, J.; Liu, H.; Katta, V. Structural characterization of intact antibodies by high-resolution LTQ Orbitrap mass spectrometry. *J. Mass Spectrom.* **2010,** *45* (1), 112.

(37) Wang, B.; Gucinski, A. C.; Keire, D. A.; Buhse, L. F.; Boyne, M. T., 2nd. Structural comparison of two anti-CD20 monoclonal antibody drug products using middle-down mass spectrometry. *Analyst* **2013,** *138* (10), 3058.

(38) Fornelli, L.; Ayoub, D.; Aizikov, K.; Beck, A.; Tsybin, Y. O. Middle-down analysis of monoclonal antibodies with electron transfer dissociation orbitrap fourier transform mass spectrometry. *Anal. Chem.* **2014,** *86* (6), 3005.

(39) An, Y.; Zhang, Y.; Mueller, H. M.; Shameem, M.; Chen, X. A new tool for monoclonal antibody analysis: application of IdeS proteolysis in IgG domain-specific characterization. *MAbs* **2014,** *6* (4), 879.

(40) Wang, D.; Wynne, C.; Gu, F.; Becker, C.; Zhao, J.; Mueller, H. M.; Li, H.; Shameem, M.; Liu, Y. H. Characterization of drug-product-related impurities and variants of a therapeutic monoclonal antibody by higher energy C-trap dissociation mass spectrometry. *Anal. Chem.* **2015,** *87* (2), 914.

(41) Nicolardi, S.; Deelder, A. M.; Palmblad, M.; van der Burgt, Y. E. Structural analysis of an intact monoclonal antibody by online electrochemical reduction of disulfide bonds and Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **2014,** *86* (11), 5376.

(42) Fornelli, L.; Damoc, E.; Thomas, P. M.; Kelleher, N. L.; Aizikov, K.; Denisov, E.; Makarov, A.; Tsybin, Y. O. Analysis of intact monoclonal antibody IgG1 by electron transfer dissociation Orbitrap FTMS. *Mol. Cell. Proteomics.* **2012,** *11* (12), 1758.

(43) Zang, L.; Carlage, T.; Murphy, D.; Frenkel, R.; Bryngelson, P.; Madsen, M.; Lyubarskaya, Y. Residual metals cause variability in methionine oxidation measurements in protein pharmaceuticals using LC-UV/MS peptide mapping. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2012,** *895-896*, 71.

(44) Dick, L. W., Jr.; Mahon, D.; Qiu, D.; Cheng, K. C. Peptide mapping of therapeutic monoclonal antibodies: improvements for increased speed and fewer artifacts. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2009,** *877* (3), 230.

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

(46) Hahne, H.; Pachl, F.; Ruprecht, B.; Maier, S. K.; Klaeger, S.; Helm, D.; Medard, G.; Wilm, M.; Lemeer, S.; Kuster, B. DMSO enhances electrospray response, boosting sensitivity of proteomic experiments. *Nat. Methods* **2013,** *10* (10), 989.

(47) Swaney, D. L.; Wenger, C. D.; Coon, J. J. Value of using multiple proteases for large-scale mass spectrometry-based proteomics. *J. Proteome. Res.* **2010,** *9* (3), 1323.

(48) Wu, C.; Tran, J. C.; Zamdborg, L.; Durbin, K. R.; Li, M.; Ahlf, D. R.; Early, B. P.; Thomas, P. M.; Sweedler, J. V.; Kelleher, N. L. A protease for 'middle-down' proteomics. *Nat. Methods* **2012,** *9* (8), 822.

(49) Yuan, Z. F.; Arnaudo, A. M.; Garcia, B. A. Mass spectrometric analysis of histone proteoforms. *Annu. Rev. Anal. Chem.* **2014,** *7*, 113.

(50) Tan, Y. J.; Wang, W. H.; Zheng, Y.; Dong, J.; Stefano, G.; Brandizzi, F.; Garavito, R. M.; Reid, G. E.; Bruening, M. L. Limited proteolysis via millisecond digestions in protease-modified membranes. *Anal. Chem.* **2012,** *84* (19), 8357.

(51) Good, D. M.; Wirtala, M.; McAlister, G. C.; Coon, J. J. Performance characteristics of electron transfer dissociation mass spectrometry. *Mol. Cell. Proteomics* **2007,** *6* (11), 1942.

(52) Han, W.; Yamauchi, M.; Hasegawa, U.; Noda, M.; Fukui, K.; van der Vlies, A. J.; Uchiyama, S.; Uyama, H. Pepsin immobilization on an aldehyde-modified polymethacrylate monolith and its application for protein analysis. *J. Biosci. Bioeng.* **2015,** *119* (5), 505.

(53) Long, Y.; Wood, T. D. Immobilized pepsin microreactor for rapid peptide mapping with nanoelectrospray ionization mass spectrometry. *J. Am. Soc. Mass. Spectrom.* **2015,** *26* (1), 194.

(54) Udeshi, N. D.; Compton, P. D.; Shabanowitz, J.; Hunt, D. F.; Rose, K. L. Methods for analyzing peptides and proteins on a chromatographic timescale by electron-transfer dissociation mass spectrometry. *Nat. Protoc.* **2008,** *3* (11), 1709.

(55) Tan, Y. J.; Sui, D.; Wang, W. H.; Kuo, M. H.; Reid, G. E.; Bruening, M. L. Phosphopeptide enrichment with TiO2-modified membranes and investigation of tau protein phosphorylation. *Anal. Chem.* **2013,** *85* (12), 5699.

(56) Beck, A.; Debaene, F.; Diemer, H.; Wagner-Rousset, E.; Colas, O.; Van Dorsselaer, A.; Cianferani, S. Cutting-edge mass spectrometry characterization of originator, biosimilar and biobetter antibodies. *J. Mass Spectrom.* **2015,** *50* (2), 285.

(57) Marcoux, J.; Champion, T.; Colas, O.; Wagner-Rousset, E.; Corvaia, N.; Van Dorsselaer, A.; Beck, A.; Cianferani, S. Native mass spectrometry and ion mobility characterization of trastuzumab emtansine, a lysine-linked antibody drug conjugate. *Protein Sci.* **2015**, *24*, 1210.

(58) Lew, C.; Gallegos-Perez, J. L.; Fonslow, B.; Lies, M.; Guttman, A. Rapid level-3 characterization of therapeutic antibodies by capillary electrophoresis electrospray ionization mass spectrometry. *J. Chromatogr. Sci.* **2015,** *53* (3), 443.

(59) Gahoual, R.; Burr, A.; Busnel, J. M.; Kuhn, L.; Hammann, P.; Beck, A.; Franccois, Y. N.; Leize-Wagner, E. Rapid and multi-level characterization of trastuzumab using sheathless capillary electrophoresis-tandem mass spectrometry. *Mabs* **2013,** *5* (3), 479.

(60) Xie, H.; Chakraborty, A.; Ahn, J.; Yu, Y. Q.; Dakshinamoorthy, D. P.; Gilar, M.; Chen, W.; Skilton, S. J.; Mazzeo, J. R. Rapid comparison of a candidate biosimilar to an innovator monoclonal antibody with advanced liquid chromatography and mass spectrometry technologies. *MAbs* **2010,** *2* (4), 379.

(61) Olsen, J. V.; Macek, B.; Lange, O.; Makarov, A.; Horning, S.; Mann, M. Higher-energy C-trap dissociation for peptide modification analysis. *Nat. Methods* **2007,** *4* (9), 709.

(62) Hu, D.; Qin, Z.; Xue, B.; Fink, A. L.; Uversky, V. N. Effect of methionine oxidation on the structural properties, conformational stability, and aggregation of immunoglobulin light chain LEN. *Biochemistry* **2008,** *47* (33), 8665.

Chapter 3 . Enzyme-Containing Spin Membranes for Rapid Protein Digestion

Proteolytic digestion is an important step in protein characterization using mass spectrometry (MS). This study uses pepsin- or trypsin-containing spin membranes for rapid protein digestion prior to ultrahigh-resolution Orbitrap MS analysis. Centrifugation of 100 μL of pretreated protein solutions through the functionalized membranes requires less than 1 min. Peptic and tryptic peptides from spin digestion of apomyoglobin and four commercial monoclonal antibodies (mAbs) cover nearly 100% of the protein sequences in direct infusion MS analysis. Increasing the spin rate leads to a higher fraction of large peptic peptides for apomyoglobin, and MS analysis of peptic and tryptic peptides reveals post-translational mAb modifications such as N-terminal pyroglutamate formation, C-terminal Lysine clipping and glycosylation. Analysis of tryptic spin digests with liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) and MaxQuant data searching gives 100% sequence coverage of all four antibody light chains, and 75.1%-98.4% coverage of the heavy chains. Compared to in-solution tryptic digestion of mAbs, trypsin spin digestion yields higher sequence coverage and a larger number of unique peptides.

3.1 Introduction

Proteolysis is often a crucial step in protein characterization, identification, and quantitation through mass spectrometry (MS) and/or tandem mass spectrometry (MS/MS) analysis.[1]

Compared to MS and MS/MS characterization of intact proteins, analyses of proteolytic peptides yield greater sequence information as well as greater resolution in separations with liquid chromatography (LC).[2] However, conventional peptide generation using in-solution digestion requires long incubation times (up to 24 h) because of the low enzyme concentrations required to avoid self-digestion.[3-7] Unfortunately, oxidation or other protein modifications may occur during long digestions.[8,9] Minimizing hydrogen/deuterium exchange during digestion is also vital to avoid back exchange in studies of protein structure based on hydrogen/deuterium exchange prior to digestion.[10] Thus, rapid digestion is important for a number of protein characterization studies.

Several research groups and companies developed immobilized-enzyme reactors for rapid protein digestion.[11,12] The high enzyme-to-protein ratio in these reactors greatly improves the digestion rate, and immobilization can also increase enzyme stability and decrease autolysis.[3-7,13] Solid supports employed to create immobilized-enzyme reactors include monoliths,[14-29] capillaries,[30,31] magnetic particles,[32-34] resins,[35-38] microfluidic chips,[39,40] and membranes.[41-44] We are particularly interested in membrane supports because they are inexpensive, and varying the flowrate through these thin structures affords solution residence times that range from msec to sec.

Initially, enzyme immobilization in membranes relied on hydrophobic interactions in poly (vinylidene difluoride).[44,45] Xu and coworkers later formed a trypsin-containing membrane through sequential adsorption of poly (styrene sulfonate) (PSS) and trypsin in porous nylon.[41] PSS absorbs strongly to nylon, presumably through multiple hydrophobic interactions, to create a negatively charged surface. With a pI of ~10.5, trypsin is positively charged in acidic solution and electrostatically adsorbs to negatively charged PSS-modified membranes. This adsorption procedure gives a membrane reactor with a local concentration of 10 mg of trypsin per mL of

membrane pores, which is 450 times higher than the typical trypsin concentration for in-solution digestion. The short radial diffusion distances within the microporous membrane pores further facilitate rapid digestion. Tan et al. used a similar strategy to form pepsin-containing membranes.[42]

Recently, we exploited a pepsin-containing membrane to facilitate monoclonal antibody (mAb) characterization.[43] By varying the antibody residence time (from 3 ms to 3 s) in the membrane, we obtained "bottom-up" (1-2 kDa) to "middle-down" (5-15 kDa) sized peptides, and these peptides cover the entire sequences of two different antibodies. However, for all the aforementioned membrane-based protein digestions, protein passage through the membrane employed relatively cumbersome systems that included syringe or peristaltic pumps. To overcome this challenge, we recently developed a membrane fitting that attaches to a disposable pipette tip.[46] This allows rapid digestion, but loading the membrane into the ferrule fitting, achieving a good seal, and extended production of the device are challenging.

This chapter describes protein digestion using spin membranes containing immobilized pepsin or trypsin. During centrifugation, protein solutions pass through the spin membrane in 1 min or less to yield proteolytic peptides for subsequent direct infusion or LC-MS/MS analysis with an Orbitrap ultrahigh resolution mass spectrometer. Direct infusion analysis of apomyoglobin and four commercial monoclonal antibodies (Herceptin, Avastin, Rituxan and Vectibix) yields nearly 100% sequence coverage. LC-MS/MS analysis of tryptic spin digests of four antibodies followed by MaxQuant data processing easily identifies the four antibodies via comparison with a protein database. In these LC-MS/MS analyses, sequence coverages are 100% for all the light chains, and range from 75.1 to 98.4% for the different antibody heavy chains.

3.2 Experimental

3.2.1 Materials

Nylon membranes (LoProdyne LP, nominal pore size 1.2 μm, 110 μm thickness) were purchased from Pall Corporation. Trastuzumab (Herceptin, Genentech), Bevacizumab (Avastin, Genentech), Rituxan (Rituximab, Genentech) and Vectibix (Panitumumab, Amgen) were obtained in their commercial formulations as a gift from Dr. Muhammad Chisti of Michigan State University. Trypsin from bovine pancreas (TPCK-treated, lyophilized powder, ⩾10,000 BAEE units/mg protein), pepsin from porcine gastric mucosa (lyophilized powder, 3200-4500 units/mg protein), ammonium bicarbonate (⩾99%), iodoacetamide (IAM, ⩾99%), dithiothreitol (DTT, ⩾99.5%), polystyrene sulfonate (PSS, average molecular weight ~70,000), formic acid (FA, ⩾98%) and acetonitrile (ACN, HPLC grade, ⩾99.9%) were purchased from Sigma Aldrich. Sequencing grade modified trypsin was obtained from Promega. NaCl (ACS grade) and HCl (ACS grade) were purchased from CCI. Other chemicals include urea (⩾ 98%, Invitrogen), tris(2-carboxyethyl) phosphine hydrochloride (TCEP-HCl, >98%, Fluka), trifluoroacetic acid (TFA, EMD), acetic acid (HOAc, ACS, Macron Fine Chemicals), and methyl alcohol (anhydrous, MeOH, Macron Fine Chemicals). Solutions were prepared in deionized water (DI water, Milli-Q, 18.2 MΩ·cm at 25 °C). C4 ZipTips were purchased from EMD Millipore, and Pierce C18 spin columns were used to isolate tryptic peptides after digestion. Amicon ultra 0.5 mL centrifugal filters (MWCO 10 kDa) were employed to desalt samples before pepsin in-membrane digestion, and an Eppendorf centrifuge (5415D) was used to conduct spin digestion.

3.2.2 Functionalized Membrane-Containing Spin Columns

Trypsin- and pepsin-containing membranes were prepared using a slight modification of our literature procedure.[41-43] Membranes were UV/ozone-cleaned for 10 min, and 10 mL of 0.02 M PSS in 0.5 M NaCl (pH=2.3) was circulated through the membrane for 10 min using a peristaltic pump, followed by rinsing with 30 mL of DI water. For trypsin-containing membranes after adsorption of PSS, 5 mL of 1 mg/mL trypsin in 2.7 mM HCl was circulated through the membrane for 1 h. Subsequently, the membrane was rinsed with 30 mL of 1 mM HCl, dried with N_2, and stored in a desiccator. For pepsin-containing membranes, 4 mL of 2 mg/mL pepsin in 5% FA was circulated through the membrane for 1 h. Then, the membrane was rinsed with 30 mL of 5% FA, dried with N_2, and stored in a desiccator. Flow rates during membrane modification were 2 mL/min. The modified membranes were embedded in spin devices at Takara/Clontech Laboratories (Mountain View, CA). These devices expose a membrane surface with a diameter of ~1.8 mm.

3.2.3 Apomyoglobin spin digestion with pepsin- and trypsin-containing membranes

Apomyoglobin (10 μg) was dissolved in 100 μL of 10 mM NH_4HCO_3 for trypsin digestion, and in 100 μL of 5% FA for pepsin digestion. The spin column was rinsed with 100 μL of 10 mM NH_4HCO_3 or 100 μL of 5% FA before tryptic or peptic spin digestion, respectively. Both enzymatic digestions were conducted at three spin rates corresponding to 500 *g* and 10,000 *g*. The centrifugation time was 1 min, and the digests were dried with a SpeedVac after spin digestion and immediately reconstituted for MS analysis.

3.2.4 mAb spin digestion with pepsin- and trypsin-containing membranes

For pepsin digestion, Trastuzumab, Bevacizumab, Rituximab and Panitumumab were each diluted in deionized water to prepare stock solutions with 1 mg/mL of antibody. Subsequently, 2 μL of 0.1 M HOAc and 2 μL of 0.1 M TCEP-HCl were added to 20 μL of mAb stock solutions, and these mixtures were incubated at 75 ℃ for 15 min. Subsequent buffer exchange with 5% FA employed 3 cycles of centrifugation with an Amicon ultra 0.5 mL centrifugal filter (MWCO 10 kDa). About 25 μL of solution remained after each centrifugation, and 475 μL of 5% FA was added prior to the following centrifugation. Residues were diluted to 200 μL with 5% FA to make 0.1 mg/mL solutions.

For trypsin digestion, 4 μL of 10 mg/mL antibody stock solutions of each of the four mAbs were diluted separately in 14 μL of 2 mM TCEP-HCl solution in 0.1% HOAc containing 8 M urea. The mixtures were incubated at 50 ℃ for 10 min prior to addition of 14 μL of 20 mM IAM in 2 M NH_4HCO_3 containing 8 M urea, and incubation in the dark for 30 min. Finally, 12 μL of 30 mM DTT in 100 mM NH_4HCO_3 containing 8 M urea was added followed by incubation in the dark for 20 min to quench the IAM. After reduction and alkylation, the residual solutions were diluted with deionized water to create 0.1 mg/mL solutions.

3.2.4.1 In-membrane spin digestion of mAbs

Within 1.5 h of antibody pretreatments, 200 μL of each nonalkylated antibody solution was added to a pepsin spin column, and 200 μL of each alkylated antibody solution was added to a trypsin spin column. The solutions were centrifuged through the membrane for 1 min at 500 *g*. Pepsin spin digestion samples were collected for direct infusion MS and LC-MS/MS analysis, whereas trypsin spin digestion samples were first desalted using Pierce C18 spin cartridges

(following the manufacturer's protocol) before infusion and LC/MS analysis. The C18 spin column was activated with 50% MeOH, and equilibrated in 0.5% TFA in 5% ACN. Then, the sample was loaded onto the column, followed by washing with of 0.5% TFA in 5% ACN. Finally, the peptides were eluted from the spin column with 70% ACN. Antibody in-membrane spin digestions were also monitored by SDS-PAGE. The reproducibility of the spin-membrane digestion was tested by running triplicate trypsin and pepsin digestions of Bevacizumab with a separate membrane for each digestion.

3.2.4.2 In-solution trypsin digestion of mAbs

We conducted in-solution tryptic digestion of four antibodies to compare in-membrane and spin digestion. Five microliters of 0.2 μg/μL sequencing grade modified trypsin solution was added to 200 μL of the 0.1 mg/mL alkylated antibody solution prior to incubation at 37 ℃ for 16 h. The reaction was quenched by adding 5 μL of acetic acid. Samples were then desalted with a C18 spin column and dried with a SpeedVac before reconstitution and infusion MS or LC-MS/MS analysis.

3.2.5 Mass Spectrometry and Data Analysis

For direct infusion MS, in-membrane spin digests and in-solution digests were dried with a SpeedVac and reconstituted in 1% acetic acid, 49% H_2O, and 50% methanol within 1 day. Then, 40 μL of each sample was loaded into a Whatman multichem 96-well plate (Sigma−Aldrich) and sealed with Teflon Ultrathin Sealing Tape (Analytical Sales and Services, Prompton Plains, NJ). An Advion Triversa Nanomate nanoelectrospray ionization (nESI) source (Advion, Ithaca, NY) was used to introduce the sample into a high-resolution accurate mass Thermo Fisher Scientific LTQ Orbitrap Velos mass spectrometer (San Jose, CA) that was equipped with a dual pressure

ion trap, HCD cell, and ETD. The spray voltage and gas pressure were set to 1.4 kV and 1.0 psi, respectively. The ion-source interface had an inlet temperature of 200 ℃ with an S-Lens value of 57%. High-resolution mass spectra were acquired in positive ionization mode across the *m/z* range of 400−1800, using the FT analyzer operating at a mass resolving power of 100,000. Spectra were the average of 100 scans. Signals with >1% of the highest peak intensities and S/N>3 were analyzed. Peptide identification was performed manually using ProteinProspector (v 5.14.1, University of California, San Francisco, CA). Mass tolerance was set to 10 ppm.

For LC-MS/MS, Nano-Ultra High Performance Liquid Chromatography MS/MS was performed essentially as described previously.[47] Briefly, 2uL injections corresponding to 500 ng of spin-digested tryptic protein (reconstituted in 0.1% FA) were loaded onto a 100 mm x 75 um C18-BEH column (Waters Billerica, MA), and separated over a 90 min gradient from 5-35%B on a nano-Acquity system (Waters) flowing at 500 nL/min. Solution A was 0.1% FA in H_2O, and solution B was 0.1% FA in ACN. MS/MS was performed on an LTQ-Velos Orbitrap-FTMS (Thermo, San Jose, CA) running a top-20 data dependent method, where a single MS at a resolution of 60,000 was acquired, and the top-20 precursors were selected for fragmentation.

Raw LC-MS/MS files were processed by MaxQuant version 1.5.6.0. MS/MS spectra were searched against the Cricetulus griseus (Chinese hamster) proteome (23,884 proteins). The database also included common contaminants and the antibody sequences. MaxQuant analysis parameters included a precursor mass tolerance of 20 ppm for the initial search, a precursor mass tolerance of 6 ppm for the main search, and an FTMS MS/MS match tolerance of 20 ppm. We set trypsin as the specific enzyme. Variable modifications included oxidation (M), deamidation (NQ), and Gln->pyro-Glu, while the fixed modification was carbamidomethyl on cysteine. The

minimal peptide length was set to 6 amino acids, the maximum peptide mass was 8000 Da, and the maximum number of missed cleavages was 5.

3.3 Results and discussion

3.3.1 Workflow for Digestion in Membrane-Containing Spin Columns

Figure 3.1 shows the workflow that we use to conduct protein digestion in spin membranes. After protein pretreatment, the solution simply passes through the membrane reactor during centrifugation. Digestion of 100 µL of protein solution requires a centrifugation time less than 30 sec. The high concentration of enzyme in the membrane pores affords rapid digestion of protein, and we can control the digestion by varying the spin rate. Moreover, the tiny dead volume (0.275 µL) of the spin membrane should minimize peptide loss during digestion.

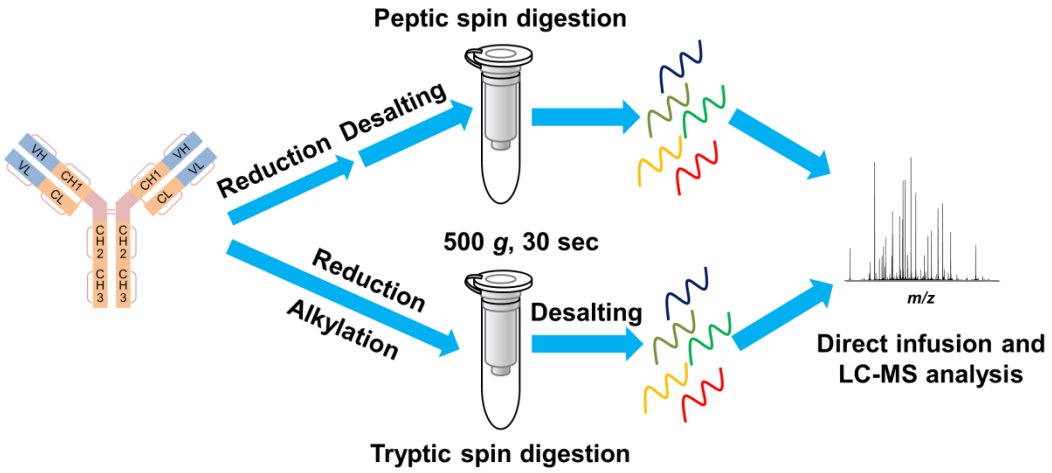


Figure 3.1. Workflow for protein spin digestion and analysis. [Acronyms: VH-variable region of the heavy chain; CH1, CH2, and CH3- different constant regions of the heavy chain; CL- constant region of the light chain; VL- variable region of the light chain; LC- light chain; HC- heavy chain.]

3.3.2 Apomyoglobin spin digestion with pepsin- and trypsin-containing membranes

We chose apomyoglobin, a common standard for peptide mapping, to initially test the spin digestion. Apomyoglobin has a compact hydrophobic core at neutral pH and undergoes slow in-solution proteolysis at pH 8.[48] Reduction and alkylation are not necessary for digesting this protein because it has no disulfide bonds. Using both trypsin and pepsin spin membranes and different spin rates prior to infusion MS analysis, we always observed 100% apomyoglobin sequence coverage (percent of protein sequence covered by identified proteolytic peptides) after spin digestion in a single pass through the membrane.

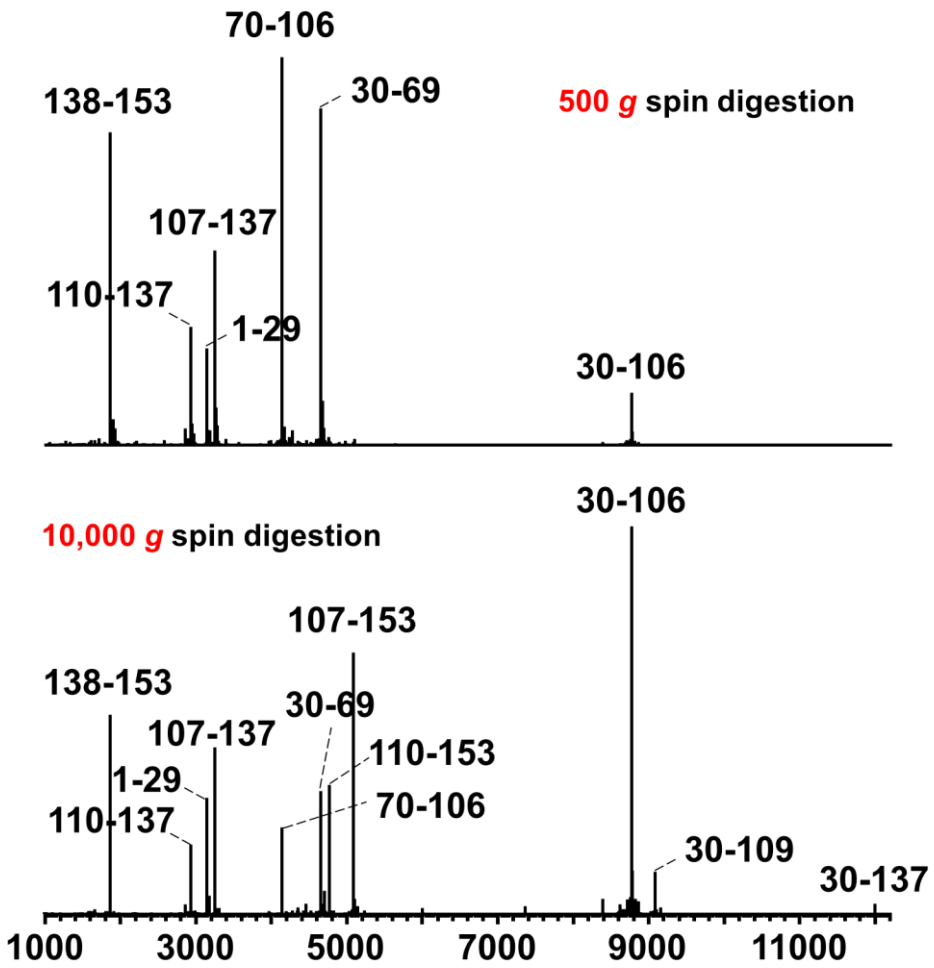


Figure 3.2. Deconvoluted ESI-Orbitrap mass spectra of apomyoglobin peptic digests obtained through 500 *g* (top) and 10,000 *g* (bottom) spin digestion. Deconvoluted mass spec were generated with Xtract software.

Figure 3.2 shows ESI-Orbitrap deconvoluted mass spectra of peptic apomyoglobin digests obtained using spin digestion at 500 *g* and 10,000 *g*. Four peptides, 1-29, 30-106, 107-137, and 138-153 cover the whole sequence after digestion at 500 *g*, whereas three peptides 1-29, 30-106, and 107-153 do the same thing after digestion at 10,000 *g*. When the spin rate increases to 10,000 *g*, the signals of large peptides such as amino acids 30-106 and 107-153 increase dramatically. At the same time, signals for smaller peptides, including amino acids 30-69 and 70-

106 decrease. Increasing the spin rate leads to a higher fraction of large peptides, presumably because lower residence times in the membrane decrease the proteolysis time to generate more missed cleavages. By varying the centrifugation rate, we can obtain overlapping peptides using a single enzyme.

The aspartic protease pepsin exhibits less specificity than trypsin.[49] However, extensive studies of pepsin digestion show that this protease prefers to cleave peptide bonds after phenylalanine (F) and leucine (L).[50] Our results match the cleavage site preferences for pepsin. Peptic peptides 30-69, 70-106, 107-153, and 138-153 result from cleavage of the 29L-30I, 69L-70T, 106F-107I, and 137L-138F bonds. Table 3.1 and Table 3.2 show the peptic peptides identified from digestion at 500 *g* and 10,000 *g*, respectively.

Trypsin is the most common enzyme used for protein digestion, particularly for bottom-up proteomics.[51] Compared with pepsin, it has higher specificity, cleaving proteins and peptides at the C-terminus of K and R, except when followed by P. Moreover, at low pH tryptic peptides carry at least two positive charges, which benefits downstream collision-induced dissociation tandem mass spectrometry (CID-MS/MS) analysis. Figure 4.3 shows the deconvoluted ESI-Orbitrap mass spectrum of a tryptic spin digest of apomyoglobin. A spin at 500 *g* gives complete digestion of apomyoglobin (no intact apomyoglobin) in one pass through the membrane. We identified 26 tryptic peptides, and as few as seven tryptic peptides cover 100% of the sequence: amino acids 1-31, 32-47, 48-63, 64-77, 78-96, 97-133, and 134-153. Table 3.3 gives a full list of the identified peptides, most of which contain 1 or more missed cleavage sites. Different from peptic spin digestion, we don't see the emergence of large peptides at higher rates of centrifugation. Instead we see signals of undigested, intact protein. This may stem from the compact structure of apomyoglobin at pH 7-8 as well as the high activity of trypsin. After an

initial cleavage, structures of the resulting peptides likely open rapidly to allow further digestion, although the initial cleavage is slow. Thus, changing the spin rate yields intact protein rather than limited proteolysis. In future studies, covalent linking of trypsin to the membrane may decrease its activity to better obtain limited proteolysis.

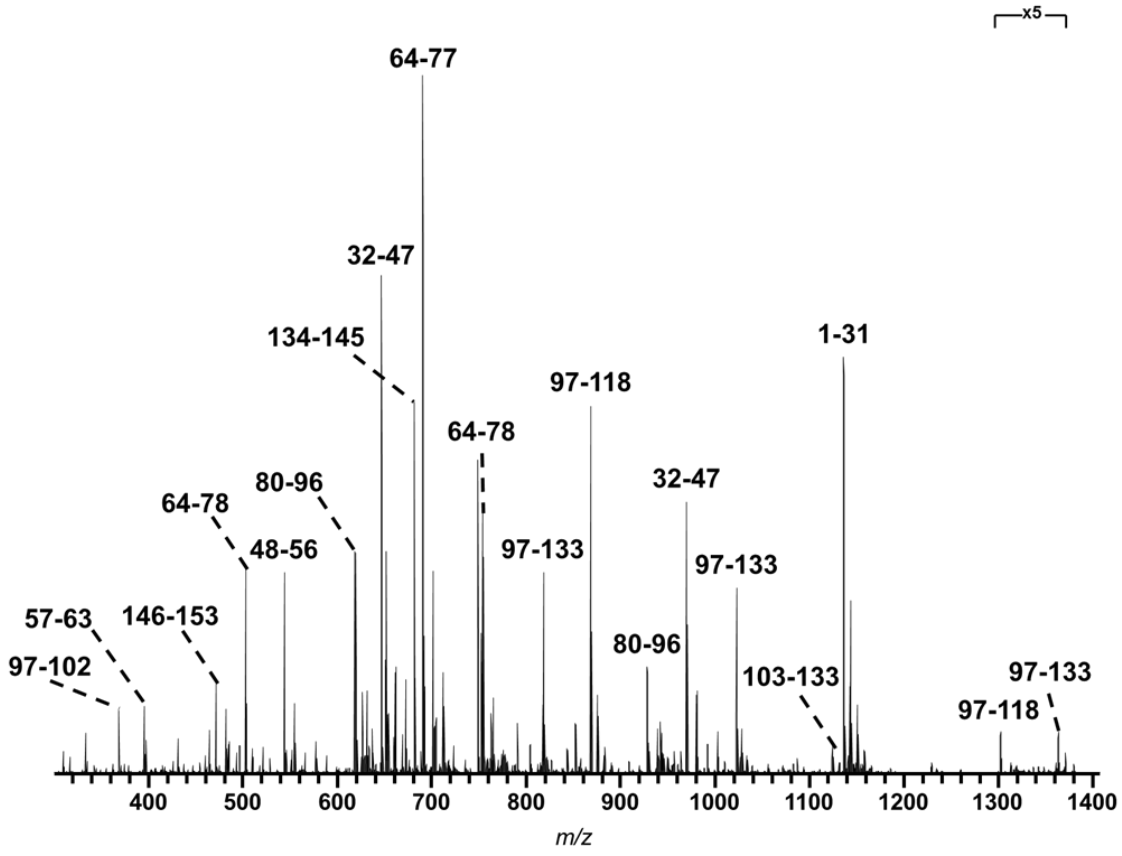


Figure 3.3. Part of the mass spectrum of a tryptic spin digests (500 *g*) of apomyoglobin. Labels show the amino acid (not all of the peptide signals are labelled).

Table 3.1. Apomyoglobin peptides identified from a spin-membrane (spun at 500 *g*) peptic digest.

m/z of $[M+H]^+$	Peptide Sequence	Amino Acids
3134.5542	(-)GLSDGEWQQVLNVWGKVEADIAGHGQE VL(I)	1-29
4651.5069	(L)IRLFTGHPETLEKFDKFKHLKTEAEMKASEDL KKHGTVVL(T)	30-69
8764.8436	(L)IRLFTGHPETLEKFDKFKHLKTEAEMKASEDL KKHGTVVLTALGGILKKKGHHEAELKPLAQSHA TKHKIPIKYLEF(I)	30-106
4132.3546	(L)TALGGILKKKGHHEAELKPLAQSHATKHKIPI KYLEF(I)	70-106
3242.6627	(F)ISDAIIHVLHSKHPGDFGADAQGAMTKALEL(F)	107-137
2927.5196	(D)AIIHVLHSKHPGDFGADAQGAMTKALE L(F)	110-137
1856.9654	(L)FRNDIAAKYKELGFQG(-)	138-153

Table 3.2. Apomyoglobin peptides identified from a spin-membrane (spun at 10,000 *g*) peptic digest.

m/z of $[M+H]^+$	Peptide Sequence	Amino Acids
3134.5542	(-)GLSDGEWQQVLNVWGKVEADIAGHGQE VL(I)	1-29
4651.5069	(L)IRLFTGHPETLEKFDKFKHLKTEAEMKASEDLKK HGTVVL(T)	30-69
8764.8436	(L)IRLFTGHPETLEKFDKFKHLKTEAEMKASEDLKK HGTVVLTALGGILKKKGHHEAELKPLAQSHATKHK IPIKYLEF(I)	30-106
9079.9867	(L)IRLFTGHPETLEKFDKFKHLKTEAEMKASEDLKK HGTVVLTALGGILKKKGHHEAELKPLAQSHATKHK IPIKYLEFISD(A)	30-109
11988.4885	(L)IRLFTGHPETLEKFDKFKHLKTEAEMKASEDLKK HGTVVLTALGGILKKKGHHEAELKPLAQSHATKHK IPIKYLEFISDAIIHVLHSKHPGDFGADAQGAMTKAL EL(F)	30-137
4132.3546	(L)TALGGILKKKGHHEAELKPLAQSHATKHKIPIKY LEF(I)	70-106
3242.6627	(F)ISDAIIHVLHSKHPGDFGADAQGAMTKALEL(F)	107-137
5080.6102	(F)ISDAIIHVLHSKHPGDFGADAQGAMTKALELFRN DIAAKYKELGFQG(-)	107-153
2927.5196	(D)AIIHVLHSKHPGDFGADAQGAMTKALE L(F)	110-137
1856.9654	(L)FRNDIAAKYKELGFQG(-)	138-153

Table 3.3. Apomyoglobin peptides identified from a spin-membrane (spun at 500 *g*) tryptic digest.

m/z of $[M+H]^+$	Peptide Sequence	Amino Acids
1815.9024	(-)GLSDGEWQQVLNVWGK(V)	1-16
3403.7393	(-)GLSDGEWQQVLNVWGKVEADIAGHGQE VLIR(L)	1-31
1606.8547	(K)VEADIAGHGQEVLIR(L)	17-31
1271.663	(R)LFTGHPETLEK(F)	32-42
1937.0167	(R)LFTGHPETLEKFDKFK(H)	32-47
3004.5601	(R)LFTGHPETLEKFDKFKHLKTEAEMK(A)	32-56
1086.5612	(K)HLKTEAEMK(A)	48-56
1857.9739	(K)HLKTEAEMKASEDLKK(H)	48-63
790.4305	(K)ASEDLKK(H)	57-63
1378.8417	(K)HGTVVLTALGGILK(K)	64-77
1506.9366	(K)HGTVVLTALGGILKK(K)	64-78
1635.0316	(K)HGTVVLTALGGILKKK(G)	64-79
2110.1516	(K)KKGHHEAELKPLAQSHATK(H)	78-96
1982.0566	(K)KGHHEAELKPLAQSHATK(H)	79-96
1853.9617	(K)GHHEAELKPLAQSHATK(H)	80-96
735.4876	(K)HKIPIK(Y)	97-102
2601.4915	(K)HKIPIKYLEFISDAIIHVLHSK(H)	97-118
4085.143	(K)HKIPIKYLEFISDAIIHVLHSKHPGDFGADAQG AMTK(A)	97-133
3819.9891	(K)IPIKYLEFISDAIIHVLHSKHPGDFGADAQGAM TK(A)	99-133
1885.0218	(K)YLEFISDAIIHVLHSK(H)	103-118
3368.6732	(K)YLEFISDAIIHVLHSKHPGDFGADAQGAMTK(A)	103-133
1502.6693	(K)HPGDFGADAQGAMTK(A)	119-133
1360.7583	(K)ALELFRNDIAAK(Y)	134-145
2283.2132	(K)ALELFRNDIAAKYKELGFQG(-)	134-153
922.4993	(R)NDIAAKYK(E)	140-147
941.4727	(K)YKELGFQG(-)	146-153

3.3.3 mAb spin digestion with pepsin-containing membranes

Antibodies have unique Y-shape structures that include inter- and intra-chain disulfide bonds.[52] Thus, we used mAbs as model proteins with disulfide bonds. Moreover, enzymatic digestion is crucial for antibody characterization, structure analysis and quality control. We previously reported using a cumbersome homemade setup with a syringe pump for controlled antibody in-

membrane peptic digestion.[43] Here we examine spin digestion using both trypsin- and pepsin-containing membranes of four different therapeutic antibodies.

Proper antibody pretreatment, which is vital to effective, reproducible digestion, is different for pepsin and trypsin. For peptic digestion, we used TCEP as the reducing agent. Unlike DTT, TCEP can reduce disulfide bonds under acidic conditions where pepsin has maximum activity. Moreover, antibodies partially denature at pH 2, and this should increase access to cleavage sites. Acidic conditions also prevent reformation of disulfide bonds after reduction and avoid the need for adding chaotropic and alkylation agents. Different from our previous workflow, just prior to the spin digestion we added a buffer exchange step using 10 kDa molecular weight cutoff membranes. Because of the high concentration of salt in the commercial antibody formulation,[53] desalting is important for downstream MS or LC/MS analysis. In digestion, we employed 30 s for spinning 100 μL of desalted antibody solutions through membranes at 500 *g*, but the time required for the solution to pass through the membrane is actually less than 30 s.

Direct-infusion MS analysis of peptic peptides from Herceptin (He), Avastin (Av), Rituxan (Ri) and Vectibix (Ve) gives 100% sequence coverage for all of the antibodies. Figure 3.4 is the sequence map for Avastin, and Figure 3.5-3.7 give sequence maps for the other antibodies.

Figure 3.4. Sequence map of the peptides identified from infusion ESI-Orbitrap analysis of peptic digest of Avastin. Light green color "N" represents the glycosylation site. Red color "K" represents the C-terminal clipping.

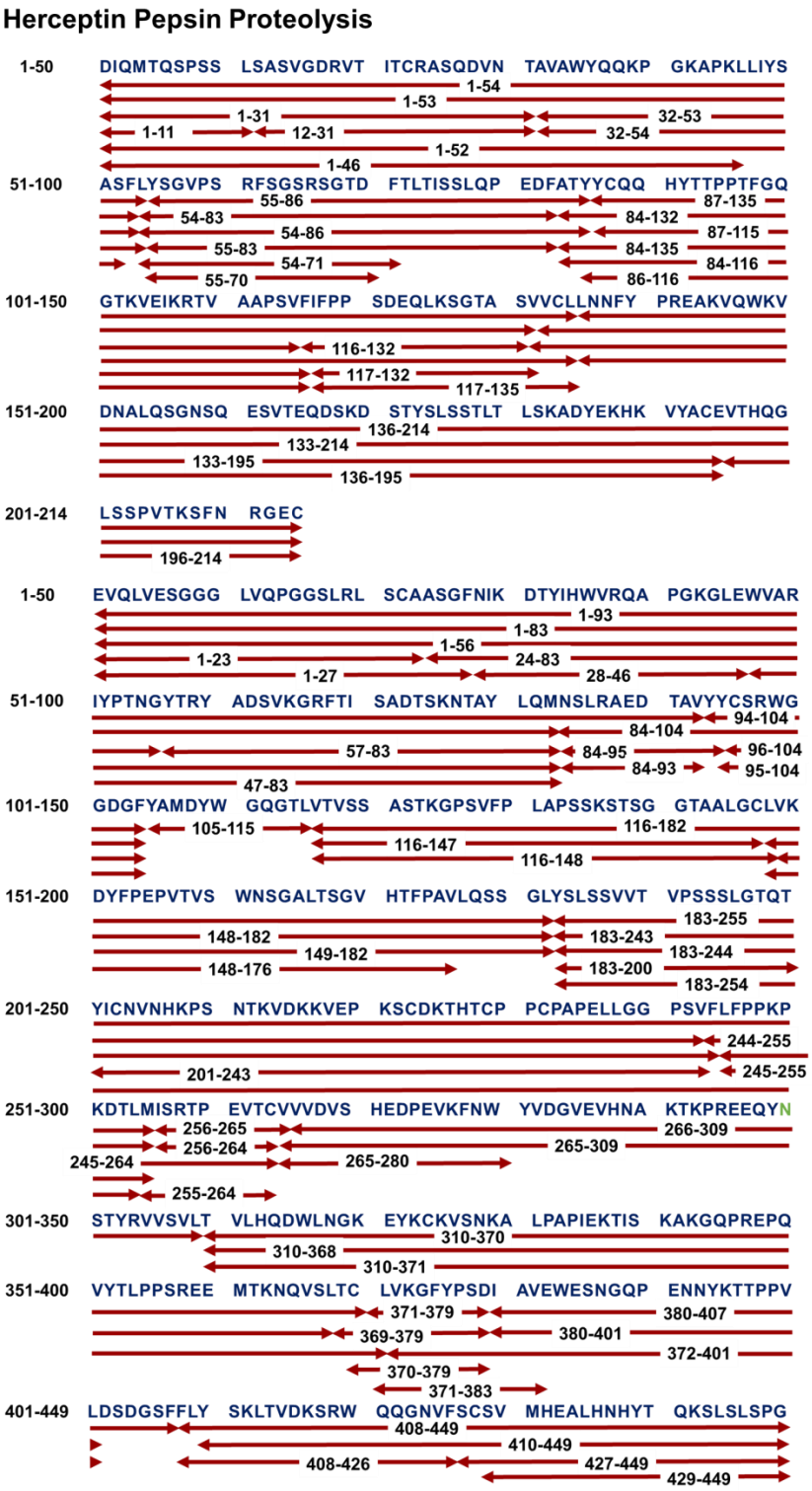


Figure 3.5. Sequence map of the peptides identified from infusion ESI-Orbitrap analysis of peptic digest of Herceptin. Light green color "N" represents the glycosylation site.

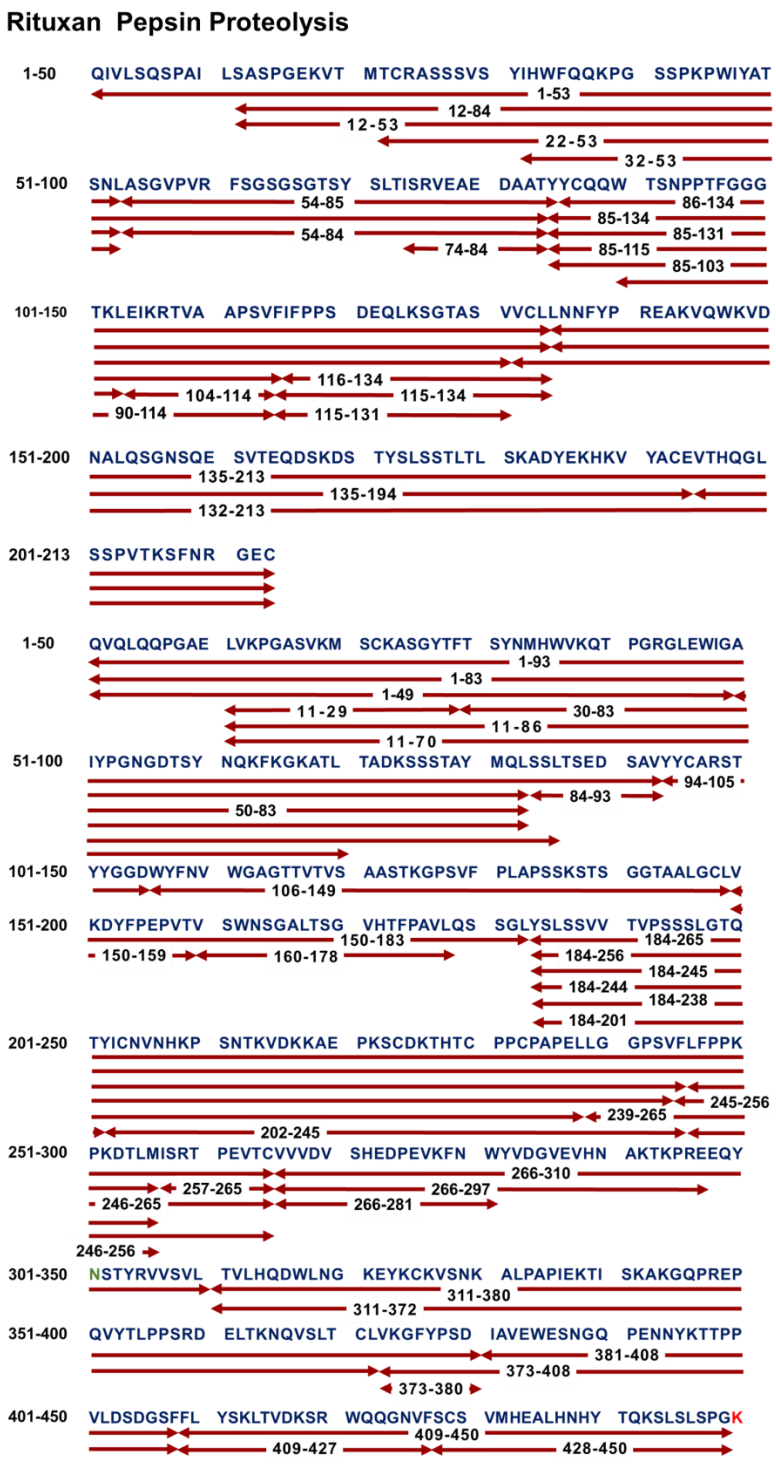


Figure 3.6. Sequence map of the peptides identified from infusion ESI-Orbitrap analysis of peptic digest of Rituxan. Light green color "N" represents the glycosylation site. Red color "K" represents the C-terminal clipping.

Figure 3.7. Sequence map of the peptides identified from infusion ESI-Orbitrap analysis of peptic digest of Vectibix. Light green color "N" represents the glycosylation site. Red color "K" represents the C-terminal clipping.

The average length of the 365 identified peptic peptides from He, Av, Ri and Ve spin digestions is 36 amino acids, indicating that spin digestion at 500 *g* generates middle-down sized peptides that will enable rapid identification of protein posttranslational modifications (PTMs). We identified glycosylation on the heavy chains of all four antibodies. N-terminal pyroglutamate formation appeared on the Ri light chain (Lc), Ri heavy chain (Hc) and Ve Hc. We also saw C-terminal Lysine clipping on Av, Ri and Ve.

Figures 3.8-3.11 present the original mass spectra of the four mAbs, and Tables 3.4-3.7 list the peptic peptides identified for the four antibodies. Some signals with the same *m/z* value are present in the spectra of all four antibodies. For example, signals corresponding to M+H of 8858.2650 result from light chain 136-214 (L136-214) of He, L136-214 of Av, L135-213 of Ri, and L136-214 of Ve. Another peptide with M+H of 4823.3821 stems from heavy chain 408-449 (H408-449) of He, H411-452 of Av, H409-450 of Ri, and H403-444 of Ve. More examples appear in the peptide list in Table 3.4-3.7, as we would expect because the antibodies He, Av, Ri and Ve share the same sequences in a large part of the light-chain and heavy-chain constant regions. The presence of the same peptides in spectra of four antibodies shows that rapid spin digestion is a powerful method for comparing proteins with similar sequences (see chapter 4). Signals that are present in one mass spectrum but not another give hints for the parts that are different in two proteins.

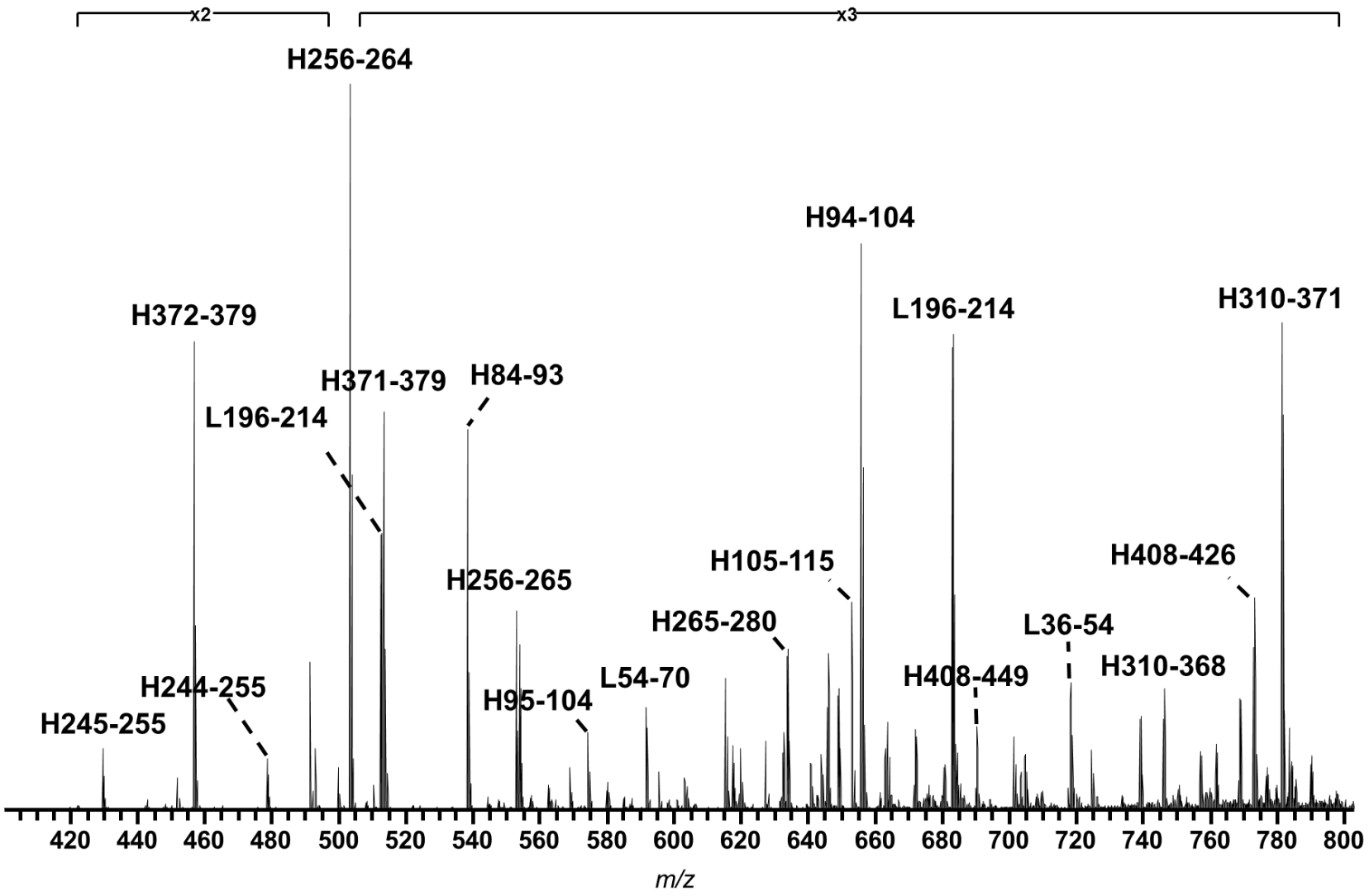


Figure 3.8. Part of the mass spectrum of a peptic spin digest of Herceptin. Labels show the amino acids on the Hc and Lc. (not all of the peptide signals are labelled).

Figure 3.8 (cont'd)

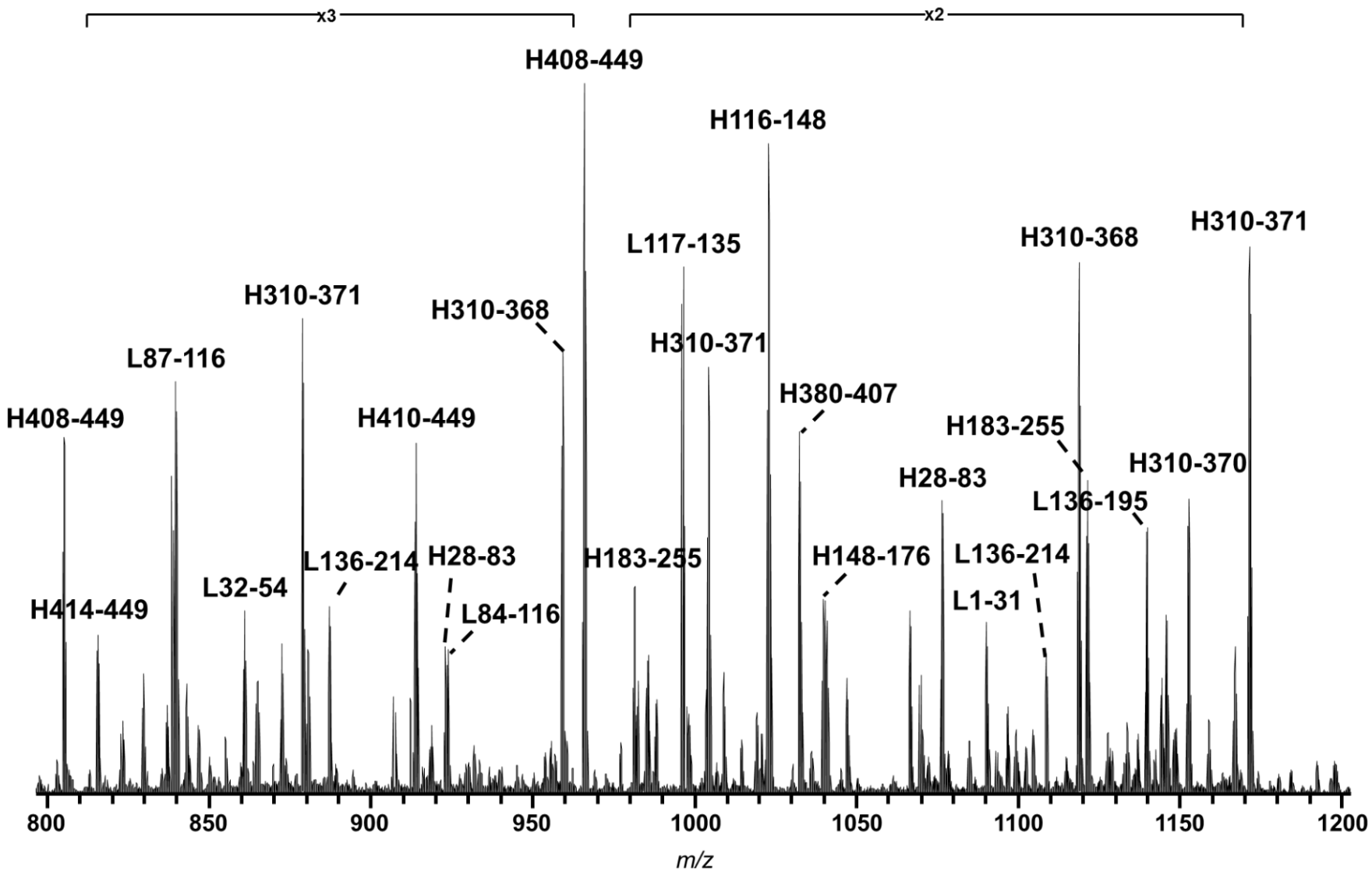
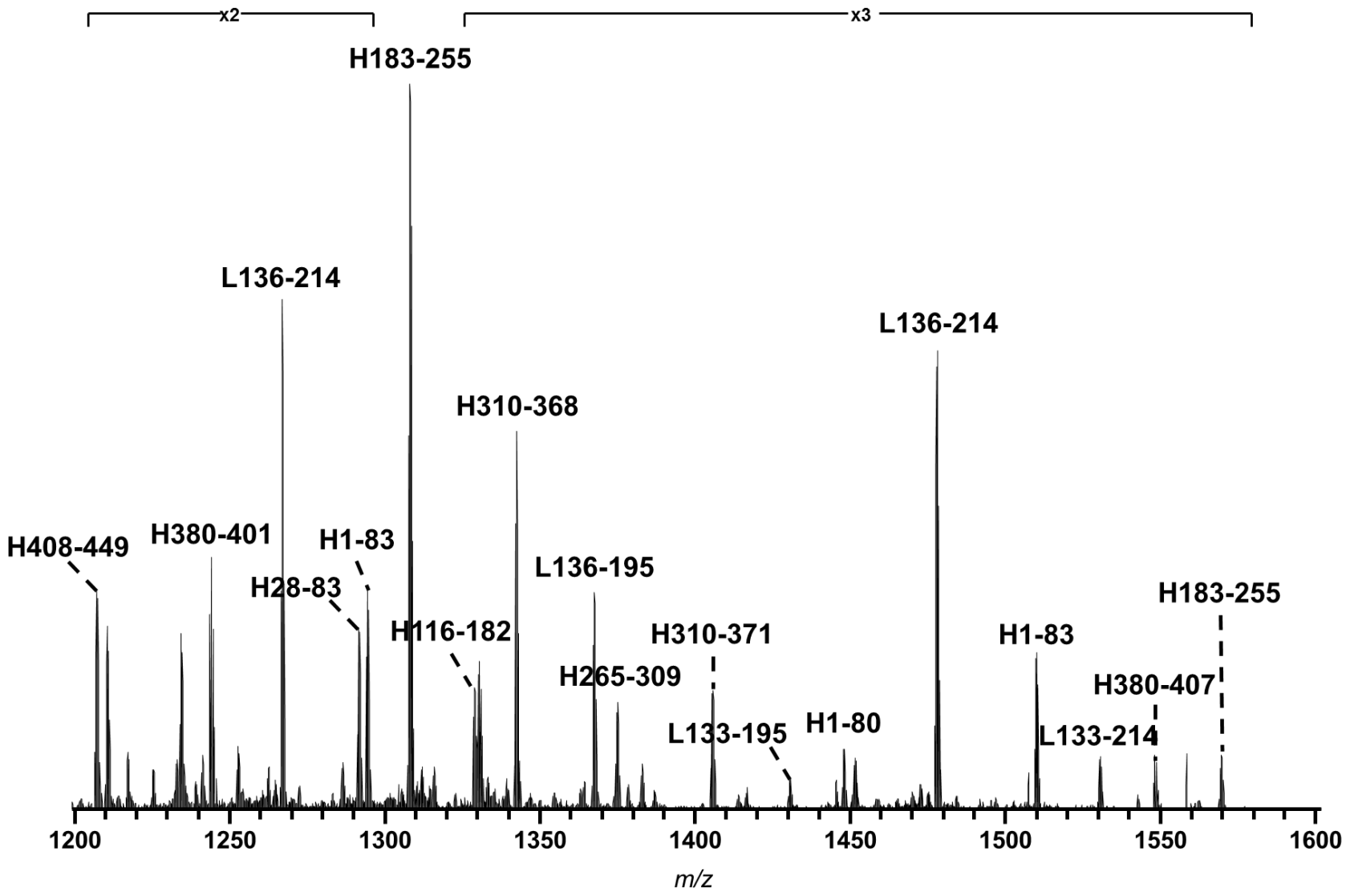


Figure 3.8 (cont'd)



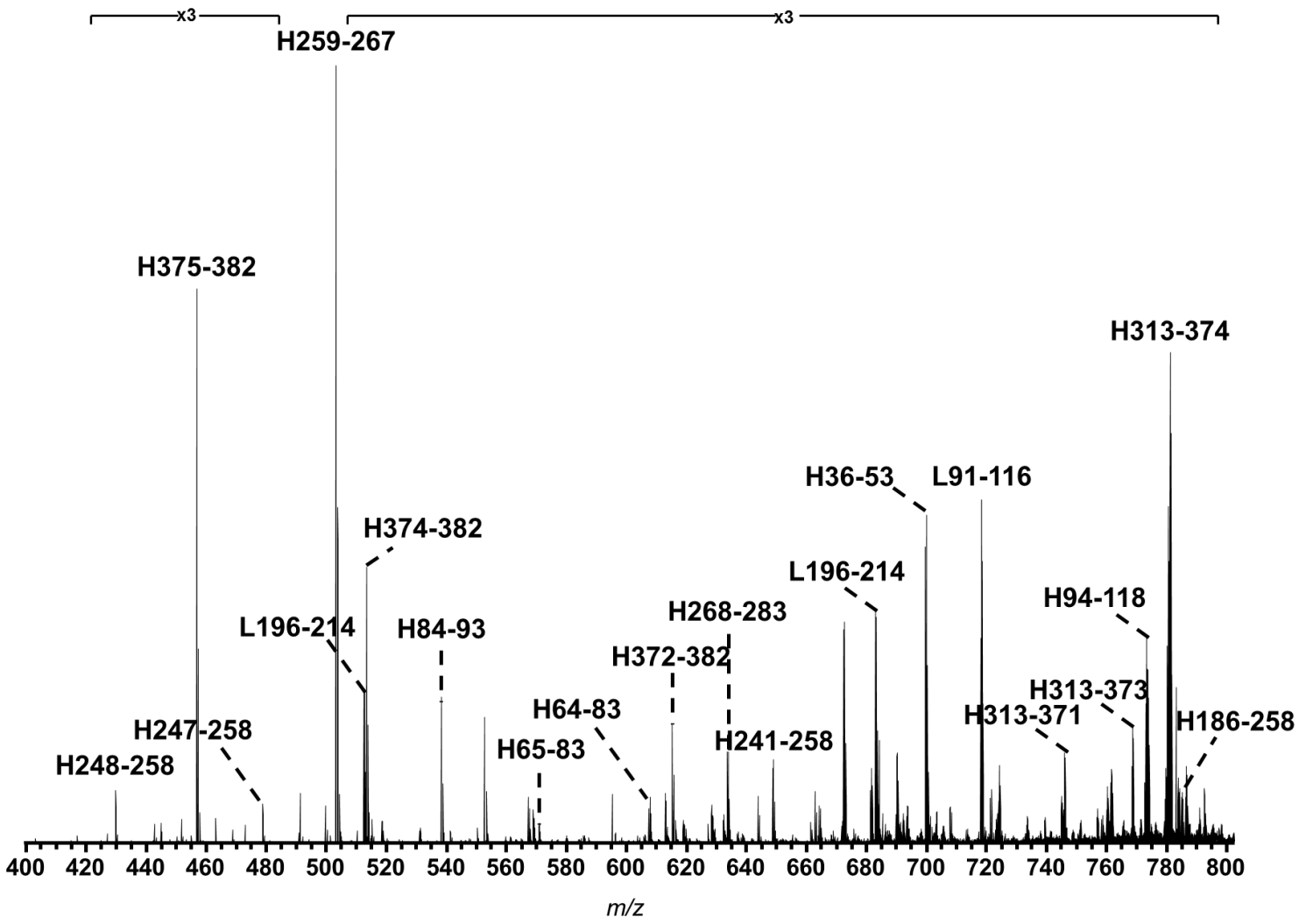


Figure 3.9. Part of the mass spectrum of a peptic spin digest of Avastin. Labels show the amino acids on the Hc and Lc. (not all of the peptide signals are labelled).

Figure 3.9 (cont'd)

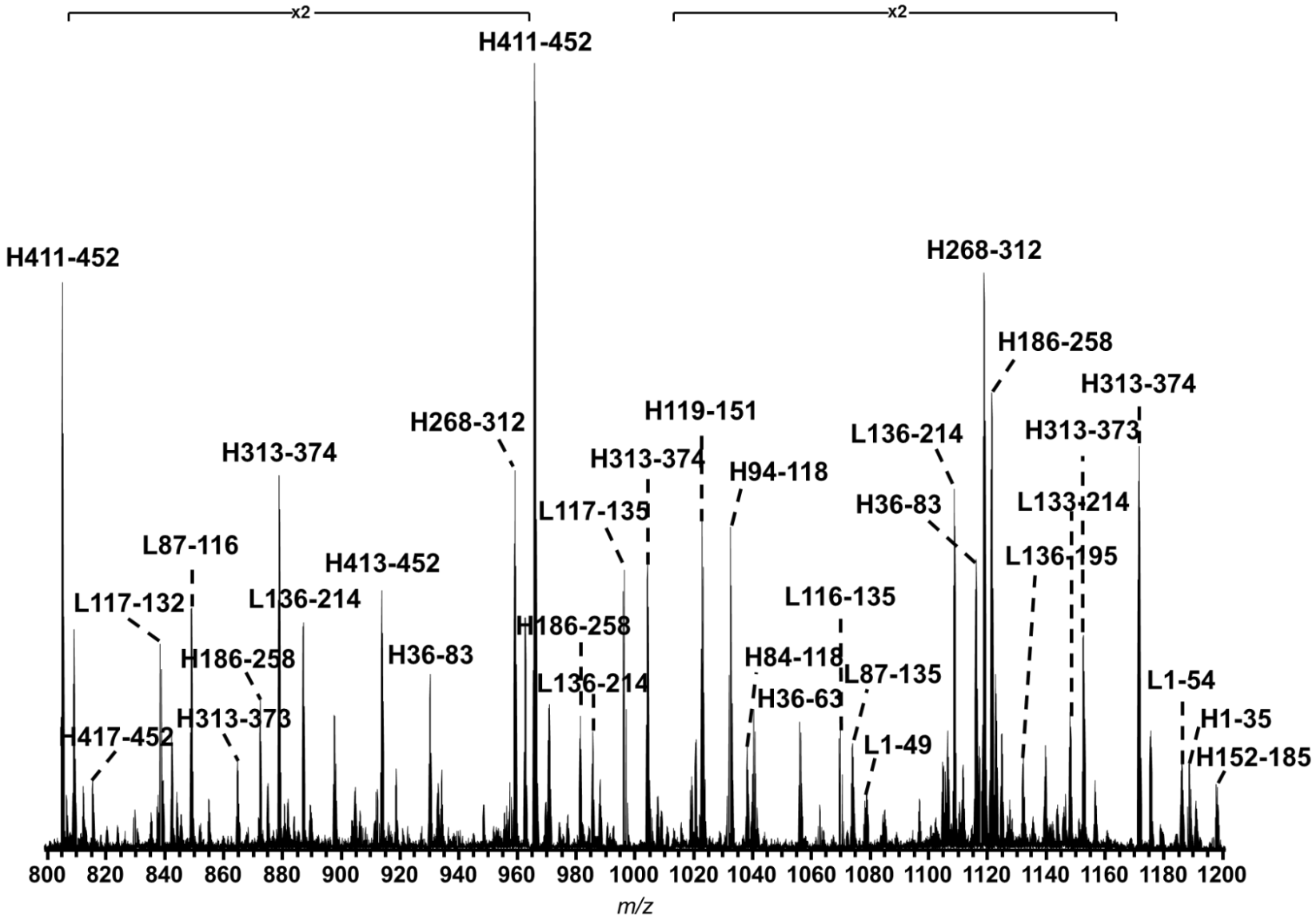
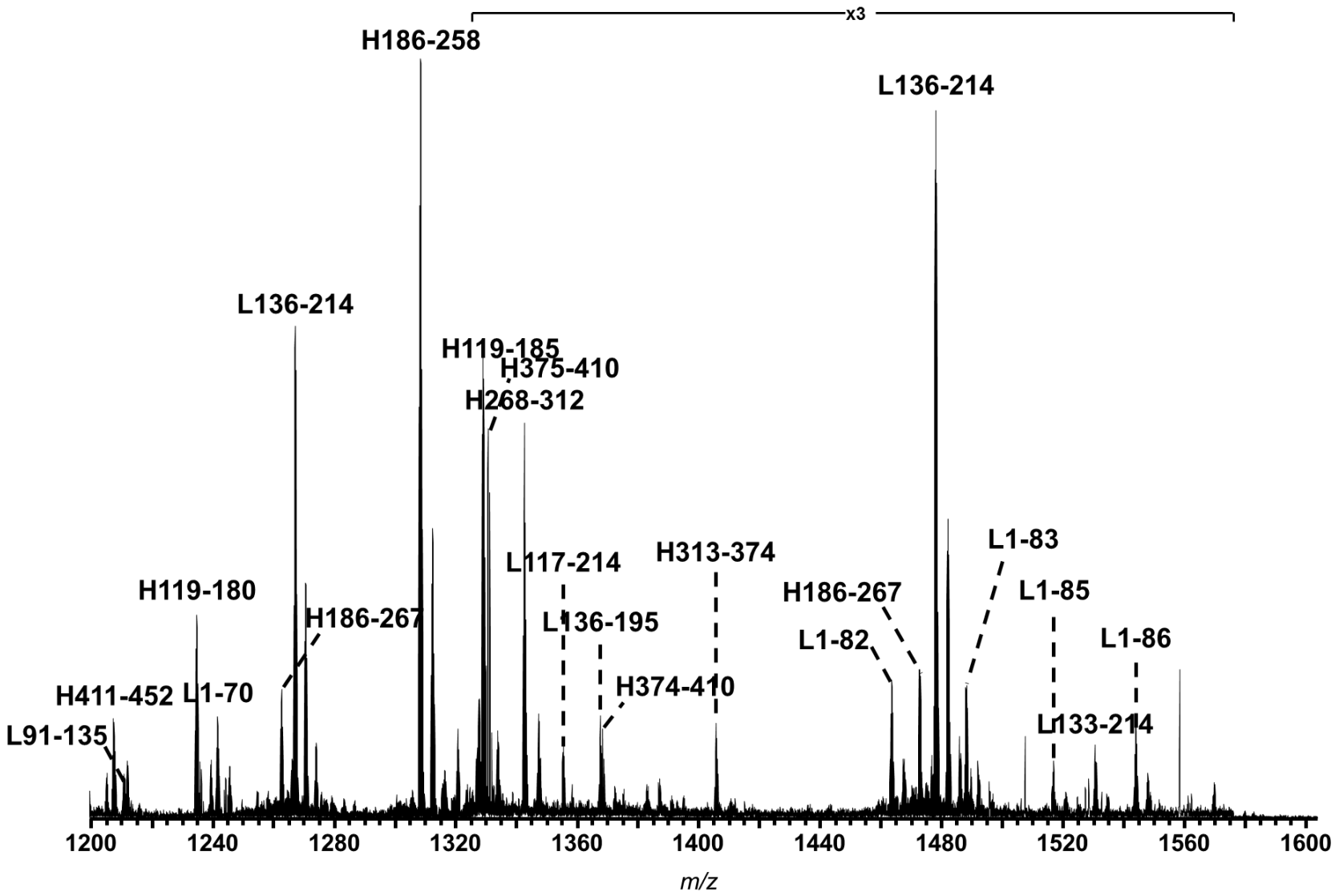


Figure 3.9 (cont'd)



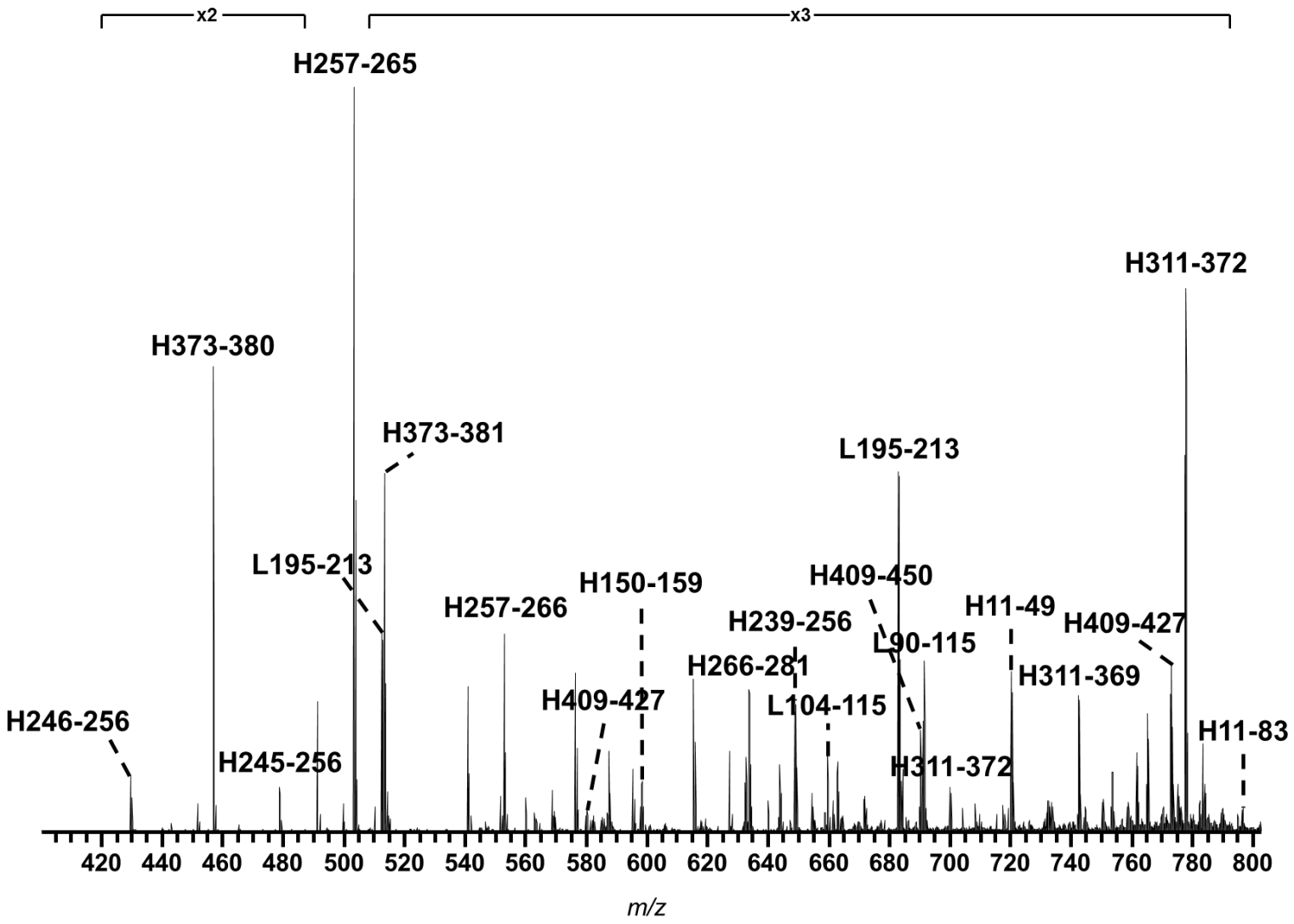


Figure 3.10. Part of the mass spectrum of a peptic spin digest of Rituxan. Labels show the amino acids on the Hc and Lc. (not all of the peptide signals are labelled).

Figure 3.10 (cont'd)

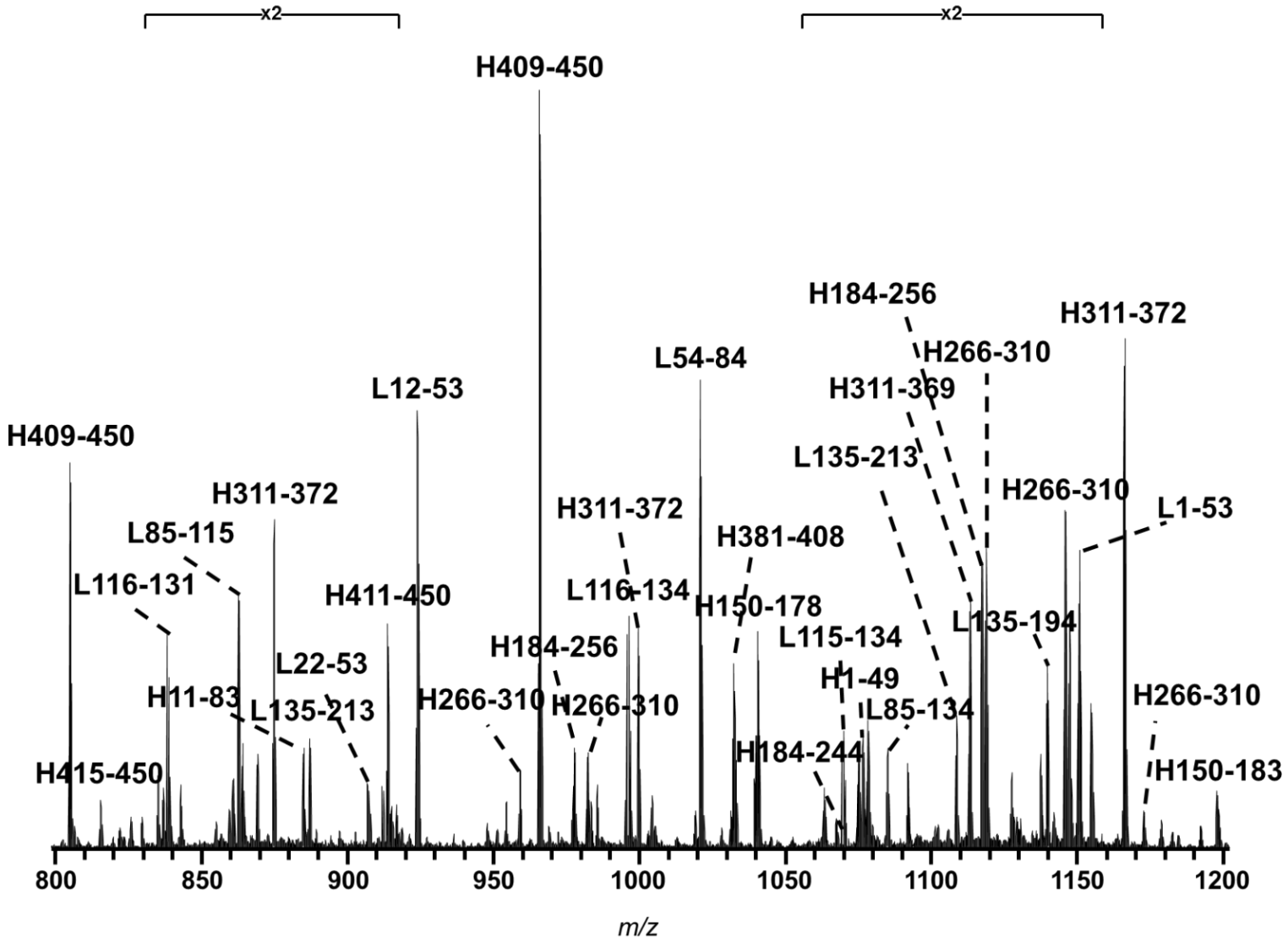
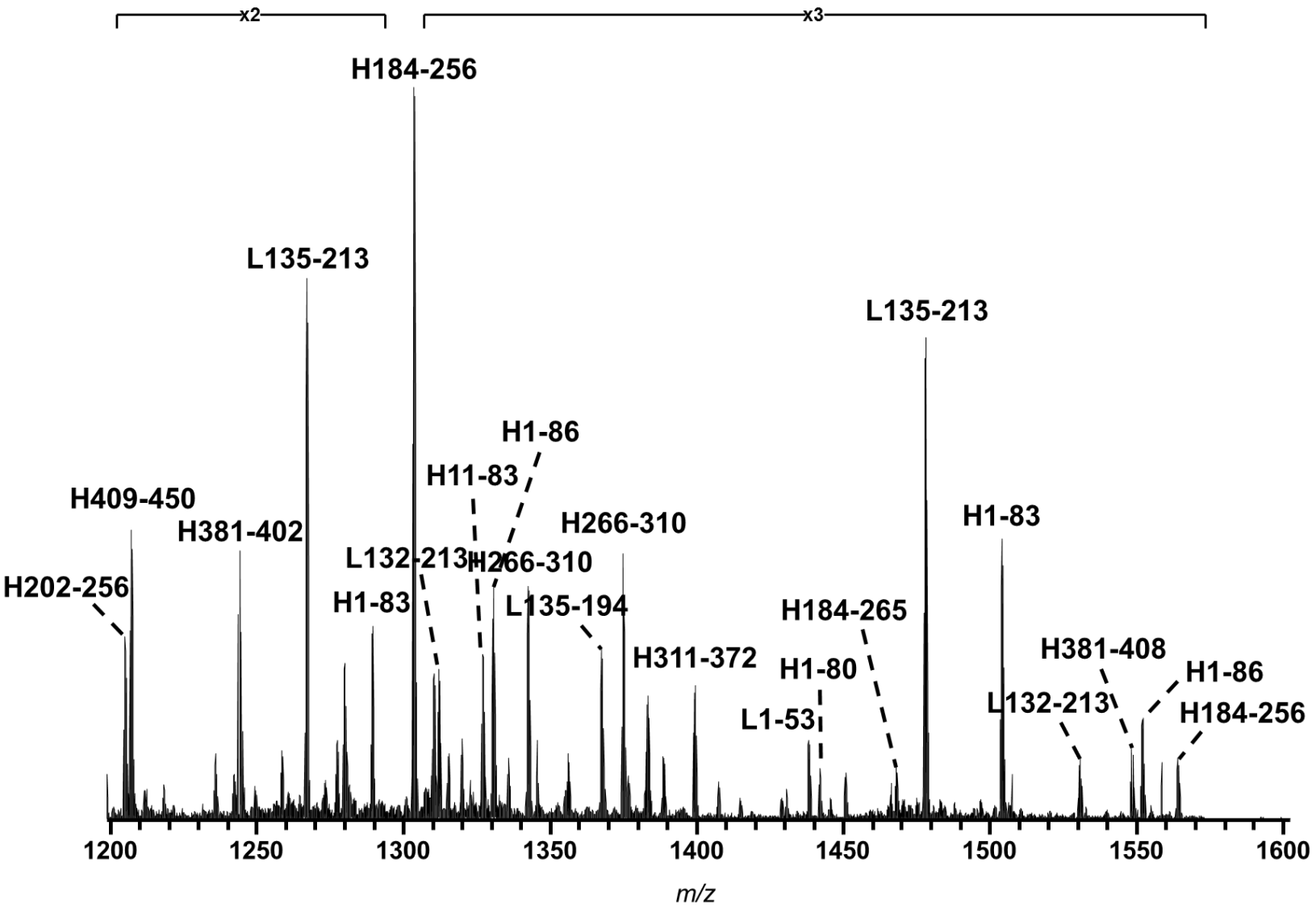


Figure 3.10 (cont'd)



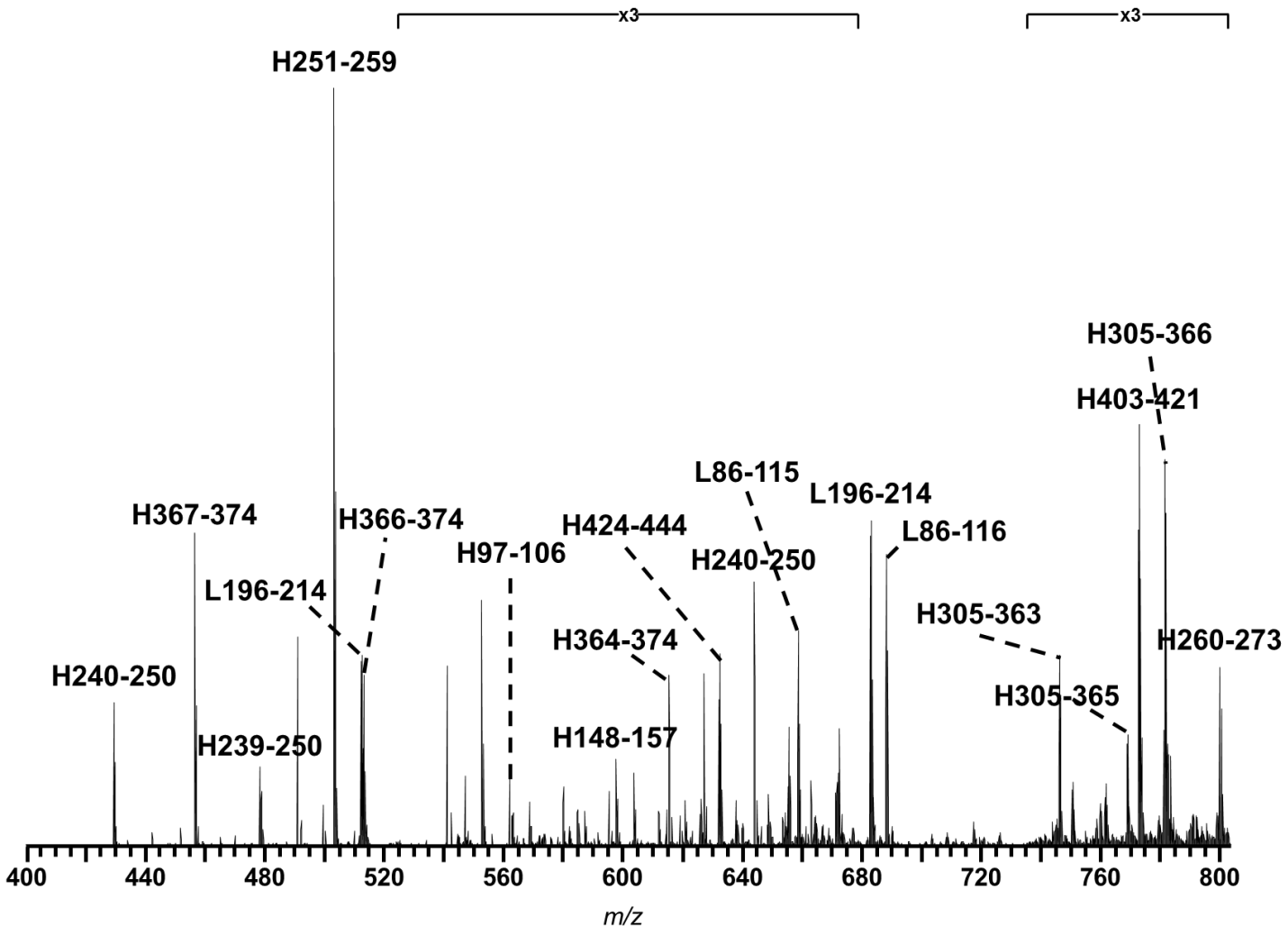


Figure 3.11. Part of the mass spectrum of a peptic spin digest of Vectibix. Labels show the amino acids on the Hc and Lc. (not all of the peptide signals are labelled).

Figure 3.11 (cont'd)

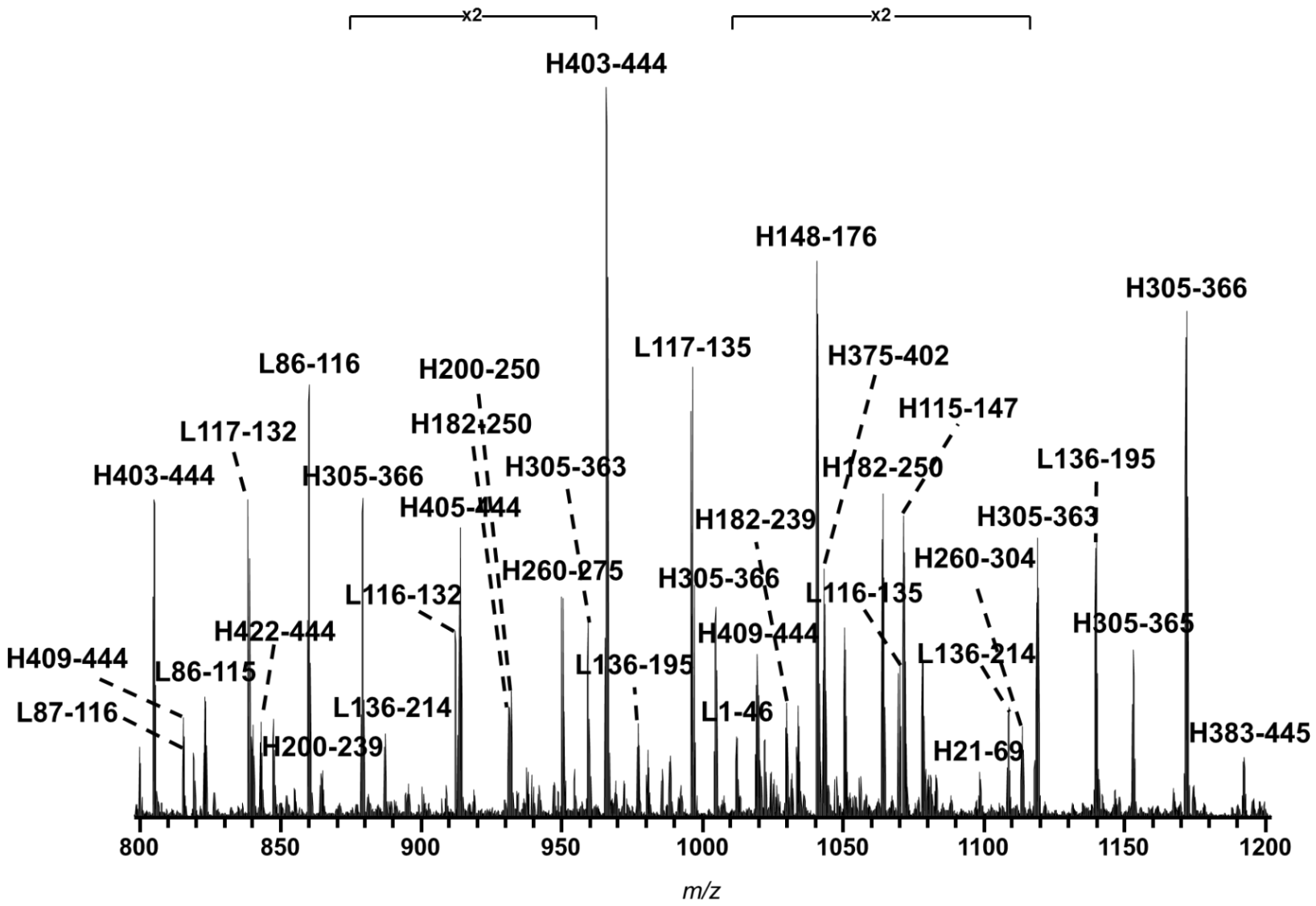


Figure 3.11 (cont'd)

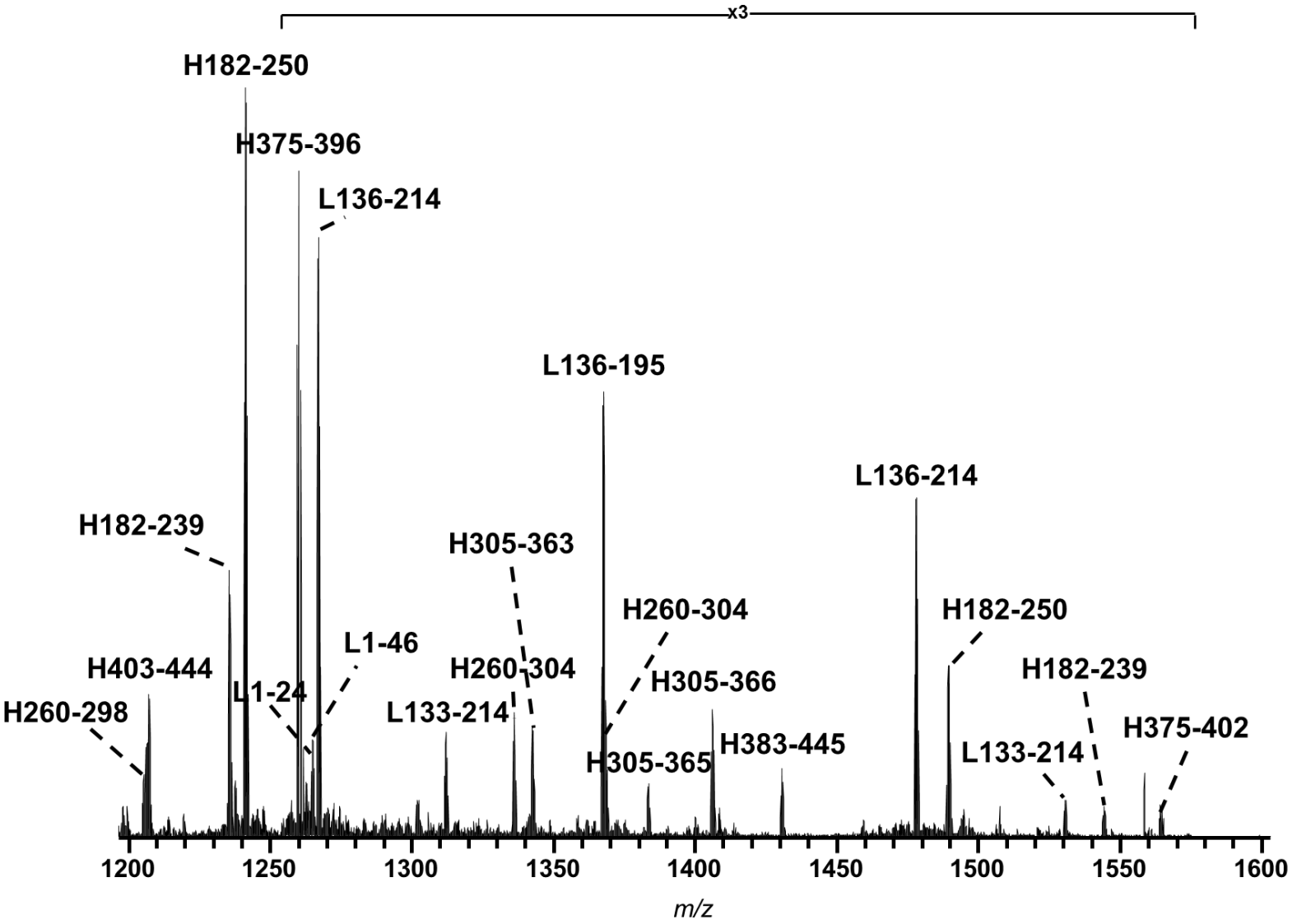


Table 3.4. Light- and heavy-chain peptides identified from a spin-membrane (spun at 500 *g*) peptic digest of Herceptin.

m/z of [M+H]+	Peptide Sequence	Amino Acids
1206.5671	(-)DIQMTQSPSSL(S)	L1-11
3267.558	(-)DIQMTQSPSSLSASVGDRVTITCRASQDVNT(A)	L1-31
3694.7799	(-)DIQMTQSPSSLSASVGDRVTITCRASQDVNTAVA W(Y)	L1-35
4933.4935	(-)DIQMTQSPSSLSASVGDRVTITCRASQDVNTAVA WYQQKPGKAPKL(L)	L1-46
5567.8261	(-)DIQMTQSPSSLSASVGDRVTITCRASQDVNTAVA WYQQKPGKAPKLLIYSAS(F)	L1-52
5714.8945	(-)DIQMTQSPSSLSASVGDRVTITCRASQDVNTAVA WYQQKPGKAPKLLIYSASF(L)	L1-53
5827.9786	(-)DIQMTQSPSSLSASVGDRVTITCRASQDVNTAVA WYQQKPGKAPKLLIYSASFL(Y)	L1-54
2080.0088	(L)SASVGDRVTITCRASQDVNT(A)	L12-31
3745.9443	(L)SASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKL(L)	L12-46
1684.9533	(T)AVAWYQQKPGKAPKL(L)	L32-46
2466.3544	(T)AVAWYQQKPGKAPKLLIYSASF(L)	L32-53
2579.4384	(T)AVAWYQQKPGKAPKLLIYSASFL(Y)	L32-54
2039.1324	(W)YQQKPGKAPKLLIYSASF(L)	L36-53
2152.2165	(W)YQQKPGKAPKLLIYSASFL(Y)	L36-54
1772.8562	(F)LYSGVPSRFSGSRSGTD(F)	L54-70
1919.9246	(F)LYSGVPSRFSGSRSGTDF(T)	L54-71
3104.5171	(F)LYSGVPSRFSGSRSGTDFTLTISSLQPED(F)	L54-82
3251.5855	(F)LYSGVPSRFSGSRSGTDFTLTISSLQPEDF(A)	L54-83
3586.7336	(F)LYSGVPSRFSGSRSGTDFTLTISSLQPEDFAT Y(Y)	L54-86
1659.7721	(L)YSGVPSRFSGSRSGTD(F)	L55-70
1806.8406	(L)YSGVPSRFSGSRSGTDF(T)	L55-71
2991.433	(L)YSGVPSRFSGSRSGTDFTLTISSLQPED(F)	L55-82
3138.5014	(L)YSGVPSRFSGSRSGTDFTLTISSLQPEDF(A)	L55-83
3473.6496	(L)YSGVPSRFSGSRSGTDFTLTISSLQPEDFATY(Y)	L55-86
3689.8421	(F)ATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSV F(I)	L84-116
5346.678	(F)ATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPPS DEQLKSGTASV(V)	L84-132
5661.8396	(F)ATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPPS DEQLKSGTASVVCL(L)	L84-135
3517.7573	(T)YYCQQHYTTPPTFGQGTKVEIKRTVAAPSVF(I)	L86-116
5174.5932	(T)YYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPPSDE QLKSGTASV(V)	L86-132
5489.7549	(T)YYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPPSDE QLKSGTASVVCL(L)	L86-135
3207.6255	(Y)YCQQHYTTPPTFGQGTKVEIKRTVAAPSV(F)	L87-115

Table 3.4 (cont'd)

3354.694	(Y)YCQQHYTTPPTFGQGTKVEIKRTVAAPSVF(I)	L87-116
5011.5299	(Y)YCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPPSDEQ LKSGTASV(V)	L87-132
5326.6915	(Y)YCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPPSDEQ LKSGTASVVCL(L)	L87-135
3191.6306	(Y)CQQHYTTPPTFGQGTKVEIKRTVAAPSVF(I)	L88-116
1822.9222	(V)FIFPPSDEQLKSGTASV(V)	L116-132
2138.0838	(V)FIFPPSDEQLKSGTASVVCL(L)	L116-135
1675.8537	(F)IFPPSDEQLKSGTASV(V)	L117-132
1991.0154	(F)IFPPSDEQLKSGTASVVCL(L)	L117-135
7145.4419	(V)VCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDS KDSTYSLSSTLTLSKADYEKHKVYACE(V)	L133-195
9173.4266	(V)VCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDS KDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTK SFNRGEC(-)	L133-214
6830.2803	(L)LNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDS TYSLSSTLTLSKADYEKHKVYACE(V)	L136-195
8858.265	(L)LNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDS TYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFN RGEC(-)	L136-214
2047.0025	(E)VTHQGLSSPVTKSFNRGEC(-)	L196-214
2256.1653	(-)EVQLVESGGGLVQPGGSLRLSCA(A)	H1-23
2618.3243	(-)EVQLVESGGGLVQPGGSLRLSCAASGF(N)	H1-27
5982.0872	(-)EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYI HWVRQAPGKGLEWVARIYPTNG(Y)	H1-56
8678.4066	(-)EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYI HWVRQAPGKGLEWVARIYPTNGYTRYADSVKGRFTISA DTSKNTAY(L)	H1-80
9050.5897	(-)EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYI HWVRQAPGKGLEWVARIYPTNGYTRYADSVKGRFTISA DTSKNTAYLQM(N)	H1-83
10107.1097	(-)EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYI HWVRQAPGKGLEWVARIYPTNGYTRYADSVKGRFTISA DTSKNTAYLQMNSLRAEDTAV(Y)	H1-93
4775.3576	(L)SCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGY TRYAD(S)	H21-62
7074.5207	(L)SCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGY TRYADSVKGRFTISADTSKNTAYLQM(N)	H21-83
6813.4423	(A)ASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRY ADSVKGRFTISADTSKNTAYLQM(N)	H24-83
2225.1826	(F)NIKDTYIHWVRQAPGKGLE(W)	H28-46
2411.2619	(F)NIKDTYIHWVRQAPGKGLEW(V)	H28-47
3382.7807	(F)NIKDTYIHWVRQAPGKGLEWVARIYPTNG(Y)	H28-56
6079.1002	(F)NIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYADSV KGRFTISADTSKNTAY(L)	H28-80
6451.2833	(F)NIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYADSV KGRFTISADTSKNTAYLQM(N)	H28-83

Table 3.4 (cont'd)

7507.8033	(F)NIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYADSV KGRFTISADTSKNTAYLQMNSLRAEDTAV(Y)	H28-93
4245.1186	(E)WVARIYPTNGYTRYADSVKGRFTISADTSKNTAYLQM (N)	H47-83
3087.5204	(G)YTRYADSVKGRFTISADTSKNTAYLQM(N)	H57-83
1447.721	(Y)LQMNSLRAEDTAV(Y)	H81-93
2367.0459	(M)NSLRAEDTAVYYCSRWGGDGF(Y)	H84-104
1075.5378	(M)NSLRAEDTAV(Y)	H84-93
1401.6645	(M)NSLRAEDTAVYY(C)	H84-95
1310.5259	(V)YYCSRWGGDGF(Y)	H94-104
1147.4626	(Y)YCSRWGGDGF(Y)	H95-104
984.3992	(Y)CSRWGGDGF(Y)	H96-104
1304.5616	(F)YAMDYWGQGTL(V)	H105-115
2952.4983	(L)VTVSSASTKGPSVFPLAPSSKSTSGGTAALGC(L)	H116-147
3065.5823	(L)VTVSSASTKGPSVFPLAPSSKSTSGGTAALGC L(V)	H116-148
6637.3811	(L)VTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYF PEPVTVSWNSGALTSGVHTFPAVLQSSGL(Y)	H116-182
3118.5884	(C)LVKDYFPEPVTVSWNSGALTSGVHTFPAV(L)	H148-176
3703.9006	(C)LVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGL(Y)	H148-182
3590.8166	(L)VKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGL(Y)	H149-182
1812.9226	(L)YSLSSVVTVPSSSLGTQT(Y)	H183-200
5914.9122	(L)YSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEP KSCDKTHTCPPCPAPEL(L)	H183-237
6425.1924	(L)YSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEP KSCDKTHTCPPCPAPELLGGPSV(F)	H183-243
6572.2608	(L)YSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEP KSCDKTHTCPPCPAPELLGGPSVF(L)	H183-244
7708.9202	(L)YSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEP KSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTL(M)	H183-254
7839.9607	(L)YSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEP KSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLM(I)	H183-255
4631.2877	(T)YICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPELL GGPSV(F)	H201-243
1944.0663	(L)LGGPSVFLFPPKPKDTLM(I)	H238-255
1433.7861	(V)FLFPPKPKDTLM(I)	H244-255
1286.7177	(F)LFPPKPKDTLM(I)	H245-255
2273.2032	(F)LFPPKPKDTLMISRTPEVTC(V)	H245-264
1136.5438	(L)MISRTPEVTC(V)	H255-264
1005.5034	(M)ISRTPEVTC(V)	H256-264
1104.5718	(M)ISRTPEVTCV(V)	H256-265
1898.9283	(C)VVVDVSHEDPEVKFNW(Y)	H265-280
6705.1696	(C)VVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQY NSTYRVVSVL(T)*	H265-309
6867.2114	(C)VVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQY NSTYRVVSVL(T)**	H265-309

Table 3.4 (cont'd)

1799.8599	(V)VVDVSHEDPEVKFNW(Y)	H266-280
6606.0918	(V)VVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVL(T)***	H266-309
6703.5491	(L)TVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSL(T)	H310-368
6907.606	(L)TVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTC(L)	H310-370
7020.69	(L)TVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCL(V)	H310-371
1229.5871	(L)TCLVKGFYPSD(I)	H369-379
1128.5394	(T)CLVKGFYPSD(I)	H370-379
1025.5302	(C)LVKGFYPSD(I)	H371-379
1437.7624	(C)LVKGFYPSDIAVE(W)	H371-383
912.4462	(L)VKGFYPSD(I)	H372-379
1324.6783	(L)VKGFYPSDIAVE(W)	H372-383
3379.6481	(L)VKGFYPSDIAVEWESNGQPENNYKTTPPVL(D)	H372-401
2486.2198	(D)IAVEWESNGQPENNYKTTPPVL(D)	H380-401
3094.4276	(D)IAVEWESNGQPENNYKTTPPVLDSDGSF(F)	H380-407
2316.2135	(F)FLYSKLTVDKSRWQQGNVF(S)	H408-426
4823.3821	(F)FLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG(-)	H408-449
4563.2296	(L)YSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG(-)	H410-449
2526.1864	(F)SCSVMHEALHNHYTQKSLSLSPG(-)	H427-449
2336.1452	(C)SVMHEALHNHYTQKSLSLSPG(-)	H429-449

*6705.1696 is the monoisotopic mass for H265-309 with G0F glycosylation.
**6867.2114 is the monoisotopic mass for H265-309 with G1F glycosylation.
***6606.0918 is the monoisotopic mass for H266-309 with G0F glycosylation.

Table 3.5. Light- and heavy-chain peptides identified from a spin-membrane (spun at 500 *g*) peptic digest of Avastin.

m/z of [M+H]+	Peptide Sequence	Amino Acids
5388.6879	(-)DIQMTQSPSSLSASVGDRVTITCSASQDISNYLNWYQQKPGKAPKVLIY(F)	L1-49
5923.9521	(-)DIQMTQSPSSLSASVGDRVTITCSASQDISNYLNWYQQKPGKAPKVLIYFTSSL(H)	L1-54
7439.6223	(-)DIQMTQSPSSLSASVGDRVTITCSASQDISNYLNWYQQKPGKAPKVLIYFTSSLHSGVPSRFSGSGSGTD(F)	L1-70

Table 3.5 (cont'd)

8771.2832	(-)DIQMTQSPSSLSASVGDRVTITCSASQDISNYL NWYQQKPGKAPKVLIYFTSSLHSGVPSRFSGSGSGTDF TLTISSLQPED(F)	L1-82
8918.3517	(-)DIQMTQSPSSLSASVGDRVTITCSASQDISNYL NWYQQKPGKAPKVLIYFTSSLHSGVPSRFSGSGSGTDF TLTISSLQPEDF(A)	L1-83
9090.4364	(-)DIQMTQSPSSLSASVGDRVTITCSASQDISNYL NWYQQKPGKAPKVLIYFTSSLHSGVPSRFSGSGSGTDF TLTISSLQPEDFAT(Y)	L1-85
9253.4998	(-)DIQMTQSPSSLSASVGDRVTITCSASQDISNYL NWYQQKPGKAPKVLIYFTSSLHSGVPSRFSGSGSGTDF TLTISSLQPEDFATY(Y)	L1-86
2069.9523	(Y)FTSSLHSGVPSRFSGSGSGTD(F)	L50-70
3013.4174	(L)HSGVPSRFSGSGSGTDFTLTISSLQPEDF(A)	L55-83
3185.5022	(L)HSGVPSRFSGSGSGTDFTLTISSLQPEDFAT(Y)	L55-85
3726.8625	(F)ATYYCQQYSTVPWTFGQGTKVEIKRTVAAPSVF(I)	L84-116
3244.6459	(Y)YCQQYSTVPWTFGQGTKVEIKRTVAAPSV(F)	L87-115
3391.7144	(Y)YCQQYSTVPWTFGQGTKVEIKRTVAAPSVF(I)	L87-116
5048.5503	(Y)YCQQYSTVPWTFGQGTKVEIKRTVAAPSVFIFPPSDE QLKSGTASV(V)	L87-132
5363.7119	(Y)YCQQYSTVPWTFGQGTKVEIKRTVAAPSVFIFPPSDE QLKSGTASVVCL(L)	L87-135
2722.4563	(Q)YSTVPWTFGQGTKVEIKRTVAAPSV(F)	L91-115
2869.5247	(Q)YSTVPWTFGQGTKVEIKRTVAAPSVF(I)	L91-116
4526.3606	(Q)YSTVPWTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKS GTASV(V)	L91-132
4841.5223	(Q)YSTVPWTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKS GTASVVCL(L)	L91-135
1822.9222	(V)FIFPPSDEQLKSGTASV(V)	L116-132
2138.0838	(V)FIFPPSDEQLKSGTASVVCL(L)	L116-135
1675.8537	(F)IFPPSDEQLKSGTASV(V)	L117-132
1991.0154	(F)IFPPSDEQLKSGTASVVCL(L)	L117-135
10830.2625	(F)IFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDN ALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHK VYACEVTHQGLSSPVTKSFNRGEC(-)	L117-214
7145.4419	(V)VCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQD SKDSTYSLSSTLTLSKADYEKHKVYACE(V)	L133-195
9173.4266	(V)VCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQD SKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPV TKSFNRGEC(-)	L133-214
6830.2803	(L)LNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKD STYSLSSTLTLSKADYEKHKVYACE(V)	L136-195
8858.265	(L)LNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKD STYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKS FNRGEC(-)	L136-214
2047.0025	(E)VTHQGLSSPVTKSFNRGEC(-)	L196-214

Table 3.5 (cont'd)

3562.6941	(-)EVQLVESGGGLVQPGGSLRLSCAASGYTFTNY GMN(W)	H1-35
2097.1029	(N)WVRQAPGKGLEWVGWINT(Y)	H36-53
3165.5429	(N)WVRQAPGKGLEWVGWINTYTGEPTYAAD(F)	H36-63
3312.6113	(N)WVRQAPGKGLEWVGWINTYTGEPTYAADF(K)	H36-64
5573.7739	(N)WVRQAPGKGLEWVGWINTYTGEPTYAADFKRRFTF SLDTSKSTAYLQM(N)	H36-83
6630.2939	(N)WVRQAPGKGLEWVGWINTYTGEPTYAADFKRRFTF SLDTSKSTAYLQMNSLRAEDTAV(Y)	H36-93
3495.6889	(T)YTGEPTYAADFKRRFTFSLDTSKSTAYLQM(N)	H54-83
2427.2489	(D)FKRRFTFSLDTSKSTAYLQM(N)	H64-83
2280.1805	(F)KRRFTFSLDTSKSTAYLQM(N)	H65-83
4147.8708	(M)NSLRAEDTAVYYCAKYPHYYGSSHWYFDVWGQGT L(V)	H54-118
1075.5378	(M)NSLRAEDTAV(Y)	H84-93
3091.3508	(V)YYCAKYPHYYGSSHWYFDVWGQGTL(V)	H94-118
3065.5823	(L)VTVSSASTKGPSVFPLAPSSKSTSGGTAALGCL(V)	H119-151
6165.1529	(L)VTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDY FPEPVTVSWNSGALTSGVHTFPAVL(Q)	H119-180
6637.3811	(L)VTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDY FPEPVTVSWNSGALTSGVHTFPAVLQSSGL(Y)	H119-185
3118.5884	(L)VKDYFPEPVTVSWNSGALTSGVHTFPAVL(Q)	H152-180
3590.8166	(L)VKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGL(Y)	H152-185
6425.1924	(L)YSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV EPKSCDKTHTCPPCPAPELLGGPSV(F)	H186-246
6572.2608	(L)YSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV EPKSCDKTHTCPPCPAPELLGGPSVF(L)	H186-247
7839.9607	(L)YSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV EPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLM(I)	H186-258
8826.4462	(L)YSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV EPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISR TPEVTC(V)	H186-267
6046.056	(T)YICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPE LLGGPSVFLFPPKPKDTLM(I)	H204-258
1944.0663	(L)LGGPSVFLFPPKPKDTLM(I)	H241-258
1433.7861	(V)FLFPPKPKDTLM(I)	H247-258
1286.7177	(F)LFPPKPKDTLM(I)	H248-258
2273.2032	(F)LFPPKPKDTLMISRTPEVTC(V)	H248-267
1005.5034	(M)ISRTPEVTC(V)	H259-267
998.4789	(C)VVVDVSHED(P)	H268-276
1898.9283	(C)VVVDVSHEDPEVKFNW(Y)	H268-283
6705.1942	(C)VVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQ YNSTYRVVSVL(T)*	H268-312
6867.244	(C)VVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQ YNSTYRVVSVL(T)**	H268-312

Table 3.5 (cont'd)

5725.726	(D)PEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVL(T)***	H277-312
6703.5491	(L)TVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSL(T)	H313-371
6907.606	(L)TVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTC(L)	H313-373
7020.69	(L)TVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCL(V)	H313-374
7914.1183	(L)TVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSD(I)	H313-382
1229.5871	(L)TCLVKGFYPSD(I)	H372-382
1025.5302	(C)LVKGFYPSD(I)	H374-382
4100.94	(C)LVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSF(F)	H374-410
912.4462	(L)VKGFYPSD(I)	H375-382
3987.8559	(L)VKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSF(F)	H375-410
2486.2198	(D)IAVEWESNGQPENNYKTTPPVL(D)	H383-404
3094.4276	(D)IAVEWESNGQPENNYKTTPPVLDSDGSF(F)	H383-410
2316.2135	(F)FLYSKLTVDKSRWQQGNVF(S)	H411-429
4823.3821	(F)FLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG(K)	H411-452
4563.2296	(L)YSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG(K)	H413-452
4071.9552	(L)TVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG(K)	H417-452
2526.1864	(F)SCSVMHEALHNHYTQKSLSLSPG(K)	H430-452

*6705.1942 is the monoisotopic mass for H268-312 with G0F glycosylation.
**6867.244 is the monoisotopic mass for H268-312 with G1F glycosylation.
***5725.726 is the monoisotopic mass for H277-312 with G0F glycosylation.

Table 3.6. Light- and heavy-chain peptides identified from a spin-membrane (spun at 500 *g*) peptic digest of Rituxan.

m/z of [M+H]+	Peptide Sequence	Amino Acids
5763.9302	(-)QIVLSQSPAILSASPGEKVTMTCRASSSVSYIHWFQQKPGSSPKPWIYATSNL(A)	L1-53
4614.2544	(L)SASPGEKVTMTCRASSSVSYIHWFQQKPGSSPKPWIYATSNL(A)	L12-53
7654.7282	(L)SASPGEKVTMTCRASSSVSYIHWFQQKPGSSPKPWIYATSNLASGVPVRFSGSGSGTSYSLTISRVEAEDAAT(Y)	L12-84
3626.7849	(M)TCRASSSVSYIHWFQQKPGSSPKPWIYATSNL(A)	L22-53
2585.33	(Y)IHWFQQKPGSSPKPWIYATSNL(A)	L32-53

Table 3.6 (cont'd)

3059.4916	(L)ASGVPVRFSGSGSGTSYSLTISRVEAEDAA T(Y)	L54-84
3222.5549	(L)ASGVPVRFSGSGSGTSYSLTISRVEAEDAAT Y(Y)	L54-85
1161.5746	(T)ISRVEAEDAAT(Y)	L74-84
2147.9855	(T)YYCQQWTSNPPTFGGGTKL(E)	L85-103
3446.7202	(T)YYCQQWTSNPPTFGGGTKLEIKRTVAAPSVF(I)	L85-115
5103.5561	(T)YYCQQWTSNPPTFGGGTKLEIKRTVAAPSVFIFPPSD EQLKSGTASV(V)	L85-131
5418.7177	(T)YYCQQWTSNPPTFGGGTKLEIKRTVAAPSVFIFPPSD EQLKSGTASVVCL(L)	L85-134
3283.6568	(Y)YCQQWTSNPPTFGGGTKLEIKRTVAAPSVF(I)	L86-115
5255.6544	(Y)YCQQWTSNPPTFGGGTKLEIKRTVAAPSVFIFPPSDE QLKSGTASVVCL(L)	L86-134
2614.3988	(Q)WTSNPPTFGGGTKLEIKRTVAAPSV(F)	L90-114
2761.4672	(Q)WTSNPPTFGGGTKLEIKRTVAAPSVF(I)	L90-115
4733.4647	(Q)WTSNPPTFGGGTKLEIKRTVAAPSVFIFPPSDEQLKS GTASVVCL(L)	L90-134
1170.6841	(L)EIKRTVAAPSV(F)	L104-114
1317.7525	(L)EIKRTVAAPSVF(I)	L104-115
2974.5884	(L)EIKRTVAAPSVFIFPPSDEQLKSGTASV(V)	L104-131
3289.7501	(L)EIKRTVAAPSVFIFPPSDEQLKSGTASVVCL(L)	L104-134
1822.9222	(V)FIFPPSDEQLKSGTASV(V)	L115-131
2138.0838	(V)FIFPPSDEQLKSGTASVVCL(L)	L115-134
1675.8537	(F)IFPPSDEQLKSGTASV(V)	L116-131
1991.0154	(F)IFPPSDEQLKSGTASVVCL(L)	L116-134
9173.4266	(V)VCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQD SKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPV TKSFNRGEC(-)	L132-213
6830.2803	(L)LNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKD STYSLSSTLTLSKADYEKHKVYACE(V)	L135-194
8858.265	(L)LNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKD STYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKS FNRGEC(-)	L135-213
2047.0025	(E)VTHQGLSSPVTKSFNRGEC(-)	L195-213
5393.702	(-)QVQLQQPGAELVKPGASVKMSCKASGYTFTSY NMHWVKQTPGRGLEWIG(A)	H1-49
7647.8403	(-)QVQLQQPGAELVKPGASVKMSCKASGYTFTSY NMHWVKQTPGRGLEWIGAIYPGNGDTSYNQKFKGKA TL(T)	H1-70
8659.2912	(-)QVQLQQPGAELVKPGASVKMSCKASGYTFTSY NMHWVKQTPGRGLEWIGAIYPGNGDTSYNQKFKGKA TLTADKSSSTAY(M)	H1-80
9031.4743	(-)QVQLQQPGAELVKPGASVKMSCKASGYTFTSY NMHWVKQTPGRGLEWIGAIYPGNGDTSYNQKFKGKA TLTADKSSSTAYMQL(S)	H1-83

Table 3.6 (cont'd)

9318.6224	(-)QVQLQQPGAELVKPGASVKMSCKASGYTFTSY NMHWVKQTPGRGLEWIGAIYPGNGDTSYNQKFKGKA TLTADKSSSTAYMQLSSL(T)	H1-86
10007.9092	(-)QVQLQQPGAELVKPGASVKMSCKASGYTFTSY NMHWVKQTPGRGLEWIGAIYPGNGDTSYNQKFKGKA TLTADKSSSTAYMQLSSLTSEDSAV(Y)	H1-93
1974.0187	(E)LVKPGASVKMSCKASGYTF(T)	H11-29
4315.1613	(E)LVKPGASVKMSCKASGYTFTSYNMHWVKQTPGRG LEWIG(A)	H11-49
6569.2996	(E)LVKPGASVKMSCKASGYTFTSYNMHWVKQTPGRG LEWIGAIYPGNGDTSYNQKFKGKATL(T)	H11-70
7952.9336	(E)LVKPGASVKMSCKASGYTFTSYNMHWVKQTPGRG LEWIGAIYPGNGDTSYNQKFKGKATLTADKSSSTAYM QL(S)	H11-83
8240.0817	(E)LVKPGASVKMSCKASGYTFTSYNMHWVKQTPGRG LEWIGAIYPGNGDTSYNQKFKGKATLTADKSSSTAYM QLSSL(T)	H11-86
5997.9327	(F)TSYNMHWVKQTPGRGLEWIGAIYPGNGDTSYNQKF KGKATLTADKSSSTAYMQL(S)	H30-83
3656.7901	(G)AIYPGNGDTSYNQKFKGKATLTADKSSSTAYMQL(S)	H50-83
995.4528	(L)SSLTSEDSAV(Y)	H84-93
1418.5681	(V)YYCARSTYYGGD(W)	H94-105
4332.1645	(D)WYFNVWGAGTTVTVSAASTKGPSVFPLAPSSKSTSG GTAALGCL(V)	H106-149
1194.6041	(L)VKDYFPEPVT(V)	H150-159
3118.5884	(L)VKDYFPEPVTVSWNSGALTSGVHTFPAVL(Q)	H150-178
3590.8166	(L)VKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGL(Y)	H150-183
1943.0021	(T)VSWNSGALTSGVHTFPAVL(Q)	H160-178
1812.9226	(L)YSLSSVVTVPSSSLGTQT(Y)	H184-201
5886.8809	(L)YSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKA EPKSCDKTHTCPPCPAPEL(L)	H184-238
6397.1611	(L)YSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKA EPKSCDKTHTCPPCPAPELLGGPSV(F)	H184-244
6544.2295	(L)YSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKA EPKSCDKTHTCPPCPAPELLGGPSVF(L)	H184-245
7811.9294	(L)YSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKA EPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLM(I)	H184-256
8798.4149	(L)YSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKA EPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISR TPEVTC(V)	H184-265
4092.9762	(T)YICNVNHKPSNTKVDKKAEPKSCDKTHTCPPCPAPE L(L)	H202-238
4750.3248	(T)YICNVNHKPSNTKVDKKAEPKSCDKTHTCPPCPAPE LLGGPSVF(L)	H202-245
6018.0247	(T)YICNVNHKPSNTKVDKKAEPKSCDKTHTCPPCPAPE LLGGPSVFLFPPKPKDTLM(I)	H202-256

Table 3.6 (cont'd)

1944.0663	(L)LGGPSVFLFPPKPKDTLM(I)	H239-256
2930.5518	(L)LGGPSVFLFPPKPKDTLMISRTPEVTC(V)	H239-265
1433.7861	(V)FLFPPKPKDTLM(I)	H245-256
1286.7177	(F)LFPPKPKDTLM(I)	H246-256
2273.2032	(F)LFPPKPKDTLMISRTPEVTC(V)	H246-265
1005.5034	(M)ISRTPEVTC(V)	H257-265
1104.5718	(M)ISRTPEVTCV(V)	H257-266
1898.9283	(C)VVVDVSHEDPEVKFNW(Y)	H266-281
3721.8609	(C)VVVDVSHEDPEVKFNWYVDGVEVHNAKTKPRE(E)	H266-297
6705.1660	(C)VVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQ YNSTYRVVSVL(T)*	H266-310
6867.2164	(C)VVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQ YNSTYRVVSVL(T)**	H266-310
7029.2668	(C)VVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQ YNSTYRVVSVL(T)***	H266-310
6671.577	(L)TVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG QPREPQVYTLPPSRDELTKNQVSL(T)	H311-369
6988.718	(L)TVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG QPREPQVYTLPPSRDELTKNQVSLTCL(V)	H311-372
7882.1463	(L)TVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG QPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSD(I)	H311-380
912.4462	(L)VKGFYPSD(I)	H373-380
1025.5302	(L)VKGFYPSDI(A)	H373-381
3987.8559	(L)VKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSF(F)	H373-408
2486.2198	(D)IAVEWESNGQPENNYKTTPPVL(D)	H381-402
3094.4276	(D)IAVEWESNGQPENNYKTTPPVLDSDGSF(F)	H381-408
2316.2135	(F)FLYSKLTVDKSRWQQGNVF(S)	H409-427
4823.3821	(F)FLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQK SLSLSPG(K)	H409-450
4563.2296	(L)YSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSL SLSPG(K)	H411-450
4071.9552	(L)TVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP G(K)	H415-450
2526.1864	(F)SCSVMHEALHNHYTQKSLSLSPG(K)	H428-450

*6705.1660 is the monoisotopic mass for H266-310 with G0F glycosylation.
**6867.2164 is the monoisotopic mass for H266-310 with G1F glycosylation.
***7029.2668 is the monoisotopic mass for H266-310 with G2F glycosylation.

Table 3.7. Light- and heavy-chain peptides identified from a spin-membrane (spun at 500 *g*) peptic digest of Vectibix.

m/z of [M+H]+	Peptide Sequence	Amino Acids
1206.5671	(-)DIQMTQSPSSL(S)	L1-11
2524.2018	(-)DIQMTQSPSSLSASVGDRVTITCQ(A)	L1-24
3239.5155	(-)DIQMTQSPSSLSASVGDRVTITCQASQDISN(Y)	L1-31
5054.4986	(-)DIQMTQSPSSLSASVGDRVTITCQASQDISNYL NWYQQKPGKAPKL(L)	L1-46
5830.8691	(-)DIQMTQSPSSLSASVGDRVTITCQASQDISNYL NWYQQKPGKAPKLLIYDASN(L)	L1-53
5943.9532	(-)DIQMTQSPSSLSASVGDRVTITCQASQDISNYL NWYQQKPGKAPKLLIYDASNL(E)	L1-54
9116.4368	(-)DIQMTQSPSSLSASVGDRVTITCQASQDISNYL NWYQQKPGKAPKLLIYDASNLETGVPSRFSGSGSG TDFTFTISSLQPEDIAT(Y)	L1-85
1336.6525	(L)SASVGDRVTITCQ(A)	L12-24
2051.9662	(L)SASVGDRVTITCQASQDISN(Y)	L12-31
4643.3199	(L)SASVGDRVTITCQASQDISNYLNWYQQKPGKAP KLLIYDASN(L)	L12-53
2549.3147	(Q)ASQDISNYLNWYQQKPGKAPKL(L)	L25-46
3325.6852	(Q)ASQDISNYLNWYQQKPGKAPKLLIYDASN(L)	L25-53
1834.001	(N)YLNWYQQKPGKAPKL(L)	L32-46
2723.4555	(N)YLNWYQQKPGKAPKLLIYDASNL(E)	L32-54
3288.6986	(T)YFCQHFDHLPLAFGGGTKVEIKRTVAAPSV(F)	L86-115
3435.7671	(T)YFCQHFDHLPLAFGGGTKVEIKRTVAAPSVF(I)	L86-116
5092.603	(T)YFCQHFDHLPLAFGGGTKVEIKRTVAAPSVFIFPP SDEQLKSGTASV(V)	L86-132
5407.7646	(T)YFCQHFDHLPLAFGGGTKVEIKRTVAAPSVFIFPP SDEQLKSGTASVVCL(L)	L86-135
3125.6353	(Y)FCQHFDHLPLAFGGGTKVEIKRTVAAPSV(F)	L87-115
3272.7037	(Y)FCQHFDHLPLAFGGGTKVEIKRTVAAPSVF(I)	L87-116
1822.9222	(V)FIFPPSDEQLKSGTASV(V)	L116-132
2138.0838	(V)FIFPPSDEQLKSGTASVVCL(L)	L116-135
1675.8537	(F)IFPPSDEQLKSGTASV(V)	L117-132
1991.0154	(F)IFPPSDEQLKSGTASVVCL(L)	L117-135
1877.9313	(I)FPPSDEQLKSGTASVVCL(L)	L118-135
9173.4266	(V)VCLLNNFYPREAKVQWKVDNALQSGNSQESVT EQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQ GLSSPVTKSFNRGEC(-)	L133-214
6830.2803	(L)LNNFYPREAKVQWKVDNALQSGNSQESVTEQD SKDSTYSLSSTLTLSKADYEKHKVYACE(V)	L136-195

Table 3.7 (cont'd)

8858.265	(L)LNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC(-)	L136-214
2047.0025	(E)VTHQGLSSPVTKSFNRGEC(-)	L196-214
2093.1114	(-)QVQLQESGPGLVKPSETLSL(T)	H1-20
1641.8694	(L)QESGPGLVKPSETLSL(T)	H5-20
5485.6335	(L)TCTVSGGSVSSGDYYWTWIRQSPGKGLEWIGHIYYSGNTNYNPSLKSRL(T)	H21-69
1240.642	(L)TISIDTSKTQF(S)	H70-80
1702.8898	(F)SLKLSSVTAADTAIYY(C)	H81-96
1123.5677	(Y)CVRDRVTGAF(D)	H97-106
2011.9477	(Y)CVRDRVTGAFDIWGQGTM(V)	H97-114
2683.3243	(M)VTVSSASTKGPSVFPLAPCSRSTSEST(A)	H115-141
3211.5973	(M)VTVSSASTKGPSVFPLAPCSRSTSESTAALGCL(V)	H115-147
1194.6041	(L)VKDYFPEPVT(V)	H148-157
3118.5884	(L)VKDYFPEPVTVSWNSGALTSGVHTFPAVL(Q)	H148-176
3590.8166	(L)VKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGL(Y)	H148-181
2938.4945	(F)PEPVTVSWNSGALTSGVHTFPAVLQSSGL(Y)	H153-181
1873.9178	(L)YSLSSVVTVPSSNFGTQT(Y)	H182-199
6023.8745	(L)YSLSSVVTVPSSNFGTQTYTCNVDHKPSNTKVDKTVERKCCVECPPCPAPPVAGPSV(F)	H182-238
6170.9429	(L)YSLSSVVTVPSSNFGTQTYTCNVDHKPSNTKVDKTVERKCCVECPPCPAPPVAGPSVF(L)	H182-239
7307.6023	(L)YSLSSVVTVPSSNFGTQTYTCNVDHKPSNTKVDKTVERKCCVECPPCPAPPVAGPSVFLFPPKPKDTL(M)	H182-249
7438.6427	(L)YSLSSVVTVPSSNFGTQTYTCNVDHKPSNTKVDKTVERKCCVECPPCPAPPVAGPSVFLFPPKPKDTLM(I)	H182-250
4703.2183	(F)GTQTYTCNVDHKPSNTKVDKTVERKCCVECPPCPAPPVAGPSVF(L)	H196-239
5970.9182	(F)GTQTYTCNVDHKPSNTKVDKTVERKCCVECPPCPAPPVAGPSVFLFPPKPKDTLM(I)	H196-250
4168.9745	(T)YTCNVDHKPSNTKVDKTVERKCCVECPPCPAPPVAGPSV(F)	H200-238
4316.0429	(T)YTCNVDHKPSNTKVDKTVERKCCVECPPCPAPPVAGPSVF(L)	H200-239
5583.7428	(T)YTCNVDHKPSNTKVDKTVERKCCVECPPCPAPPVAGPSVFLFPPKPKDTLM(I)	H200-250
3948.9227	(C)NVDHKPSNTKVDKTVERKCCVECPPCPAPPVAGPSVF(L)	H203-239

Table 3.7 (cont'd)

5216.6226	(C)NVDHKPSNTKVDKTVERKCCVECPPCPAPPVAGPSVFLFPPKPKDTLM(I)	H203-250
1433.7861	(V)FLFPPKPKDTLM(I)	H239-250
1286.7177	(F)LFPPKPKDTLM(I)	H240-250
2273.2032	(F)LFPPKPKDTLMISRTPEVTC(V)	H240-259
1173.6336	(L)FPPKPKDTLM(I)	H241-250
1136.5438	(L)MISRTPEVTC(V)	H250-259
1005.5034	(M)ISRTPEVTC(V)	H251-259
1598.7697	(C)VVVDVSHEDPEVQF(N)	H260-273
1898.8919	(C)VVVDVSHEDPEVQFNW(Y)	H260-275
6019.7183	(C)VVVDVSHEDPEVQFNWYVDGVEVHNAKTKPREEQFNSTF(R)*	H260-298
6181.7644	(C)VVVDVSHEDPEVQFNWYVDGVEVHNAKTKPREEQFNSTF(R)**	H260-298
6673.1343	(C)VVVDVSHEDPEVQFNWYVDGVEVHNAKTKPREEQFNSTFRVVSVL(T)***	H260-304
6835.1836	(C)VVVDVSHEDPEVQFNWYVDGVEVHNAKTKPREEQFNSTFRVVSVL(T)****	H260-304
4139.8438	(W)YVDGVEVHNAKTKPREEQFNSTF(R)*****	H276-298
4301.8954	(W)YVDGVEVHNAKTKPREEQFNSTF(R)******	H276-298
4793.2674	(W)YVDGVEVHNAKTKPREEQFNSTFRVVSVL(T)******	H276-304
4955.3153	(W)YVDGVEVHNAKTKPREEQFNSTFRVVSVL(T)*******	H276-304
6705.5284	(L)TVVHQDWLNGKEYKCKVSNKGLPAPIEKTISKTKGQPREPQVYTLPPSREEMTKNQVSL(T)	H305-363
6909.5852	(L)TVVHQDWLNGKEYKCKVSNKGLPAPIEKTISKTKGQPREPQVYTLPPSREEMTKNQVSLTC(L)	H305-365
7022.6693	(L)TVVHQDWLNGKEYKCKVSNKGLPAPIEKTISKTKGQPREPQVYTLPPSREEMTKNQVSLTCL(V)	H305-366
1229.5871	(L)TCLVKGFYPSD(I)	H364-374
1025.5302	(C)LVKGFYPSD(I)	H366-374
912.4462	(L)VKGFYPSD(I)	H367-374
2518.1919	(D)IAVEWESNGQPENNYKTTPPML(D)	H375-396
3126.3997	(D)IAVEWESNGQPENNYKTTPPMLDSDGSF(F)	H375-402
2316.2135	(F)FLYSKLTVDKSRWQQGNVF(S)	H403-421
2506.2547	(F)FLYSKLTVDKSRWQQGNVFSC(S)	H403-423
4823.3821	(F)FLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG(K)	H403-444
2056.061	(L)YSKLTVDKSRWQQGNVF(S)	H405-421
4563.2296	(L)YSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG(K)	H405-444

Table 3.7 (cont'd)

4071.9552	(L)TVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG(K)	H409-444
2526.1864	(F)SCSVMHEALHNHYTQKSLSLSPG(K)	H422-444
2336.1452	(C)SVMHEALHNHYTQKSLSLSPG(K)	H424-444

*6019.7183 is the monoisotopic mass for H260-298 with G0F glycosylation.
**6181.7644 is the monoisotopic mass for H260-298 with G1F glycosylation.
***6673.1343 is the monoisotopic mass for H260-304 with G0F glycosylation.
****6835.1836 is the monoisotopic mass for H260-304 with G1F glycosylation.
*****4139.8438 is the monoisotopic mass for H276-298 with G0F glycosylation.
******4301.8954 is the monoisotopic mass for H276-298 with G1F glycosylation.
*******4793.2674 is the monoisotopic mass for H276-304 with G0F glycosylation.
********4955.3153 is the monoisotopic mass for H276-304 with G1F glycosylation.

To test the reproducibility of spin proteolysis, we digested Av three times using three different spin columns. For the twenty highest signals in the mass spectra, standard deviations of the signal intensities (relative to the most intense peak in the spectrum) from triplicate digestion are <6% (Figure 3.12). We did not see signals of intact protein, and gel electrophoresis further confirms complete digestion of the four antibodies (see Figure 3.13).

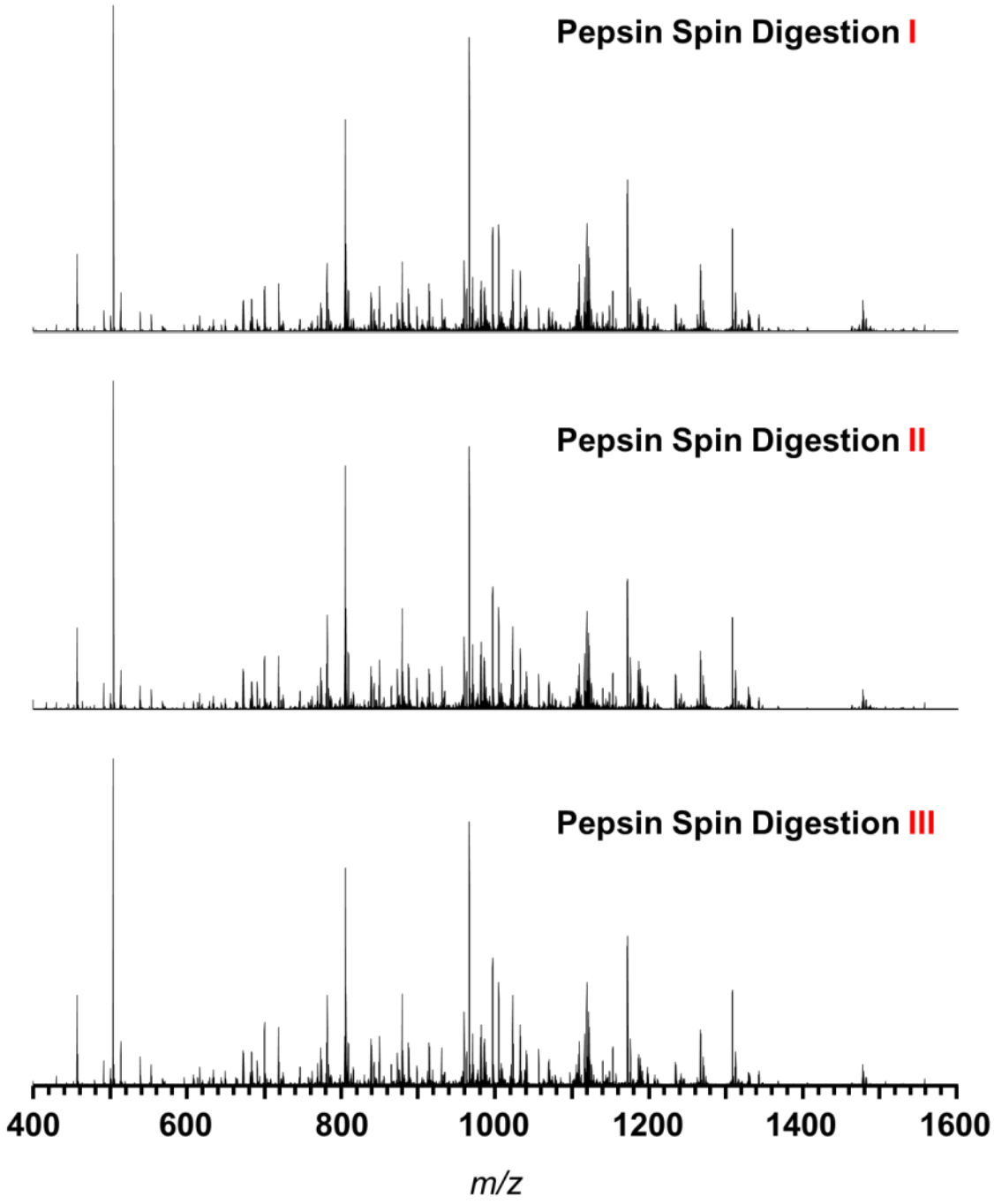


Figure 3.12. Mass spectra of 3 different spin-membrane digests of Avastin. Each digestion employed a different peptic spin column. For the twenty highest signals in the mass spectra, standard deviations of the signals (relative to the most intense peak in the spectrum) from triplicate digestion are <6%.

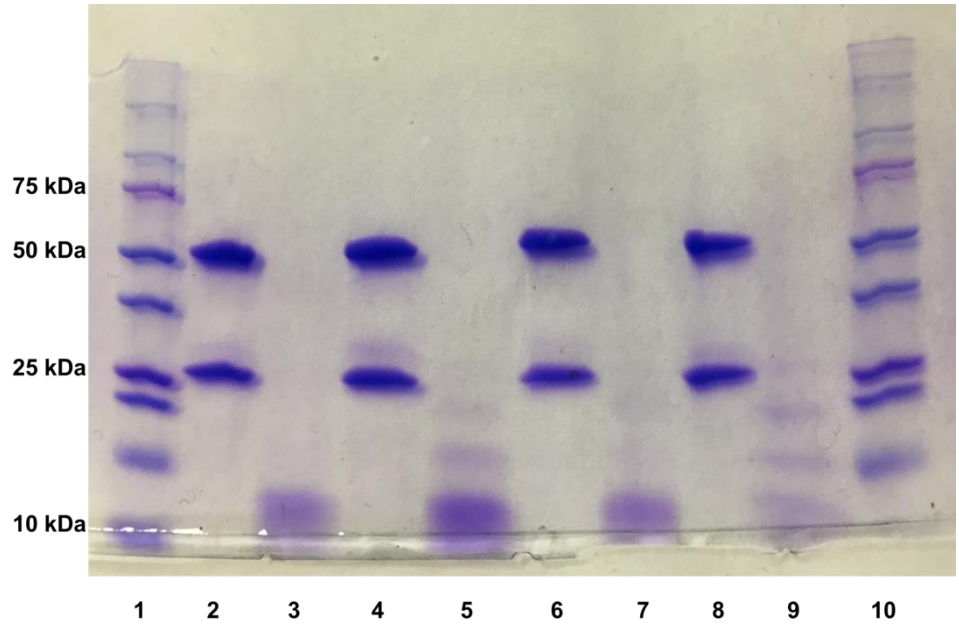


Figure 3.13. Gel electrophoresis (SDS-PAGE) analysis of antibodies before and after digestion in a peptic spin column. Lanes 1 and 10: protein standards; Lanes 2, 4, 6 and 8: 5 µg of Herceptin, Avastin, Rituxan and Vectibix, respectively; Lanes 3, 5, 7 and 9: 5 µg of Herceptin, Avastin, Rituxan and Vectibix peptic spin digests (spun at 500 *g*), respectively.

3.3.4 mAb spin digestion with trypsin-containing membranes

We previously digested an antibody using a trypsin membrane connected to a pipet tip, even without protein alkylation.[46] However, reforming and/or scrambling of the disulfide bonds might occur under basic conditions. In developing a general protein pretreatment, we decided to conduct protein alkylation. Desalting is also necessary because of the large amount of denaturation and alkylation agents. We initially tried to desalt the reduced antibody before trypsin spin digestion. However, when desalting at the protein level using a C4 ZipTip, we did not detect identifiable peptides after spin digestion, presumably because of a large sample loss.

Zhao et al. found that reduced antibodies tend to precipitate during elution with 50% ACN and 0.1% FA.[54]

Considering the instability of reduced antibodies, we performed the desalting step after tryptic spin digestion. We were concerned that the salt and chaotropic agents in the digestion mixture, especially 0.8 M urea, might overcome electrostatic interactions between PSS and trypsin in the digestion membrane. However, this protocol yielded detectable tryptic peptides that cover 100% of the He, Av and Ve sequences, and 84% of the Ri sequence. Figure 3.14 presents the original mass spectrum of Avastin tryptic digestion, and Figure 3.15-3.18 give the sequence map of trypsin digestion of four antibodies. One of the missing pieces from Ri is H68-125, which has a theoretical monoisotopic mass of 6247.8700. The *m/z* of this peptide, which contains two missed cleavage sites, might be outside the mass range we set (300-1800). Another missing piece is H293-305, which contains the glycosylation site. A low ionization efficiency of these glycosylated peptides might explain why they did not give detectable signals in the MS spectrum.

Prior studies indicate that in-solution tryptic digestion generates peptides with average lengths of ~14 amino acids. In contrast, in spin digestion, we found tryptic peptides with up to 10 missed cleavages. Based on 285 tryptic peptides identified from He, Av, Ri and Ve, the average antibody tryptic peptide length is 40 after spin digestion. Limited proteolysis time apparently leads to incomplete peptide digestion, but the larger peptides may facilitate characterization of antibody complementarity determining regions (CDRs). For example, a large tryptic peptide, L1-108 of Ve, covers all the light chain CDRs, which makes characterization of three CDRs possible with a single peptide. As with peptic digestion, spin-membrane tryptic digestion enables identification of PTMs such as glycosylation, N-terminal glutamate formation and C-terminal Lysine clipping. Tables 3.8-3.11 present the full list of identified tryptic peptides.

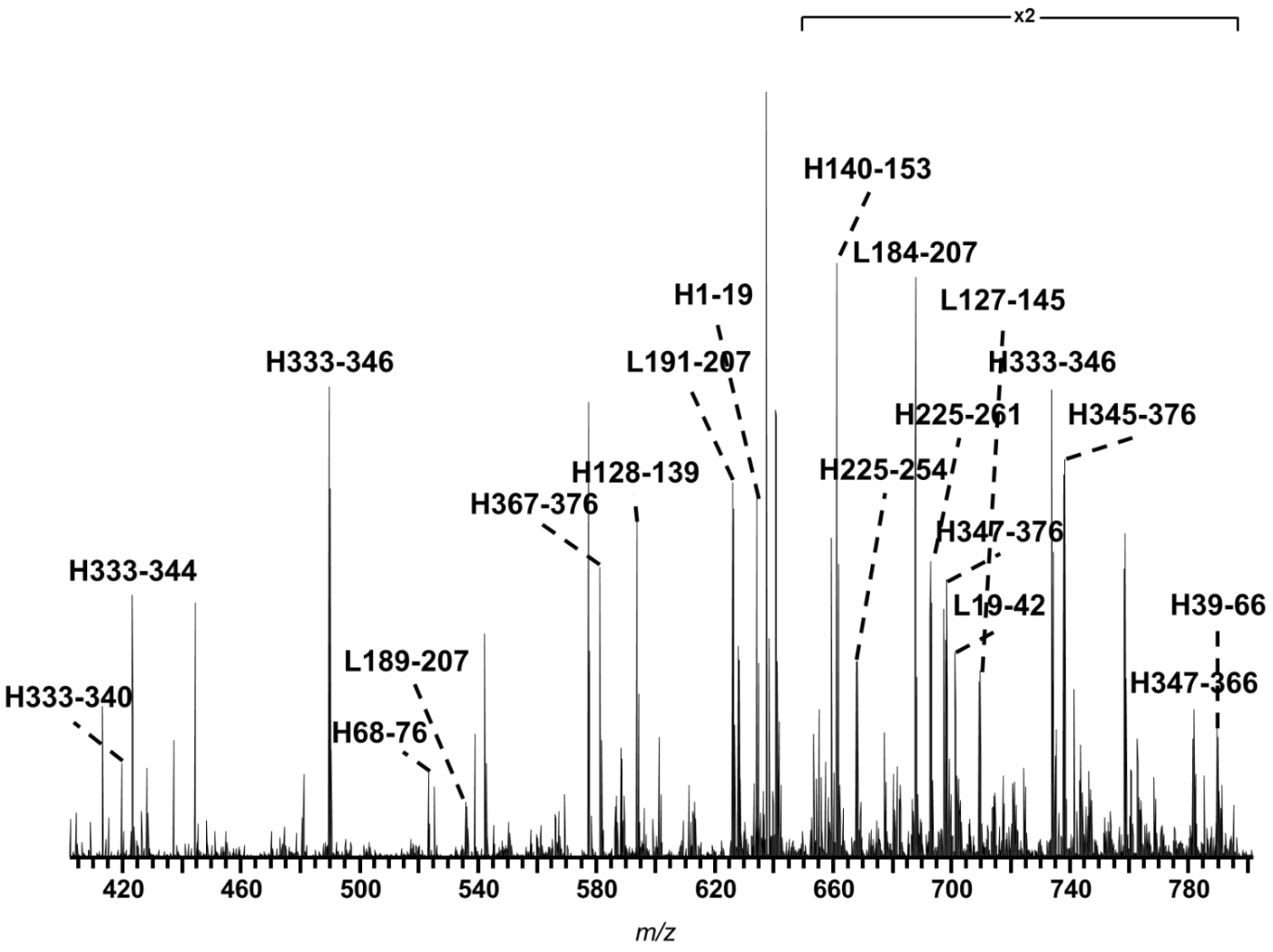


Figure 3.14. Part of the mass spectrum of a tryptic spin digest of Avastin. Labels show the amino acids on the Hc and Lc. (not all of the peptide signals are labelled).

Figure 3.14 (cont'd)

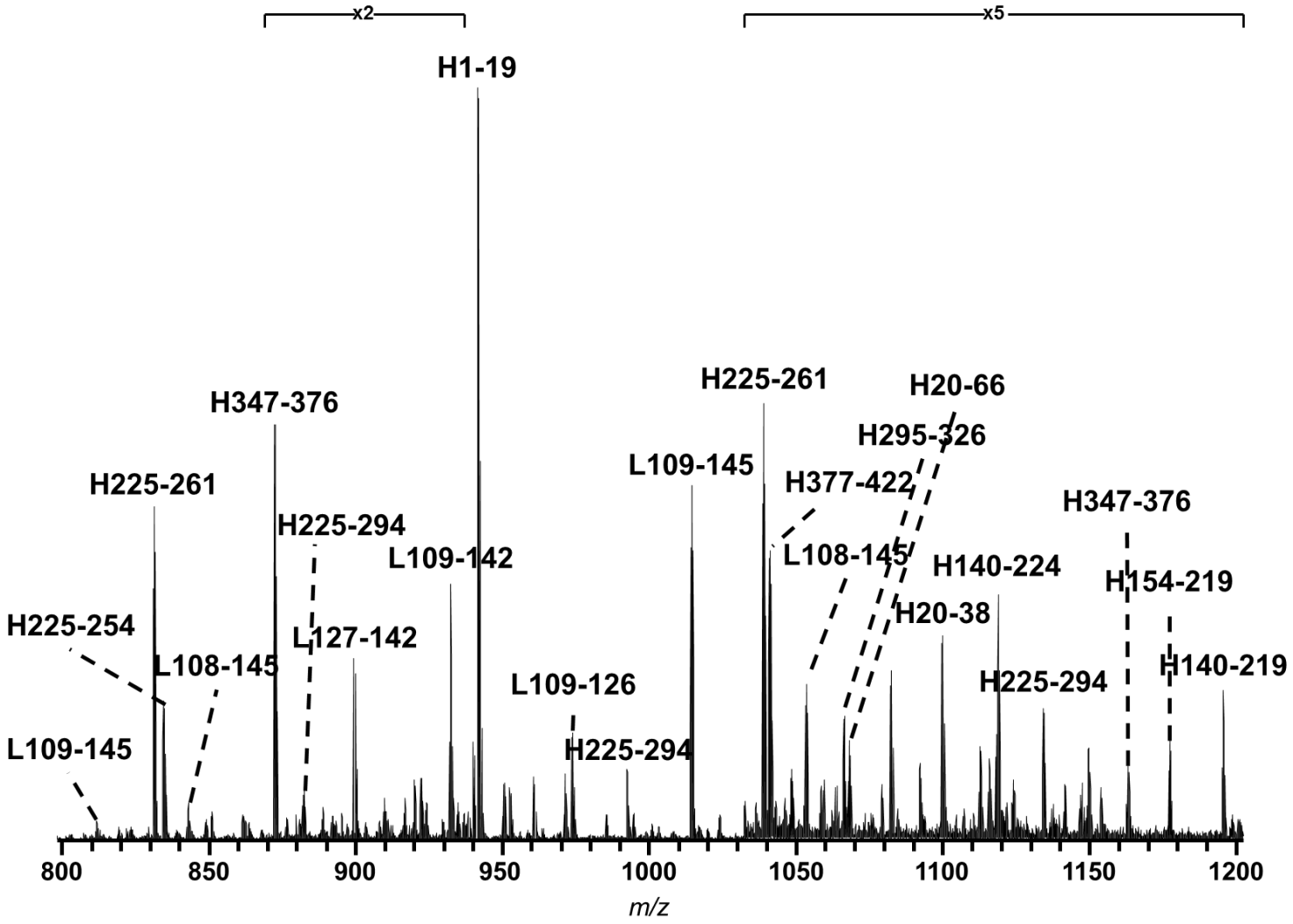
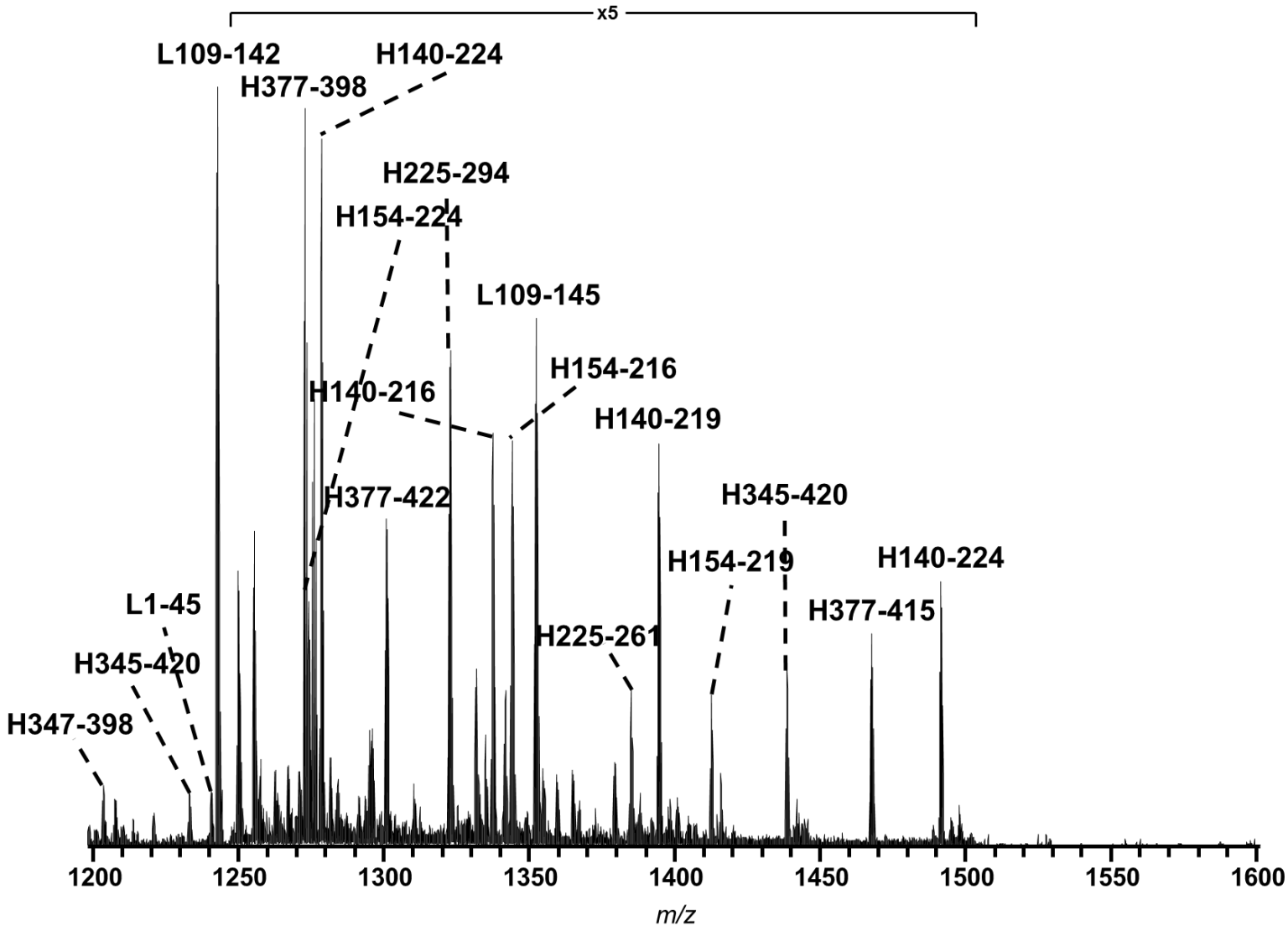


Figure 3.14 (cont'd)



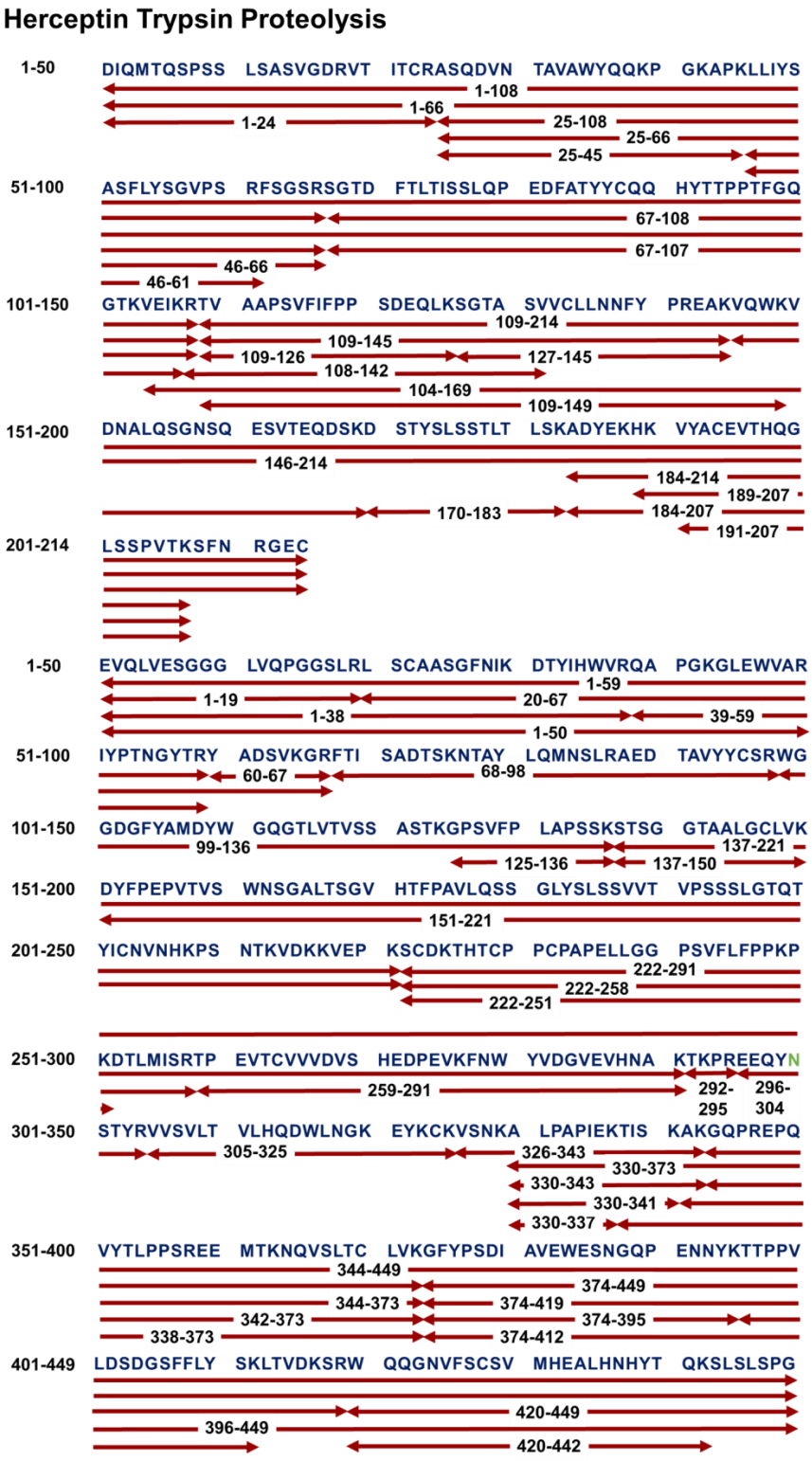


Figure 3.15. Sequence map of the peptides identified from infusion ESI-Orbitrap analysis of a tryptic digest of Herceptin. Light green color "N" represents the glycosylation site.

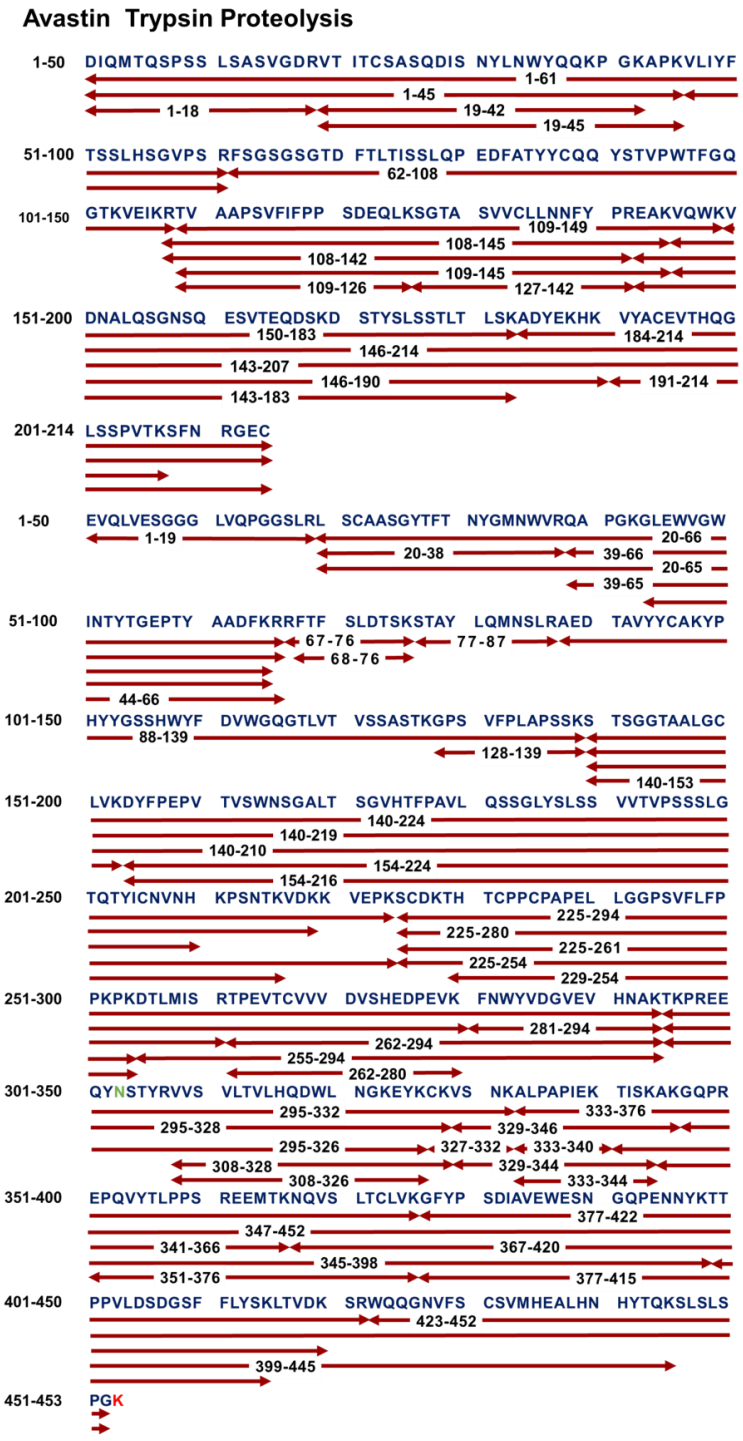


Figure 3.16. Sequence map of the peptides identified from infusion ESI-Orbitrap analysis of a tryptic digest of Avastin. Light green color "N" represents the glycosylation site. Red color "K" represents the C-terminal clipping.

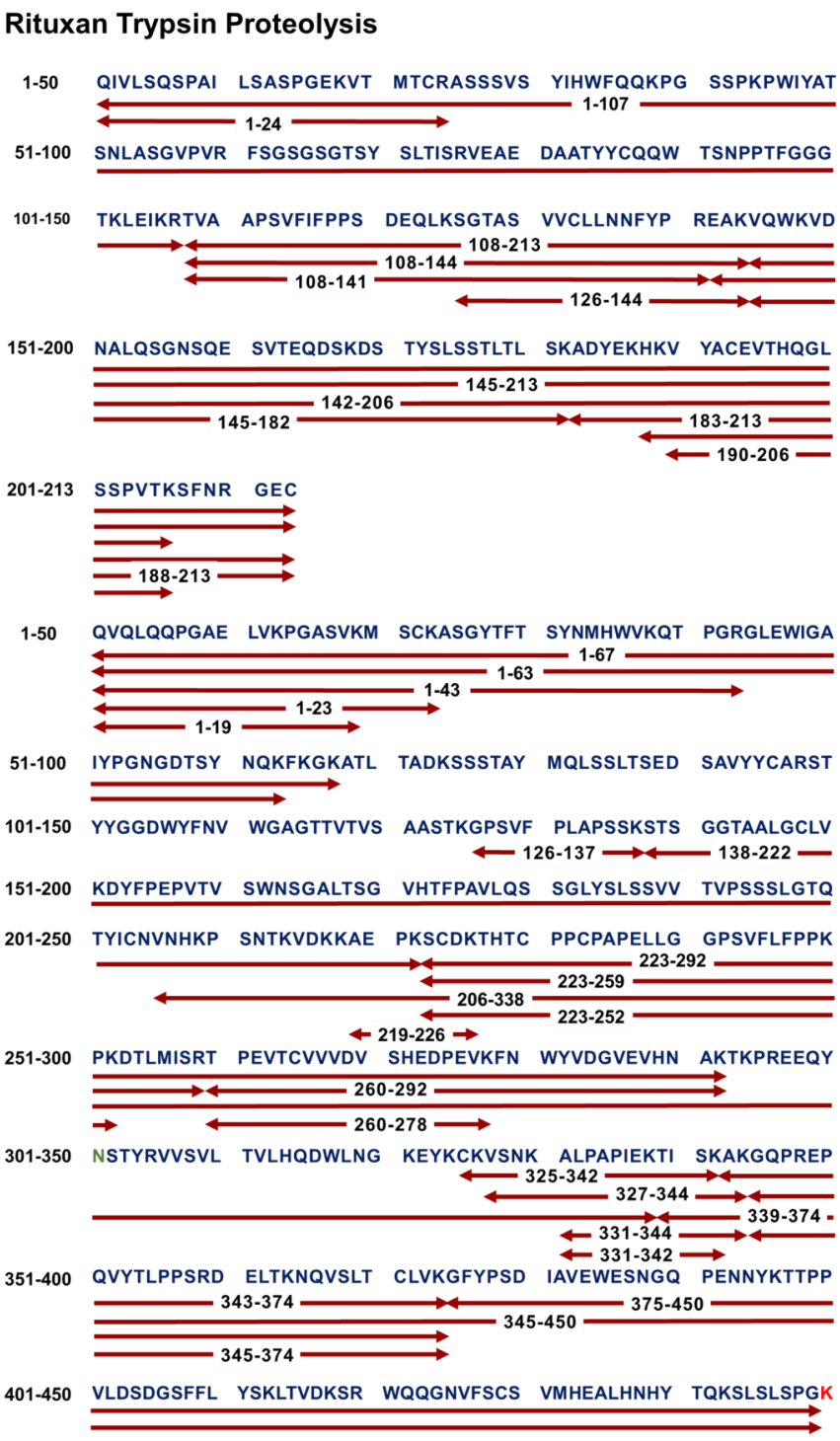


Figure 3.17. Sequence map of the peptides identified from infusion ESI-Orbitrap analysis of a tryptic digest of Rituxan. Light green color "N" represents the glycosylation site. Red color "K" represents the C-terminal clipping.

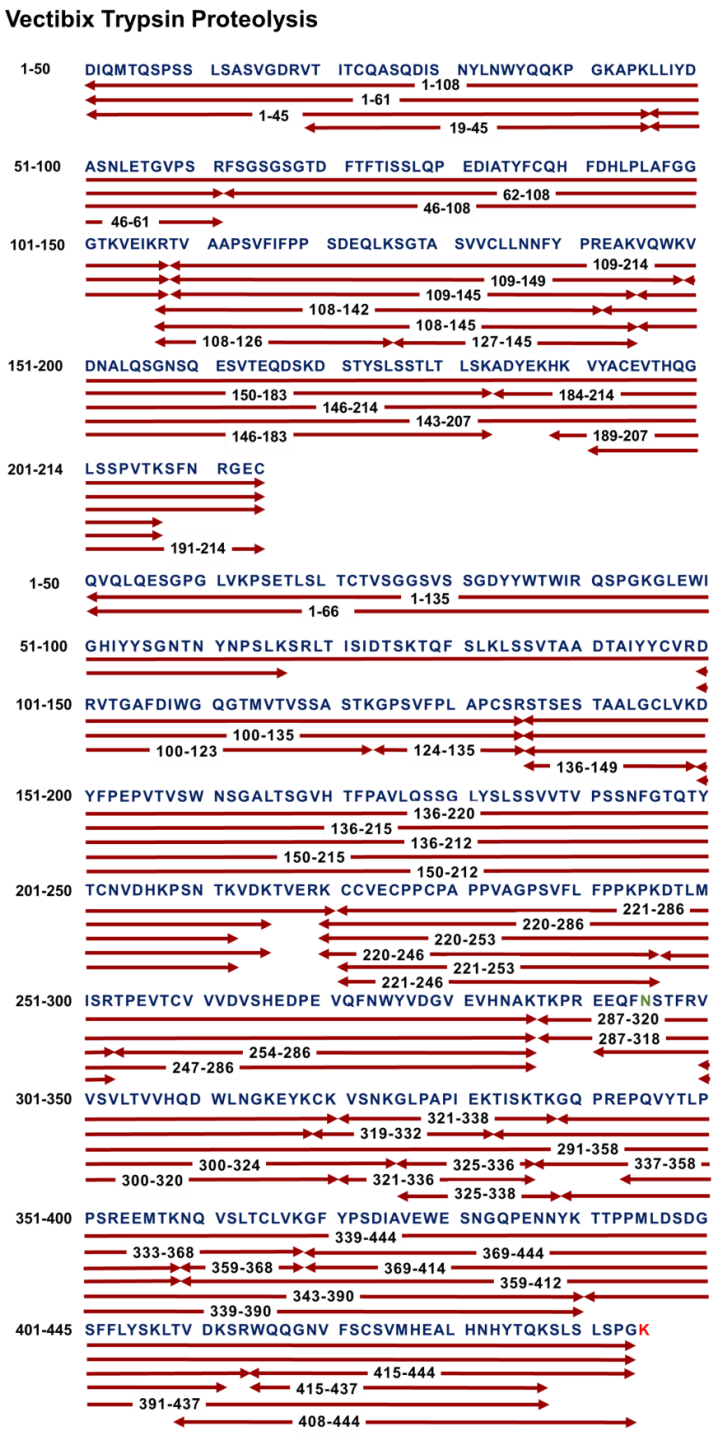


Figure 3.18. Sequence map of the peptides identified from infusion ESI-Orbitrap analysis of a tryptic digest of Vectibix. Light green color "N" represents the glycosylation site. Red color "K" represents the C-terminal clipping.

Table 3.8. Light- and heavy-chain peptides identified from a spin-membrane (spun at 500 *g*) tryptic digest of Herceptin.

m/z of [M+H]+	Peptide Sequence	Amino Acids
2609.2658	(-)DIQMTQSPSSLSASVGDRVTITC(Carbamidomethyl)R(A)	L1-24
7165.6262	(-)DIQMTQSPSSLSASVGDRVTITC(Carbamidomethyl)RASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPSRFSGSR(S)	L1-66
11959.9174	(-)DIQMTQSPSSLSASVGDRVTITC(Carbamidomethyl)RASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPSRFSGSRSGTDFTLTISSLQPEDFATYYC(Carbamidomethyl)QQHYTTPPTFGQGTKVEIKR(T)	L1-108
2287.183	(R)ASQDVNTAVAWYQQKPGKAPK(L)	L25-45
4575.3783	(R)ASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPSRFSGSR(S)	L25-66
9369.6695	(R)ASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPSRFSGSRSGTDFTLTISSLQPEDFATYYC(Carbamidomethyl)QQHYTTPPTFGQGTKVEIKR(T)	L25-108
1772.9581	(K)LLIYSASFLYSGVPSR(F)	L46-61
2307.2132	(K)LLIYSASFLYSGVPSRFSGSR(S)	L46-66
4657.2079	(R)SGTDFTLTISSLQPEDFATYYC(Carbamidomethyl)QQHYTTPPTFGQGTKVEIK(R)	L67-107
4813.309	(R)SGTDFTLTISSLQPEDFATYYC(Carbamidomethyl)QQHYTTPPTFGQGTKVEIKR(T)	L67-108
7336.7223	(K)VEIKRTVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSK(D)	L104-169
3881.0055	(K)RTVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPR(E)	L108-142
4209.1801	(K)RTVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPREAK(V)	L108-145
1946.027	(R)TVAAPSVFIFPPSDEQLK(S)	L109-126
3724.9043	(R)TVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPR(E)	L109-142
4053.079	(R)TVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPREAK(V)	L109-145
4594.3803	(R)TVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPREAKVQWK(V)	L109-149
10923.3997	(R)TVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L109-207
11773.7389	(R)TVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTKSFNRGEC(Carbamidomethyl)(-)	L109-214
2126.0699	(K)SGTASVVC(Carbamidomethyl)LLNNFYPREAK(V)	L127-145

Table 3.8 (cont'd)

7739.6777	(K)VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTL SKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTK SFNRGEC(Carbamidomethyl)(-)	L146-214
1502.7584	(K)DSTYSLSSTLTLSK(A)	L170-183
2747.3457	(K)ADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTK (S)	L184-207
3597.6849	(K)ADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTK SFNRGEC(Carbamidomethyl)(-)	L184-214
2141.0808	(K)HKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L189-207
1875.9269	(K)VYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L191-207
2726.2661	(K)VYAC(Carbamidomethyl)EVTHQGLSSPVTKSFNRGEC(Carbamidomethyl)(-)	L191-214
1882.0029	(-)EVQLVESGGGLVQPGGSLR(L)	H1-19
4101.0975	(-)EVQLVESGGGLVQPGGSLRLSC(Carbamido-methyl)AASGFNIKDTYIHWVR(Q)	H1-38
5393.7964	(-)EVQLVESGGGLVQPGGSLRLSC(Carbamido-methyl)AASGFNIKDTYIHWVRQAPGKGLEWVAR(I)	H1-50
6459.3208	(-)EVQLVESGGGLVQPGGSLRLSC(Carbamido-methyl)AASGFNIKDTYIHWVRQAPGKGLEWVARIYPTN GYTR(Y)	H1-59
2238.1124	(R)LSC(Carbamidomethyl)AASGFNIKDTYIHWVR (Q)	H20-38
3530.8114	(R)LSC(Carbamidomethyl)AASGFNIKDTYIHWVRQAPGK GLEWVAR(I)	H20-50
5472.7811	(R)LSC(Carbamidomethyl)AASGFNIKDTYIHWVRQAPGK GLEWVARIYPTNGYTRYADSVKGR(F)	H20-67
1089.5476	(K)DTYIHWVR(Q)	H31-38
2377.2411	(R)QAPGKGLEWVARIYPTNGYTR(Y)	H39-59
895.4632	(R)YADSVKGR(F)	H60-67
3576.6734	(R)FTISADTSKNTAYLQMNSLRAEDTAVYYC(Carbamido methyl)SR(W)	H68-98
3951.8898	(R)WGGDGFYAMDYWGQGTLVTVSSASTKGPSVFPLAP SSK(S)	H99-136
1186.6467	(K)GPSVFPLAPSSK(S)	H125-136
1321.678	(K)STSGGTAALGC(Carbamidomethyl)LVK(D)	H137-150
8939.5187	(K)STSGGTAALGC(Carbamidomethyl)LVKDYFPEPVTVS WNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQ TYIC(Carbamidomethyl)NVNHKPSNTKVDKKVEPK(S)	H137-221
7636.8585	(K)DYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSS VVTVPSSSLGTQTYIC(Carbamidomethyl)NVNHKPSNTKV DKKVEPK(S)	H151-221
3334.6421	(K)SC(Carbamidomethyl)DKTHTC(Carbamidomethyl)PPC(C arbamidomethyl)PAPELLGGPSVFLFPPKPK(D)	H222-251
4151.0585	(K)SC(Carbamidomethyl)DKTHTC(Carbamidomethyl)PPC(C arbamidomethyl)PAPELLGGPSVFLFPPKPKDTLMISR(T)	H222-258

Table 3.8 (cont'd)

7929.8522	(K)SC(Carbamidomethyl)DKTHTC(Carbamidomethyl)PPC(Carbamidomethyl)PAPELLGGPSVFLFPPKPKDTLMISRTPEVTC(Carbamidomethyl)VVVDVSHEDPEVKFNWYVDGVEVHNAK(T)	H222-291
3797.8116	(R)TPEVTC(Carbamidomethyl)VVVDVSHEDPEVKFNWYVDGVEVHNAK(T)	H259-291
501.3144	(K)TKPR(E)	H292-295
2634.238	(R)EEQYNSTYR(V)*	H296-304
2516.333	(R)VVSVLTVLHQDWLNGKEYKC(Carbamidomethyl)K(V)	H305-325
1895.1324	(K)VSNKALPAPIEKTISKAK(G)	H326-343
838.5033	(K)ALPAPIEK(T)	H330-337
1267.762	(K)ALPAPIEKTISK(A)	H330-341
1466.8941	(K)ALPAPIEKTISKAK(G)	H330-343
4933.6642	(K)ALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVK(G)	H330-373
4114.1787	(K)TISKAKGQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVK(G)	H338-373
3684.92	(K)AKGQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVK(G)	H342-373
2343.1762	(K)GQPREPQVYTLPPSREEMTK(N)	H344-363
3485.7879	(K)GQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVK(G)	H344-373
12088.8484	(K)GQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSC(Carbamidomethyl)SVMHEALHNHYTQKSLSLSPG(-)	H344-449
3047.554	(R)EPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVK(G)	H348-373
1161.6296	(K)NQVSLTC(Carbamidomethyl)LVK(G)	H364-373
2544.1314	(K)GFYPSDIAVEWESNGQPENNYK(T)	H374-395
4399.0354	(K)GFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSK(L)	H374-412
5198.4906	(K)GFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSR(W)	H374-419
7980.7398	(K)GFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSC(Carbamidomethyl)SVMHEALHNHYTQK(S)	H374-442
8622.0783	(K)GFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSC(Carbamidomethyl)SVMHEALHNHYTQKSLSLSPG(-)	H374-449
6096.9647	(K)TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSC(Carbamidomethyl)SVMHEALHNHYTQKSLSLSPG(-)	H396-449
2801.2671	(R)WQQGNVFSC(Carbamidomethyl)SVMHEALHNHYTQK(S)	H420-442
3442.6056	(R)WQQGNVFSC(Carbamidomethyl)SVMHEALHNHYTQKSLSLSPG(-)	H420-449

*2634.238 is the monoisotopic mass for H296-304 with G0F glycosylation.

Table 3.9. Light- and heavy-chain peptides identified from a spin-membrane (spun at 500 *g*) tryptic digest of Avastin.

m/z of [M+H]+	Peptide Sequence	Amino Acids
1878.8862	(-)DIQMTQSPSSLSASVGDR(V)	L1-18
4957.4095	(-)DIQMTQSPSSLSASVGDRVTITC(Carbamidomethyl)SASQDISNYLNWYQQKPGKAPK(V)	L1-45
6701.3403	(-)DIQMTQSPSSLSASVGDRVTITC(Carbamidomethyl)SASQDISNYLNWYQQKPGKAPKVLIYFTSSLHSGVPSR(F)	L1-61
2801.3563	(R)VTITC(Carbamidomethyl)SASQDISNYLNWYQQKPGK(A)	L19-42
3097.5411	(R)VTITC(Carbamidomethyl)SASQDISNYLNWYQQKPGKAPK(V)	L19-45
1762.9486	(K)VLIYFTSSLHSGVPSR(F)	L46-61
5285.5048	(R)FSGSGSGTDFTLTISSLQPEDFATYYC(Carbamidomethyl)QQYSTVPWTFGQGTKVEIKR(T)	L62-108
3881.0055	(K)RTVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPR(E)	L108-142
4209.1801	(K)RTVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPREAK(V)	L108-145
1946.027	(R)TVAAPSVFIFPPSDEQLK(S)	L109-126
3724.9043	(R)TVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPR(E)	L109-142
4053.079	(R)TVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPREAK(V)	L109-145
4594.3803	(R)TVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPREAKVQWK(V)	L109-149
1797.8952	(K)SGTASVVC(Carbamidomethyl)LLNNFYPR(E)	L127-142
2126.0699	(K)SGTASVVC(Carbamidomethyl)LLNNFYPREAK(V)	L127-145
4489.1853	(R)EAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK(A)	L143-183
7217.5132	(R)EAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L143-207
4161.0106	(K)VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK(A)	L146-183
5032.4294	(K)VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHK(V)	L146-190
6889.3385	(K)VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L146-207
7739.6777	(K)VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTKSFNRGEC(Carbamidomethyl)(-)	L146-214
3619.7093	(K)VDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK(A)	L150-183
6348.0372	(K)VDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L150-207

Table 3.9 (cont'd)

1502.7584	(K)DSTYSLSSTLTLSK(A)	L170-183
2747.3457	(K)ADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L184-207
3597.6849	(K)ADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTKSFNRGEC(Carbamidomethyl)(-)	L184-214
2141.0808	(K)HKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L189-207
2991.42	(K)HKVYAC(Carbamidomethyl)EVTHQGLSSPVTKSFNRGEC(Carbamidomethyl)(-)	L189-214
1875.9269	(K)VYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L191-207
2726.2661	(K)VYAC(Carbamidomethyl)EVTHQGLSSPVTKSFNRGEC(Carbamidomethyl)(-)	L191-214
1882.0029	(-)EVQLVESGGGLVQPGGSLR(L)	H1-19
2197.9794	(R)LSC(Carbamidomethyl)AASGYTFTNYGMNWVR(Q)	H20-38
5178.4189	(R)LSC(Carbamidomethyl)AASGYTFTNYGMNWVRQAPGKGLEWVGWINTYTGEPTYAADFK(R)	H20-65
5334.52	(R)LSC(Carbamidomethyl)AASGYTFTNYGMNWVRQAPGKGLEWVGWINTYTGEPTYAADFKR(R)	H20-66
2999.4574	(R)QAPGKGLEWVGWINTYTGEPTYAADFK(R)	H39-65
3155.5585	(R)QAPGKGLEWVGWINTYTGEPTYAADFKR(R)	H39-66
2674.2936	(K)GLEWVGWINTYTGEPTYAADFKR(R)	H44-66
1201.6212	(R)RFTFSLDTSK(S)	H67-76
1045.5201	(R)FTFSLDTSK(S)	H68-76
1283.6412	(K)STAYLQMNSLR(A)	H77-87
5762.7213	(R)AEDTAVYYC(Carbamidomethyl)AKYPHYYGSSHWYFDVWGQGTLVTVSSASTKGPSVFPLAPSSK(S)	H88-139
1186.6467	(K)GPSVFPLAPSSK(S)	H128-139
1321.678	(K)STSGGTAALGC(Carbamidomethyl)LVK(D)	H140-153
8015.9746	(K)STSGGTAALGC(Carbamidomethyl)LVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC(Carbamidomethyl)NVNHKPSNTK(V)	H140-216
8358.165	(K)STSGGTAALGC(Carbamidomethyl)LVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC(Carbamidomethyl)NVNHKPSNTKVDK(K)	H140-219
8486.2599	(K)STSGGTAALGC(Carbamidomethyl)LVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC(Carbamidomethyl)NVNHKPSNTKVDKK(V)	H140-210
8939.5187	(K)STSGGTAALGC(Carbamidomethyl)LVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC(Carbamidomethyl)NVNHKPSNTKVDKKVEPK(S)	H140-224
6713.3145	(K)DYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC(Carbamidomethyl)NVNHKPSNTK(V)	H154-216
7055.5048	(K)DYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC(Carbamidomethyl)NVNHKPSNTKVDK(K)	H154-219

Table 3.9 (cont'd)

7636.8585	(K)DYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC(Carbamidomethyl)NVNHKPSNTKVDKKVEPK(S)	H154-224
3334.6421	(K)SC(Carbamidomethyl)DKTHTC(Carbamidomethyl)PPC(Carbamidomethyl)PAPELLGGPSVFLFPPKPK(D)	H225-254
4151.0585	(K)SC(Carbamidomethyl)DKTHTC(Carbamidomethyl)PPC(Carbamidomethyl)PAPELLGGPSVFLFPPKPKDTLMISR(T)	H225-261
6271.0681	(K)SC(Carbamidomethyl)DKTHTC(Carbamidomethyl)PPC(Carbamidomethyl)PAPELLGGPSVFLFPPKPKDTLMISRTPEVTC(Carbamidomethyl)VVVDVSHEDPEVK(F)	H225-280
7929.8522	(K)SC(Carbamidomethyl)DKTHTC(Carbamidomethyl)PPC(Carbamidomethyl)PAPELLGGPSVFLFPPKPKDTLMISRTPEVTC(Carbamidomethyl)VVVDVSHEDPEVKFNWYVDGVEVHNAK(T)	H225-294
2844.4575	(K)THTC(Carbamidomethyl)PPC(Carbamidomethyl)PAPELLGGPSVFLFPPKPK(D)	H229-254
4614.2279	(K)DTLMISRTPEVTC(Carbamidomethyl)VVVDVSHEDPEVKFNWYVDGVEVHNAK(T)	H255-294
2139.0274	(R)TPEVTC(Carbamidomethyl)VVVDVSHEDPEVK(F)	H262-280
3797.8116	(R)TPEVTC(Carbamidomethyl)VVVDVSHEDPEVKFNWYVDGVEVHNAK(T)	H262-294
1677.802	(K)FNWYVDGVEVHNAK(T)	H281-294
5325.5188	(K)TKPREEQYNSTYRVVSVLTVLHQDWLNGKEYK(C)*	H295-326
5613.679	(K)TKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKC(Carbamidomethyl)K(V)**	H295-328
6042.0424	(K)TKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKC(Carbamidomethyl)KVSNK(A)***	H295-332
2228.2074	(R)VVSVLTVLHQDWLNGKEYK(C)	H308-326
2516.333	(R)VVSVLTVLHQDWLNGKEYKC(Carbamidomethyl)K(V)	H308-328
735.3818	(K)C(Carbamidomethyl)KVSNK(A)	H327-332
1696.0003	(K)VSNKALPAPIEKTISK(A)	H329-344
1895.1324	(K)VSNKALPAPIEKTISKAK(G)	H329-346
838.5033	(K)ALPAPIEK(T)	H333-340
1267.762	(K)ALPAPIEKTISK(A)	H333-344
1466.8941	(K)ALPAPIEKTISKAK(G)	H333-346
4933.6642	(K)ALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVK(G)	H333-376
2971.567	(K)TISKAKGQPREPQVYTLPPSREEMTK(N)	H341-366
4114.1787	(K)TISKAKGQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVK(G)	H341-376
2542.3082	(K)AKGQPREPQVYTLPPSREEMTK(N)	H346-366
3684.92	(K)AKGQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVK(G)	H345-376
6210.0335	(K)AKGQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVKGFYPSDIAVEWESNGQPENNYK(T)	H345-398

Table 3.9 (cont'd)

8621.2596	(K)AKGQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDK(S)	H345-420
2343.1762	(K)GQPREPQVYTLPPSREEMTK(N)	H347-366
3485.7879	(K)GQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVK(G)	H347-376
6010.9015	(K)GQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVKGFYPSDIAVEWESNGQPENNYK(T)	H347-398
8665.2607	(K)GQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSR(W)	H347-422
12091.8591	(K)GQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSC(Carbamidomethyl)SVMHEALHNHYTQKSLSLSPG(-)	H347-452
3047.554	(R)EPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVK(G)	H351-376
5572.6676	(R)EPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVKGFYPSDIAVEWESNGQPENNYK(T)	H351-398
1161.6296	(K)NQVSLTC(Carbamidomethyl)LVK(G)	H367-376
6097.9692	(K)NQVSLTC(Carbamidomethyl)LVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDK(S)	H367-420
2544.1314	(K)GFYPSDIAVEWESNGQPENNYK(T)	H377-398
4399.0354	(K)GFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSK(L)	H377-415
5198.4906	(K)GFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSR(W)	H377-422
7980.7398	(K)GFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSC(Carbamidomethyl)SVMHEALHNHYTQK(S)	H377-445
2673.377	(K)TTPPVLDSDGSFFLYSKLTVDKSR(W)	H399-422
5455.6263	(K)TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSC(Carbamidomethyl)SVMHEALHNHYTQK(S)	H399-445
3442.6043	(R)WQQGNVFSC(Carbamidomethyl)SVMHEALHNHYTQKSLSLSPG(-)	H423-452

*5325.5188 is the monoisotopic mass for H295-326 with G0F glycosylation.
**5613.679 is the monoisotopic mass for H295-328 with G0F glycosylation.
***6042.0424 is the monoisotopic mass for H295-332 with G0F glycosylation.

Table 3.10. Light- and heavy-chain peptides identified from a spin-membrane (spun at 500 *g*) tryptic digest of Rituxan.

m/z of [M+H]+	Peptide Sequence	Amino Acids
2556.316	(-)QIVLSQSPAILSASPGEKVTMTC(Carbamidomethyl)R(A)	L1-24
11556.7107	(-)QIVLSQSPAILSASPGEKVTMTC(Carbamidomethyl)RASSSVSYIHWFQQKPGSSPKPWIYATSNLASGVPVRFSGSGSGTSYSLTISRVEAEDAATYYC(Carbamidomethyl)QQWTSNPPTFGGGTKLEIKR(T)	L1-107
3724.9043	(R)TVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPR(E)	L108-141
4053.079	(R)TVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPREAK(V)	L108-144
10923.3997	(R)TVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L108-206
11773.7389	(R)TVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTKSFNRGEC(Carbamidomethyl)(-)	L108-213
2126.0699	(K)SGTASVVC(Carbamidomethyl)LLNNFYPREAK(V)	L126-144
7217.5132	(R)EAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L142-206
4161.0106	(K)VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK(A)	L145-182
6889.3385	(K)VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L145-206
7739.6777	(K)VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTKSFNRGEC(Carbamidomethyl)(-)	L145-213
2747.3457	(K)ADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L183-206
3597.6849	(K)ADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTKSFNRGEC(Carbamidomethyl)(-)	L183-213
2141.0808	(K)HKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L188-206
2991.42	(K)HKVYAC(Carbamidomethyl)EVTHQGLSSPVTKSFNRGEC(Carbamidomethyl)(-)	L188-213
1875.9269	(K)VYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L190-206
2726.2661	(K)VYAC(Carbamidomethyl)EVTHQGLSSPVTKSFNRGEC(Carbamidomethyl)(-)	L190-213
1960.0862	(-)QVQLQQPGAELVKPGASVK(M)	H1-19
2466.2843	(-)QVQLQQPGAELVKPGASVKMSC(Carbamidomethyl)K(A)	H1-23

Table 3.10 (cont'd)

4778.364	(-)QVQLQQPGAELVKPGASVKMSC(Carbamido methyl)KASGYTFTSYNMHWVKQTPGR(G)	H1-43
6942.3865	(-)QVQLQQPGAELVKPGASVKMSC(Carbamido methyl)KASGYTFTSYNMHWVKQTPGRGLEWIGAIYPGN GDTSYNQK(F)	H1-63
7402.6663	(-)QVQLQQPGAELVKPGASVKMSC(Carbamido methyl)KASGYTFTSYNMHWVKQTPGRGLEWIGAIYPGN GDTSYNQKFKGK(A)	H1-67
1186.6467	(K)GPSVFPLAPSSK(S)	H126-137
8911.4874	(K)STSGGTAALGC(Carbamidomethyl)LVKDYFPEPVTVS WNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQ TYIC(Carbamidomethyl)NVNHKPSNTKVDKKAEPK(S)	H138-222
934.4299	(K)AEPKSC(Carbamidomethyl)DK(T)	H219-226
3334.6421	(K)SC(Carbamidomethyl)DKTHTC(Carbamidomethyl)PPC(C arbamidomethyl)PAPELLGGPSVFLFPPKPK(D)	H223-252
4151.0585	(K)SC(Carbamidomethyl)DKTHTC(Carbamidomethyl)PPC(C arbamidomethyl)PAPELLGGPSVFLFPPKPKDTLMISR(T)	H223-259
7929.8522	(K)SC(Carbamidomethyl)DKTHTC(Carbamidomethyl)PPC(C arbamidomethyl)PAPELLGGPSVFLFPPKPKDTLMISRTPEV TC(Carbamidomethyl)VVVDVSHEDPEVKFNWYVDGVEV HNAK(T)	H223-292
2139.0274	(R)TPEVTC(Carbamidomethyl)VVVDVSHEDPEVK(F)	H260-278
3797.8116	(R)TPEVTC(Carbamidomethyl)VVVDVSHEDPEVKFNWYV DGVEVHNAK(T)	H260-292
3764.0568	(R)VVSVLTVLHQDWLNGKEYKC(Carbamidomethyl)KVS NKALPAPIEK(T)	H306-338
1984.126	(K)C(Carbamidomethyl)KVSNKALPAPIEKTISK(A)	H325-342
1696.0003	(K)VSNKALPAPIEKTISK(A)	H327-342
1895.1324	(K)VSNKALPAPIEKTISKAK(G)	H327-344
1267.762	(K)ALPAPIEKTISK(A)	H331-342
1466.8941	(K)ALPAPIEKTISKAK(G)	H331-344
4082.2067	(K)TISKAKGQPREPQVYTLPPSRDELTKNQVSLTC(Carba midomethyl)LVK(G)	H339-374
3652.9479	(K)AKGQPREPQVYTLPPSRDELTKNQVSLTC(Carbamido methyl)LVK(G)	H343-374
3453.8159	(K)GQPREPQVYTLPPSRDELTKNQVSLTC(Carbamidomet hyl)LVK(G)	H345-374
12056.8763	(K)GQPREPQVYTLPPSRDELTKNQVSLTC(Carbamidomet hyl)LVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGS FFLYSKLTVDKSRWQQGNVFSC(Carbamidomethyl)SVMH EALHNHYTQKSLSLSPG(-)	H345-450
8622.0783	(K)GFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLY SKLTVDKSRWQQGNVFSC(Carbamidomethyl)SVMHEAL HNHYTQKSLSLSPG(-)	H375-450

Table 3.11. Light- and heavy-chain peptides identified from a spin-membrane (spun at 500 *g*) tryptic digest of Vectibix.

m/z of [M+H]+	Peptide Sequence	Amino Acids
4998.436	(-)DIQMTQSPSSLSASVGDRVTITC(Carbamidomethyl)QASQDISNYLNWYQQKPGKAPK(L)	L1-45
6727.3407	(-)DIQMTQSPSSLSASVGDRVTITC(Carbamidomethyl)QASQDISNYLNWYQQKPGKAPKLLIYDASNLETGVPSR(F)	L1-61
11874.817	(-)DIQMTQSPSSLSASVGDRVTITC(Carbamidomethyl)QASQDISNYLNWYQQKPGKAPKLLIYDASNLETGVPSRFSGSGSGTDFTFTISSLQPEDIATYFC(Carbamidomethyl)QHFDHLPLAFGGGTKVEIKR(T)	L1-108
3138.5677	(R)VTITC(Carbamidomethyl)QASQDISNYLNWYQQKPGKAPK(L)	L19-45
1747.9225	(K)LLIYDASNLETGVPSR(F)	L46-61
6895.3989	(K)LLIYDASNLETGVPSRFSGSGSGTDFTFTISSLQPEDIATYFC(Carbamidomethyl)QHFDHLPLAFGGGTKVEIKR(T)	L46-108
5166.4942	(R)FSGSGSGTDFTFTISSLQPEDIATYFC(Carbamidomethyl)QHFDHLPLAFGGGTKVEIKR(T)	L62-108
2102.1281	(K)RTVAAPSVFIFPPSDEQLK(S)	L108-126
3881.0055	(K)RTVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPR(E)	L108-142
4209.1801	(K)RTVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPREAK(V)	L108-145
1946.027	(R)TVAAPSVFIFPPSDEQLK(S)	L109-126
3724.9043	(R)TVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPR(E)	L109-142
4053.079	(R)TVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPREAK(V)	L109-145
4594.3803	(R)TVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPREAKVQWK(V)	L109-149
10923.3997	(R)TVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L109-207
11773.7389	(R)TVAAPSVFIFPPSDEQLKSGTASVVC(Carbamidomethyl)LLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTKSFNRGEC(Carbamidomethyl)(-)	L109-214
2126.0699	(K)SGTASVVC(Carbamidomethyl)LLNNFYPREAK(V)	L127-145
8996.3906	(K)SGTASVVC(Carbamidomethyl)LLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L127-207
7217.5132	(R)EAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L143-207

Table 3.11 (cont'd)

8067.8524	(R)EAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTKSFNRGEC(Carbamidomethyl)(-)	L143-214
4161.0106	(K)VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK(A)	L146-183
6889.3385	(K)VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L146-207
7739.6777	(K)VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTKSFNRGEC(Carbamidomethyl)(-)	L146-214
3619.7093	(K)VDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK(A)	L150-183
6348.0372	(K)VDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L150-207
2747.3457	(K)ADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L184-207
3597.6849	(K)ADYEKHKVYAC(Carbamidomethyl)EVTHQGLSSPVTKSFNRGEC(Carbamidomethyl)(-)	L184-214
2141.0808	(K)HKVYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L189-207
1875.9269	(K)VYAC(Carbamidomethyl)EVTHQGLSSPVTK(S)	L191-207
2726.2661	(K)VYAC(Carbamidomethyl)EVTHQGLSSPVTKSFNRGEC(Carbamidomethyl)(-)	L191-214
7260.539	(-)QVQLQESGPGLVKPSETLSLTC(Carbamidomethyl)TVSGGSVSSGDYYWTWIRQSPGKGLEWIGHIYYSGNTNYNPSLK(S)	H1-66
14701.312	(-)QVQLQESGPGLVKPSETLSLTC(Carbamidomethyl)TVSGGSVSSGDYYWTWIRQSPGKGLEWIGHIYYSGNTNYNPSLKSRLTISIDTSKTQFSLKLSSVTAADTAIYYC(Carbamidomethyl)VRDRVTGAFDIWGQGTMVTVSSASTKGPSVFPLAPC(Carbamidomethyl)SR(S)	H1-135
2514.2293	(R)DRVTGAFDIWGQGTMVTVSSASTK(G)	H100-123
3782.8629	(R)DRVTGAFDIWGQGTMVTVSSASTKGPSVFPLAPC(Carbamidomethyl)SR(S)	H100-135
1287.6514	(K)GPSVFPLAPC(Carbamidomethyl)SR(S)	H124-135
1423.7097	(R)STSESTAALGC(Carbamidomethyl)LVK(D)	H136-149
8167.9492	(R)STSESTAALGC(Carbamidomethyl)LVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSNFGTQTYTC(Carbamidomethyl)NVDHKPSNTK(V)	H136-212
8510.1395	(R)STSESTAALGC(Carbamidomethyl)LVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSNFGTQTYTC(Carbamidomethyl)NVDHKPSNTKVDK(T)	H136-215
9123.4943	(R)STSESTAALGC(Carbamidomethyl)LVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSNFGTQTYTC(Carbamidomethyl)NVDHKPSNTKVDKTVERK(C)	H136-220
6763.2573	(K)DYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSNFGTQTYTC(Carbamidomethyl)NVDHKPSNTK(V)	H150-212

Table 3.11 (cont'd)

7105.4477	(K)DYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSNFGTQTYTC(Carbamidomethyl)NVDHKPSNTKVDK(T)	H150-215
3036.4966	(R)KC(Carbamidomethyl)C(Carbamidomethyl)VEC(Carbamidomethyl)PPC(Carbamidomethyl)PAPPVAGPSVFLFPPKPK(D)	H220-246
3852.913	(R)KC(Carbamidomethyl)C(Carbamidomethyl)VEC(Carbamidomethyl)PPC(Carbamidomethyl)PAPPVAGPSVFLFPPKPKDTLMISR(T)	H220-253
7631.6704	(R)KC(Carbamidomethyl)C(Carbamidomethyl)VEC(Carbamidomethyl)PPC(Carbamidomethyl)PAPPVAGPSVFLFPPKPKDTLMISRTPEVTC(Carbamidomethyl)VVVDVSHEDPEVQFNWYVDGVEVHNAK(T)	H220-286
2908.4017	(K)C(Carbamidomethyl)C(Carbamidomethyl)VEC(Carbamidomethyl)PPC(Carbamidomethyl)PAPPVAGPSVFLFPPKPK(D)	H221-246
3724.8181	(K)C(Carbamidomethyl)C(Carbamidomethyl)VEC(Carbamidomethyl)PPC(Carbamidomethyl)PAPPVAGPSVFLFPPKPKDTLMISR(T)	H221-253
7503.5754	(K)C(Carbamidomethyl)C(Carbamidomethyl)VEC(Carbamidomethyl)PPC(Carbamidomethyl)PAPPVAGPSVFLFPPKPKDTLMISRTPEVTC(Carbamidomethyl)VVVDVSHEDPEVQFNWYVDGVEVHNAK(T)	H221-286
4614.1916	(K)DTLMISRTPEVTC(Carbamidomethyl)VVVDVSHEDPEVQFNWYVDGVEVHNAK(T)	H247-286
3797.7752	(R)TPEVTC(Carbamidomethyl)VVVDVSHEDPEVQFNWYVDGVEVHNAK(T)	H254-286
5279.5233	(K)TKPREEQFNSTFRVVSVLTVVHQDWLNGKEYK(C)*	H287-318
5567.649	(K)TKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKC(Carbamidomethyl)K(V)**	H287-320
9301.4297	(R)EEQFNSTFRVVSVLTVVHQDWLNGKEYKC(Carbamidomethyl)KVSNKGLPAPIEKTISKTKGQPREPQVYTLPPSREEMTK(N)***	H291-358
2214.1917	(R)VVSVLTVVHQDWLNGKEYK(C)	H300-318
2502.3173	(R)VVSVLTVVHQDWLNGKEYKC(Carbamidomethyl)K(V)	H300-320
2930.5557	(R)VVSVLTVVHQDWLNGKEYKC(Carbamidomethyl)KVSNK(G)	H300-324
3736.0255	(R)VVSVLTVVHQDWLNGKEYKC(Carbamidomethyl)KVSNKGLPAPIEK(T)	H300-332
1540.8516	(K)C(Carbamidomethyl)KVSNKGLPAPIEK(T)	H319-332
1252.726	(K)VSNKGLPAPIEK(T)	H321-332
1681.9847	(K)VSNKGLPAPIEKTISK(T)	H321-336
1911.1273	(K)VSNKGLPAPIEKTISKTK(G)	H321-338
1253.7464	(K)GLPAPIEKTISK(T)	H325-336
1482.889	(K)GLPAPIEKTISKTK(G)	H325-338
3807.0473	(K)GLPAPIEKTISKTKGQPREPQVYTLPPSREEMTK(N)	H325-358
4949.6591	(K)GLPAPIEKTISKTKGQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVK(G)	H325-368

Table 3.11 (cont'd)

3001.5775	(K)TISKTKGQPREPQVYTLPPSREEMTK(N)	H333-358
4144.1893	(K)TISKTKGQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVK(G)	H333-368
2572.3188	(K)TKGQPREPQVYTLPPSREEMTK(N)	H337-358
3714.9306	(K)TKGQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVK(G)	H337-368
2343.1762	(K)GQPREPQVYTLPPSREEMTK(N)	H339-358
3485.7879	(K)GQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVK(G)	H339-368
6010.9015	(K)GQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVKGFYPSDIAVEWESNGQPENNYK(T)	H339-390
11479.482	(K)GQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVKGFYPSDIAVEWESNGQPENNYKTTPPMLDSDGSFFLYSKLTVDKSRWQQGNVFSC(Carbamidomethyl)SVMHEALHNHYTQK(S)	H339-437
12120.8204	(K)GQPREPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVKGFYPSDIAVEWESNGQPENNYKTTPPMLDSDGSFFLYSKLTVDKSRWQQGNVFSC(Carbamidomethyl)SVMHEALHNHYTQKSLSLSPG(-)	H339-444
1904.9422	(R)EPQVYTLPPSREEMTK(N)	H343-358
3047.554	(R)EPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVK(G)	H343-368
5572.6676	(R)EPQVYTLPPSREEMTKNQVSLTC(Carbamidomethyl)LVKGFYPSDIAVEWESNGQPENNYK(T)	H343-390
1161.6296	(K)NQVSLTC(Carbamidomethyl)LVK(G)	H359-368
6129.9413	(K)NQVSLTC(Carbamidomethyl)LVKGFYPSDIAVEWESNGQPENNYKTTPPMLDSDGSFFLYSKLTVDK(S)	H359-412
5230.4626	(K)GFYPSDIAVEWESNGQPENNYKTTPPMLDSDGSFFLYSKLTVDKSR(W)	H369-414
8012.7119	(K)GFYPSDIAVEWESNGQPENNYKTTPPMLDSDGSFFLYSKLTVDKSRWQQGNVFSC(Carbamidomethyl)SVMHEALHNHYTQK(S)	H369-437
8654.0503	(K)GFYPSDIAVEWESNGQPENNYKTTPPMLDSDGSFFLYSKLTVDKSRWQQGNVFSC(Carbamidomethyl)SVMHEALHNHYTQKSLSLSPG(-)	H369-444
5487.5984	(K)TTPPMLDSDGSFFLYSKLTVDKSRWQQGNVFSC(Carbamidomethyl)SVMHEALHNHYTQK(S)	H391-437
4242.0608	(K)LTVDKSRWQQGNVFSC(Carbamidomethyl)SVMHEALHNHYTQKSLSLSPG(-)	H408-444
2801.2671	(R)WQQGNVFSC(Carbamidomethyl)SVMHEALHNHYTQK(S)	H415-437
3442.6056	(R)WQQGNVFSC(Carbamidomethyl)SVMHEALHNHYTQKSLSLSPG(-)	H415-444

*5279.5233 is the monoisotopic mass for H287-318 with G0F glycosylation.
**5567.649 is the monoisotopic mass for H287-320 with G0F glycosylation.
***9301.4297 is the monoisotopic mass for H291-358 with G0F glycosylation.

3.3.5 LC/MS-MS analyses

Direct infusion nanoESI is a powerful method for peptide mapping because of its short sample analysis time (<3 min for data collection). Also, injection of all peptides into the mass spectrometer avoids the peptide losses that are inevitable in LC.[55] However, peptides have different ionization efficiencies,[56] and ion suppression may occur during infusion MS analysis.[57] Moreover, with protein mixtures spin digestion may generate hundreds or thousands of peptides, making effective direct infusion analysis impossible.

LC-MS/MS is widely used to analyze complex protein mixtures, and its well-developed bioinformatics software makes data analysis possible. As an initial test of whether spin-membrane digestion enables analysis of antibody sequences and modifications using LC-MS/MS, we analyzed the antibody tryptic spin digests. (Peptic digests are much more difficult to analyze due to the limited cleavage specificity.[49]) Because mAbs are typically expressed in Chinese hamster ovary cell lines,[58] we identified proteolytic peptides through comparison to the hamster reference proteome from Uniprot, with the addition of the mAb sequences to the database. Using MaxQuant data analysis with this protein data base, we compared the sequence coverage and number of unique peptides identified after tryptic spin digestion either in a spin membrane or in solution.

Table 3.12. Antibody Sequence Coverages and Numbers of Unique Peptides Obtained From LC/MS-MS Analyses of Tryptic Spin and In-solution Digests.*

Tryptic Spin Digestion					
		Herceptin	Avastin	Rituxan	Vectibix
Sequence Coverage	Light Chain	100%	100%	100%	100%
	Heavy Chain	81.3%	85.2%	75.1%	98.4%
Unique Peptides	Light Chain	40	41	30	37
	Heavy Chain	66	72	43	52
Tryptic In-solution Digestion					
		Herceptin	Avastin	Rituxan	Vectibix
Sequence Coverage	Light Chain	87.4%	93.5%	93.0%	94.9%
	Heavy Chain	72.4%	73%	75.3%	79.7%
Unique Peptides	Light Chain	15	17	17	14
	Heavy Chain	33	33	32	30

*Peptides were identified using MaxQuant Software with comparison to a Chinese hamster proteome modified with antibody sequences.

As Table 3.12 shows, for all four antibodies tryptic spin digestion gives higher or essentially equal sequence coverage and more unique peptides than in-solution digestion. The missing sequences in the light chain after in-solution digestion likely result from undetectable small peptides with 3 or 4 amino acids. Heavy-chain sequence coverages are lower in LC-MS/MS analysis than in direct infusion analysis, primarily because we didn't consider the glycosylation on the heavy chain in the MaxQuant search. Glycosylated peptides also show low ionization efficiencies.[59] Enzymatic removal of the glycans prior to digestion will likely give heavy-chain sequence coverages near 100%. In comparing in-solution and spin digestion, the additional unique peptides from spin digestion may enhance protein identification in database searching with protein mixtures. Overall, spin digestion is a powerful method for fast protein digestion, and proteolytic peptides from spin digestion are suitable for downstream direct infusion or LC-MS/MS analysis.

3.4 Conclusions

This work used pepsin/trypsin spin membranes as microreactors for reproducible proteolysis prior to MS analysis. The high concentration of enzyme in the membrane pores allows spin digestion of 100 μL of antibody solution within 1 min. Peptic spin digestion avoids protein alkylation because the acidic conditions prevent reforming of disulfide bonds, whereas tryptic spin digestion benefits from alkylation. Moreover, with peptic spin digestion we can control the proteolytic peptide size by varying the spin rate. Direct infusion MS of spin digests is fast and provides a whole picture of the peptic/tryptic digests of single proteins. Essentially 100% peptide coverage along with identification of PTMSs results from direct infusion MS analyses of peptic and tryptic spin digests of Herceptin, Avastin, Rituxan and Vectibix. MaxQuant analyses of LC-MS/MS data reveal that tryptic spin digests of four mAbs give higher sequence coverage and more unique peptides than in-solution tryptic digestion of the same antibodies. In summary, the spin-digestion platform is rapid, simple, and user-friendly, and it affords control over peptide sizes for various types of subsequent MS analyses.

3.5 Acknowledgement

We gratefully acknowledge the U.S. National Science Foundation (CHE-1506315) for funding this work. We thank Dr. Mohammad Muhsin Chisti from Michigan State University for providing Herceptin, Avastin, Rituximab and Vectibix. We also thank Dr. Liangliang Sun (Michigan State University), Dr. Matthew Champion (Mass Spectrometry and Proteomics Facility of the University of Notre Dame) and Dr. Todd Lydic (Molecular Metabolism and Disease Collaborative Mass Spectrometry Core of Michigan State University) for helping to analyze the samples.

REFERENCES

REFERENCES

(1) Angel, T. E.; Aryal, U. K.; Hengel, S. M.; Baker, E. S.; Kelly, R. T.; Robinson, E. W.; Smith, R. D. Mass spectrometry-based proteomics: existing capabilities and future directions. *Chem. Soc. Rev.* **2012,** *41* (10), 3912.

(2) Yates, J. R.; Ruse, C. I.; Nakorchevsky, A. Proteomics by mass spectrometry: approaches, advances, and applications. *Annu. Rev. Biomed. Eng.* **2009,** *11*, 49.

(3) Monzo, A.; Sperling, E.; Guttman, A. Proteolytic enzyme-immobilization techniques for MS-based protein analysis. *Trac-Trend Anal. Chem.* **2009,** *28* (7), 854.

(4) Capelo, J. L.; Carreira, R.; Diniz, M.; Fernandes, L.; Galesio, M.; Lodeiro, C.; Santos, H. M.; Vale, G. Overview on modern approaches to speed up protein identification workflows relying on enzymatic cleavage and mass spectrometry-based techniques. *Anal. Chim. Acta* **2009,** *650* (2), 151.

(5) Ma, J. F.; Zhang, L. H.; Liang, Z.; Shan, Y. C.; Zhang, Y. K. Immobilized enzyme reactors in proteomics. *Trac-Trend Anal. Chem.* **2011,** *30* (5), 691.

(6) Switzar, L.; Giera, M.; Niessen, W. M. Protein digestion: an overview of the available techniques and recent developments. *J. Proteome. Res.* **2013,** *12* (3), 1067.

(7) Regnier, F. E.; Kim, J. Accelerating trypsin digestion: the immobilized enzyme reactor. *Bioanalysis* **2014,** *6* (19), 2685.

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

(9) Zang, L.; Carlage, T.; Murphy, D.; Frenkel, R.; Bryngelson, P.; Madsen, M.; Lyubarskaya, Y. Residual metals cause variability in methionine oxidation measurements in protein pharmaceuticals using LC-UV/MS peptide mapping. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2012,** *895-896*, 71.

(10) Jones, L. M.; Zhang, H.; Vidavsky, I.; Gross, M. L. Online, high-pressure digestion system for protein characterization by hydrogen/deuterium exchange and mass spectrometry. *Anal. Chem.* **2010,** *82* (4), 1171.

(11) http://www.sigmaaldrich.com/catalog/product/sigma/tt0010?lang=en&region=US; Vol. 2016.

(12) https://www.promega.com/products/mass-spectrometry/proteases-and-surfactants/trypsin-for-protein-characterization/trypsin-reagents/immobilized-trypsin/?activeTab=2; Vol. 2016.

(13) Ma, J.; Zhang, L.; Liang, Z.; Zhang, W.; Zhang, Y. Recent advances in immobilized enzymatic reactors and their applications in proteome analysis. *Anal. Chim. Acta* **2009,** *632* (1), 1.

(14) Spross, J.; Sinz, A. A capillary monolithic trypsin reactor for efficient protein digestion in online and offline coupling to ESI and MALDI mass spectrometry. *Anal. Chem.* **2010,** *82* (4), 1434.

(15) Krenkova, J.; Lacher, N. A.; Svec, F. Highly efficient enzyme reactors containing trypsin and endoproteinase LysC immobilized on porous polymer monolith coupled to MS suitable for analysis of antibodies. *Anal. Chem.* **2009,** *81* (5), 2004.

(16) Ma, J.; Liang, Z.; Qiao, X.; Deng, Q.; Tao, D.; Zhang, L.; Zhang, Y. Organic-inorganic hybrid silica monolith based immobilized trypsin reactor with high enzymatic activity. *Anal. Chem.* **2008,** *80* (8), 2949.

(17) Nicoli, R.; Gaud, N.; Stella, C.; Rudaz, S.; Veuthey, J. L. Trypsin immobilization on three monolithic disks for on-line protein digestion. *J. Pharm. Biomed. Anal.* **2008,** *48* (2), 398.

(18) Schoenherr, R. M.; Ye, M.; Vannatta, M.; Dovichi, N. J. CE-microreactor-CE-MS/MS for protein analysis. *Anal. Chem.* **2007,** *79* (6), 2230.

(19) Ota, S.; Miyazaki, S.; Matsuoka, H.; Morisato, K.; Shintani, Y.; Nakanishi, K. High-throughput protein digestion by trypsin-immobilized monolithic silica with pipette-tip formula. *J Biochem. Biophys. Methods* **2007,** *70* (1), 57.

(20) Besanger, T. R.; Hodgson, R. J.; Green, J. R.; Brennan, J. D. Immobilized enzyme reactor chromatography: optimization of protein retention and enzyme activity in monolithic silica stationary phases. *Anal. Chim. Acta* **2006,** *564* (1), 106.

(21) Feng, S.; Ye, M.; Jiang, X.; Jin, W.; Zou, H. Coupling the immobilized trypsin microreactor of monolithic capillary with muRPLC-MS/MS for shotgun proteome analysis. *J. Proteome. Res.* **2006,** *5* (2), 422.

(22) Temporini, C.; Perani, E.; Mancini, F.; Bartolini, M.; Calleri, E.; Lubda, D.; Felix, G.; Andrisano, V.; Massolini, G. Optimization of a trypsin-bioreactor coupled with high-performance liquid chromatography-electrospray ionization tandem mass spectrometry for quality control of biotechnological drugs. *J. Chromatogr. A* **2006,** *1120* (1-2), 121.

(23) Massolini, G.; Calleri, E. Immobilized trypsin systems coupled on-line to separation methods: Recent developments and analytical applications. *J. Sep. Sci.* **2005,** *28* (1), 7.

(24) Kato, M.; Inuzuka, K.; Sakai-Kato, K.; Toyo'oka, T. Monolithic bioreactor immobilizing trypsin for high-throughput analysis. *Anal. Chem.* **2005,** *77* (6), 1813.

(25) Krenkova, J.; Bilkova, Z.; Foret, F. Characterization of a monolithic immobilized trypsin microreactor with on-line coupling to ESI-MS. *J. Sep. Sci.* **2005,** *28* (14), 1675.

(26) Calleri, E.; Temporini, C.; Perani, E.; De Palma, A.; Lubda, D.; Mellerio, G.; Sala, A.; Galliano, M.; Caccialanza, G.; Massolini, G. Trypsin-based monolithic bioreactor coupled on-line with LC/MS/MS system for protein digestion and variant identification in standard solutions and serum samples. *J. Proteome. Res.* **2005,** *4* (2), 481.

(27) Ruan, G.; Wu, Z.; Huang, Y.; Wei, M.; Su, R.; Du, F. An easily regenerable enzyme reactor prepared from polymerized high internal phase emulsions. *Biochem. Biophys. Res. Commun.* **2016,** *473* (1), 54.

(28) Zhang, Z.; Sun, L.; Zhu, G.; Cox, O. F.; Huber, P. W.; Dovichi, N. J. Nearly 1000 Protein Identifications from 50 ng of Xenopus laevis Zygote Homogenate Using Online Sample Preparation on a Strong Cation Exchange Monolith Based Microreactor Coupled with Capillary Zone Electrophoresis. *Anal. Chem.* **2016,** *88* (1), 877.

(29) Sun, L.; Zhu, G.; Dovichi, N. J. Integrated capillary zone electrophoresis-electrospray ionization tandem mass spectrometry system with an immobilized trypsin microreactor for online digestion and analysis of picogram amounts of RAW 264.7 cell lysate. *Anal. Chem.* **2013,** *85* (8), 4187.

(30) Long, Y.; Wood, T. D. Immobilized pepsin microreactor for rapid peptide mapping with nanoelectrospray ionization mass spectrometry. *J. Am. Soc. Mass. Spectrom.* **2015,** *26* (1), 194.

(31) Yamaguchi, H.; Miyazaki, M.; Honda, T.; Briones-Nagata, M. P.; Arima, K.; Maeda, H. Rapid and efficient proteolysis for proteomic analysis by protease-immobilized microreactor. *Electrophoresis* **2009,** *30* (18), 3257.

(32) Sun, L.; Zhu, G.; Yan, X.; Mou, S.; Dovichi, N. J. Uncovering immobilized trypsin digestion features from large-scale proteome data generated by high-resolution mass spectrometry. *J. Chromatogr. A* **2014,** *1337*, 40.

(33) Jeng, J.; Lin, M. F.; Cheng, F. Y.; Yeh, C. S.; Shiea, J. Using high-concentration trypsin-immobilized magnetic nanoparticles for rapid in situ protein digestion at elevated temperature. *Rapid Commun. Mass Spectrom.* **2007,** *21* (18), 3060.

(34) Li, Y.; Yan, B.; Deng, C.; Yu, W.; Xu, X.; Yang, P.; Zhang, X. Efficient on-chip proteolysis system based on functionalized magnetic silica microspheres. *Proteomics* **2007,** *7* (14), 2330.

(35) Slysz, G. W.; Lewis, D. F.; Schriemer, D. C. Detection and identification of sub-nanogram levels of protein in a nanoLC-trypsin-MS system. *J. Proteome. Res.* **2006,** *5* (8), 1959.

(36) Slysz, G. W.; Schriemer, D. C. Blending protein separation and peptide analysis through real-time proteolytic digestion. *Anal. Chem.* **2005,** *77* (6), 1572.

(37) Freije, J. R.; Mulder, P. P.; Werkman, W.; Rieux, L.; Niederlander, H. A.; Verpoorte, E.; Bischoff, R. Chemically modified, immobilized trypsin reactor with improved digestion efficiency. *J. Proteome. Res.* **2005,** *4* (5), 1805.

(38) Moore, S.; Hess, S.; Jorgenson, J. Characterization of an immobilized enzyme reactor for on-line protein digestion. *J. Chromatogr. A* **2016,** *1476*, 1.

(39) Liu, Y.; Lu, H.; Zhong, W.; Song, P.; Kong, J.; Yang, P.; Girault, H. H.; Liu, B. Multilayer-assembled microchip for enzyme immobilization as reactor toward low-level protein identification. *Anal. Chem.* **2006,** *78* (3), 801.

(40) Liuni, P.; Rob, T.; Wilson, D. J. A microfluidic reactor for rapid, low-pressure proteolysis with on-chip electrospray ionization. *Rapid Commun. Mass Spectrom.* **2010,** *24* (3), 315.

(41) Xu, F.; Wang, W. H.; Tan, Y. J.; Bruening, M. L. Facile trypsin immobilization in polymeric membranes for rapid, efficient protein digestion. *Anal. Chem.* **2010,** *82* (24), 10045.

(42) Tan, Y. J.; Wang, W. H.; Zheng, Y.; Dong, J.; Stefano, G.; Brandizzi, F.; Garavito, R. M.; Reid, G. E.; Bruening, M. L. Limited proteolysis via millisecond digestions in protease-modified membranes. *Anal. Chem.* **2012,** *84* (19), 8357.

(43) Pang, Y.; Wang, W. H.; Reid, G. E.; Hunt, D. F.; Bruening, M. L. Pepsin-Containing Membranes for Controlled Monoclonal Antibody Digestion Prior to Mass Spectrometry Analysis. *Anal. Chem.* **2015,** *87* (21), 10942.

(44) Cooper, J. W.; Chen, J.; Li, Y.; Lee, C. S. Membrane-based nanoscale proteolytic reactor enabling protein digestion, peptide separation, and protein identification using mass spectrometry. *Anal. Chem.* **2003,** *75* (5), 1067.

(45) Gao, J.; Xu, J.; Locascio, L. E.; Lee, C. S. Integrated microfluidic system enabling protein digestion, peptide separation, and protein identification. *Anal. Chem.* **2001,** *73* (11), 2648.

(46) Ning, W.; Bruening, M. L. Rapid Protein Digestion and Purification with Membranes Attached to Pipet Tips. *Anal. Chem.* **2015,** *87* (24), 11984.

(47) Alves, N. J.; Champion, M. M.; Stefanick, J. F.; Handlogten, M. W.; Moustakas, D. T.; Shi, Y.; Shaw, B. F.; Navari, R. M.; Kiziltepe, T.; Bilgicer, B. Selective photocrosslinking of functional ligands to antibodies via the conserved nucleotide binding site. *Biomaterials* **2013,** *34* (22), 5700.

(48) Fontana, A.; Zambonin, M.; Polverino de Laureto, P.; De Filippis, V.; Clementi, A.; Scaramella, E. Probing the conformational state of apomyoglobin by limited proteolysis. *J. Mol. Biol.* **1997,** *266* (2), 223.

(49) Ahn, J.; Cao, M. J.; Yu, Y. Q.; Engen, J. R. Accessing the reproducibility and specificity of pepsin and other aspartic proteases. *Biochim. Biophys. Acta* **2013,** *1834* (6), 1222.

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

(51) Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M. C.; Yates, J. R., 3rd. Protein analysis by shotgun/bottom-up proteomics. *Chem. Rev.* **2013,** *113* (4), 2343.

(52) Beck, A.; Wagner-Rousset, E.; Ayoub, D.; Van Dorsselaer, A.; Sanglier-Cianferani, S. Characterization of therapeutic antibodies and related products. *Anal. Chem.* **2013,** *85* (2), 715.

(53) Daugherty, A. L.; Mrsny, R. J. Formulation and delivery issues for monoclonal antibody therapeutics. *Adv. Drug Deliv. Rev.* **2006,** *58* (5-6), 686.

(54) Zhao, Y.; Sun, L.; Knierman, M. D.; Dovichi, N. J. Fast separation and analysis of reduced monoclonal antibodies with capillary zone electrophoresis coupled to mass spectrometry. *Talanta* **2016,** *148*, 529.

(55) Chen, S. Rapid protein identification using direct infusion nanoelectrospray ionization mass spectrometry. *Proteomics* **2006,** *6* (1), 16.

(56) Mirzaei, H.; Regnier, F. Enhancing electrospray ionization efficiency of peptides by derivatization. *Anal. Chem.* **2006,** *78* (12), 4175.

(57) Annesley, T. M. Ion suppression in mass spectrometry. *Clin. Chem.* **2003,** *49* (7), 1041.

(58) Li, F.; Vijayasankaran, N.; Shen, A. Y.; Kiss, R.; Amanullah, A. Cell culture processes for monoclonal antibody production. *MAbs* **2010,** *2* (5), 466.

(59) Song, E.; Pyreddy, S.; Mechref, Y. Quantification of glycopeptides by multiple reaction monitoring liquid chromatography/tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **2012,** *26* (17), 1941.

Chapter 4 . Membrane-base proteolytic digestion for protein sequence comparison

Chapters 2 and 3 introduced protease-containing membranes for rapid protein digestion in syringe-pump and spin-column formats. With these platforms, protein digestion and subsequent peptide mapping give important information about the protein sequence. This chapter briefly explores the use of in-membrane digestion to compare two antibodies with similar sequences.

4.1 Introduction

Because of their high specificity and long circulation lifetime in the body, mAbs are the fastest growing class of pharmaceutical drugs.[1-3] However, the high cost of therapeutic mAbs is triggering the development of less-expensive biosimilar antibodies. These biosimilar antibodies have the same primary sequences as their originators, but they are normally expressed in different cell clones.[3] Importantly, different glycosylation and heterogeneities of the biosimilar antibodies may affect their efficacy and safety.[4] Rapid comparison of the biosimilar protein and its originator is crucial for biosimilar antibody research and development. Additionally unintended sequence variants that appear during antibody expression, purification and storage may introduce side effects to the patient. Detailed characterization of these sequence variants is important for quality control.

Characterization of small sequence variations is also important in antibody engineering.[5] Point mutations or additions of one or several amino acids to the primary sequence are common steps in mAb engineering to enhance antibody stability and affinity or prevent antibody aggregation. Robust sequence validation methods will facilitate this process, and protein digestion followed

by MS analysis can easily identify differences between biosimilar and originator antibodies[6] and also verify antibody sequences before and after engineering.

A number of studies used peptide mapping to analyze sequence variation in antibodies. Chen et al. employed Lys-C and trypsin digests and liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) to compare trastuzumab with two biosimilar antibodies. They identified protein posttranslational modifications and mutations in the biosimilar.[7] Hongwei Xie and coworkers developed a rapid method for comparing a candidate biosimilar to an originator mAb using LC-MS technologies.[8] Trypsin in-solution digestion and LC-MS with data-independent acquisition successfully located a two amino acid residue variance in the heavy chain sequence of the biosimilar. Yang et al. compared the HPLC-UV/MS/MS tryptic peptide map of antibodies from four clones and found two sequence variants.[9] Fu et al. identified alanine to serine sequence variants in an IgG4 by LC-MS tryptic peptide mapping.[10] Similarly, Yantao Li and coworkers characterized alanine to valine sequence variants in the Fc region of a nivolumab biosimilar by LC-MS/MS after IdeS and trypsin digestion.[11] Glaser and coworkers used trypsin and Lys-C digestion to confirm formation of the hinge region disulfides after antibody engineering.[12] Rose et al. conducted hydrogen-deuterium exchange mass spectrometry with online pepsin digestion to study the structure change after mutation of Y407 in the CH3 domain of a human IgG,[13] and Rui Gong and coworkers used MS to confirm the engineered human antibody CH2 region.[14] Hussack et al. engineered a single-domain antibody with an additional disulfide bond to increase the protease resistance and thermal stability.[15] They performed peptide mapping with LC-MS analysis of cyanogen bromide and trypsin digests.

These studies demonstrate the power of MS technology for antibody sequence comparison. However, most of the aforementioned research used in-solution protein digestion, which usually

requires extensive sample preparation and long incubation times. Considering that in-membrane protein digestion can occur in as little as 1 min, I developed a workflow that includes proteolysis in pepsin-containing membranes and direct-infusion MS for rapid comparison of antibody sequences. Relative to LC-MS, direct infusion MS uses less sample analysis time. Simple comparison of two direct infusion ESI mass spectra takes less than 30 mins. The differences between the signals in two infusion mass spectra readily reveal sequence variations.

4.2 Experimental

4.2.1 Materials

Nylon membranes (LoProdyne LP, nominal pore size 1.2 μm, 110 μm thickness) were purchased from Pall Corporation. c13C6FR1_ZMapp and c13C6FR1_ZMapp +K antibodies were obtained in their original formulations (Tris-acetate buffer with ⩽ 200 mM Arginine, pH 7.12) as a gift from Dr. Adrian Guthals of Mapp Biopharmaceutical. Pepsin from porcine gastric mucosa (lyophilized powder, 3200-4500 units/mg protein), ammonium bicarbonate (⩾ 99%), iodoacetamide (IAM, ⩾99%), dithiothreitol (DTT, ⩾99.5%), polystyrene sulfonate (PSS, average molecular weight ~70,000), and formic acid (FA, ⩾98%) were purchased from Sigma Aldrich. Sequencing grade modified trypsin was obtained from Promega. NaCl (ACS grade) and HCl (ACS grade) were purchased from CCI. Other chemicals include urea (⩾98%, Invitrogen), tris(2-carboxyethyl) phosphine hydrochloride (TCEP-HCl, >98%, Fluka), acetic acid (HOAc, ACS, Macron Fine Chemicals), and methyl alcohol (anhydrous, MeOH, Macron Fine Chemicals). Solutions were prepared in deionized water (DI water, Milli-Q, 18.2 MΩ·cm at 25 ℃). Amicon ultra 0.5 mL centrifugal filters (MWCO 10 kDa) were employed to desalt

samples before pepsin in-membrane and in-solution digestion, and for buffer exchange before trypsin in-solution digestion. An Eppendorf centrifuge (5415D) was used to conduct in-membrane digestion.

4.2.2 Manufacture of Pepsin-containing Membrane.

The procedure for membrane fabrication is the same as described in Chapter 2.

4.2.3 Digestion of c13C6FR1_ZMapp and c13C6FR1_ZMapp +K antibodies with pepsin- containing membrane

c13C6FR1_ZMapp and c13C6FR1_ZMapp +K antibodies were each diluted in deionized water to give 1 mg/mL antibody stock solutions. Subsequently, 2 μL of 0.1 M HOAc and 2 μL of 0.1 M TCEP-HCl were added to 20 μL of mAb stock solutions. The reactions were incubated at 75 ℃ for 15 min. Buffer exchange with 5% FA employed 3 cycles of centrifugation with Amicon ultra 0.5 mL centrifugal filters (MWCO 10 kDa). About 25 μL of solution remained after each centrifugation, and 475 μL of 5% FA was added prior to the following centrifugation. Residues were diluted to 200 μL with 5% FA to make 0.1 mg/mL antibody pre-digestion solutions. Pepsin in-membrane digestion of two antibodies was conducted by passing 200 μL of each pre-digestion solution through a pepsin-containing membrane at 130 mL/h (syringe pump). Samples were dried in a SpeedVac, and reconstituted in MS buffer (50% MeOH, 49% H_2O and 1% HOAc) for direct infusion MS analysis.

4.2.4 In-solution peptic digestion of c13C6FR1_ZMapp and c13C6FR1_ZMapp +K antibodies

One μL of 1 mg/mL pepsin solution was added to 200 μL of 0.1 mg/mL antibody pre-digestion solution (prepared with buffer exchange as mentioned in the last subsection). The mixtures were incubated at 37 ℃ for 16 h, and the solutions were dried in a SpeedVac and reconstituted in MS buffer for MS analysis.

4.2.5 In-solution tryptic digestion of c13C6FR1_ZMapp and c13C6FR1_ZMapp +K antibodies

Twenty μL of 1 mg/mL mAb stock solution was dried and redissolved in 7 μL of 2 mM TCEP-HCl solution in 0.1% HOAc containing 8 M urea. The mixtures were incubated at 50 ℃ for 10 min prior to addition of 7 μL of 20 mM IAM in 2 M NH_4HCO_3 containing 8 M urea, and incubation in the dark for 30 min. Finally, 6 μL of 30 mM DTT in 100 mM NH_4HCO_3 containing 8 M urea was added followed by incubation in the dark for 20 min to quench the IAM. After reduction and alkylation, the residual solutions underwent buffer exchange with 10 mM NH_4HCO_3 using 3 centrifugation cycles. About 25 μL of solution remained after each centrifugation, and 475 μL of 10 mM NH_4HCO_3 was added prior to the following centrifugation. Finally, residues were diluted to 200 μL with 10 mM NH_4HCO_3 to make 0.1 mg/mL alkylated antibody pre-digestion solutions. Then, 5 μL of 0.2 μg/μL sequencing grade modified trypsin solution was added to 200 μL of the alkylated antibody solution prior to incubation at 37 ℃ for 16 h. The reaction was quenched by adding 5 μL of acetic acid. Samples were then dried with a SpeedVac before reconstitution and infusion MS analysis.

4.2.6 Mass Spectrometry and Data Analysis

The in-membrane digests and in-solution digests were dried with a SpeedVac and reconstituted in MS buffer within 1 day. Forty μL of each sample was loaded into a Whatman multichem 96-well plate (Sigma−Aldrich) and sealed with Teflon Ultrathin Sealing Tape (Analytical Sales and Services, Prompton Plains, NJ). An Advion Triversa Nanomate nanoelectrospray ionization (nESI) source (Advion, Ithaca, NY) was used to introduce the sample into a high-resolution accurate mass Thermo Fisher Scientific LTQ Orbitrap Velos mass spectrometer (San Jose, CA) that was equipped with a dual pressure ion trap, HCD cell, and ETD. The spray voltage and gas pressure were set to 1.4 kV and 1.0 psi, respectively. The ion source interface had an inlet temperature of 200 ℃ with an S-Lens value of 57%. High-resolution mass spectra were acquired in positive ionization mode across the *m/z* range of 300−1800, using the FT analyzer operating at a mass resolving power of 100,000. Spectra were the average of 100 scans. Signals with >1% of the highest peak intensities and S/N>3 were analyzed. Peptide identification was performed manually using comparison to masses obtained with ProteinProspector (v 5.14.1, University of California, San Francisco, CA). Mass tolerance was set to 10 ppm.

4.3 Results and discussion

c13C6FR1_ZMapp (Z) and c13C6FR1_ZMapp +K (ZK) antibodies have similar primary sequences. The only difference is ZK has an extra Lysine on the light chain. Figure 4.1 shows part of the antibody sequences of Z and ZK to illustrate the sequence difference.

Z Light Chain …KLELRTVAAPSVF…

ZK Light Chain …KLELKRTVAAPSVF…

Figure 4.1. Parts of the sequences of 13C6FR1_ZMapp (Z) and c13C6FR1_ZMapp +K (ZK) antibodies.

Interestingly, the extra Lysine is next to an Arginine. In this case, trypsin digestion is not the best option for identifying the extra lysine because the theoretical tryptic digest will contain the short peptide LELK. Even with a missed cleavage, the resulting peptide LELKR will still be small. Considering that this peptide contains two basic residues, it can carry as many as 3 protons. The *m/z* of the peptide will thus fall into the low *m/z* region (<200 *m/z*), which contains a large amount of noise.

4.3.1 Digestion of c13C6FR1_ZMapp and c13C6FR1_ZMapp +K antibodies in a pepsin-containing membrane

Chapters 2 and 3 showed that digestion using pepsin-containing membranes enables rapid mAb characterization.[16] The workflow is straightforward (Figure 4.2). In this particular application, we added a desalting step because the high salt concentration in the formulation might suppress the ionization of peptides.[17]

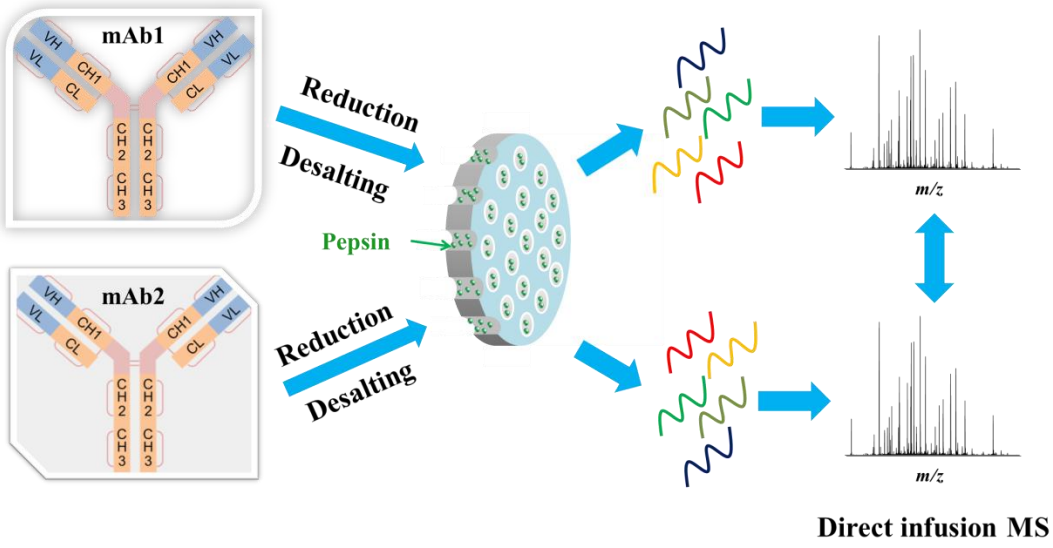


Figure 4.2. Workflow for comparison of two antibodies using pepsin-containing membranes for proteolysis.

Digestion employed flow through the membrane at 130 mL/h, which corresponds to a 3 msec residence time with the assumption of 50% membrane porosity. The short residence time generates large peptides with missed cleavages, and passage of 200 µL of protein solution through the membrane requires < 30 sec.

We did not conduct peptide mapping of the two antibodies because our goal is simply to find the differences. If a peptide signal appears in one mass spectrum, but not another, we mark it as a peptide that potentially contains the different sequence. Figure 4.3 shows an example. The +2 peptide with an *m/z* of 474.2694 appears in the spectrum of Z but not ZK, while a +2 peptide with an *m/z* of 538.3162 is present in the spectrum of ZK but not Z.

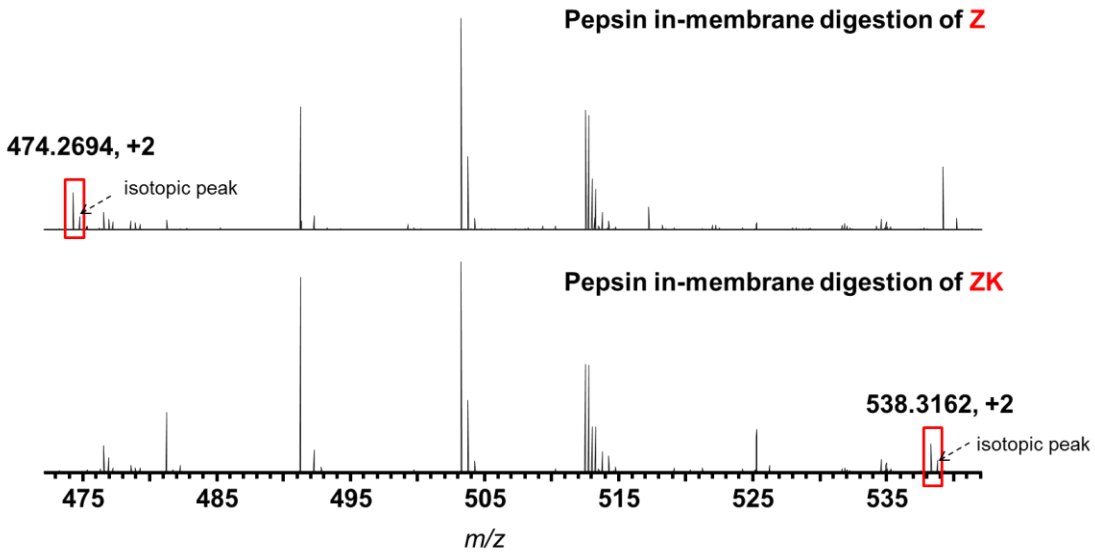


Figure 4.3. Comparison of part of the mass spectra of in-membrane peptic digests of Z (top) and ZK (bottom).

The deconvoluted mass difference between these two signals is 128.0936, which is the mass of a Lysine residue. Comparison of the peptide masses with the original antibody sequences identifies the *m/z* of 474.2694 with the Z light-chain peptide 107-115 (L107-115), and the *m/z* of 538.3162 with ZK L107-116. MS gives information about the peptide mass, and tandem mass spectrometry (MS/MS) provides peptide sequence information. Figure 4.4 shows the MS/MS spectra of these two peptides.

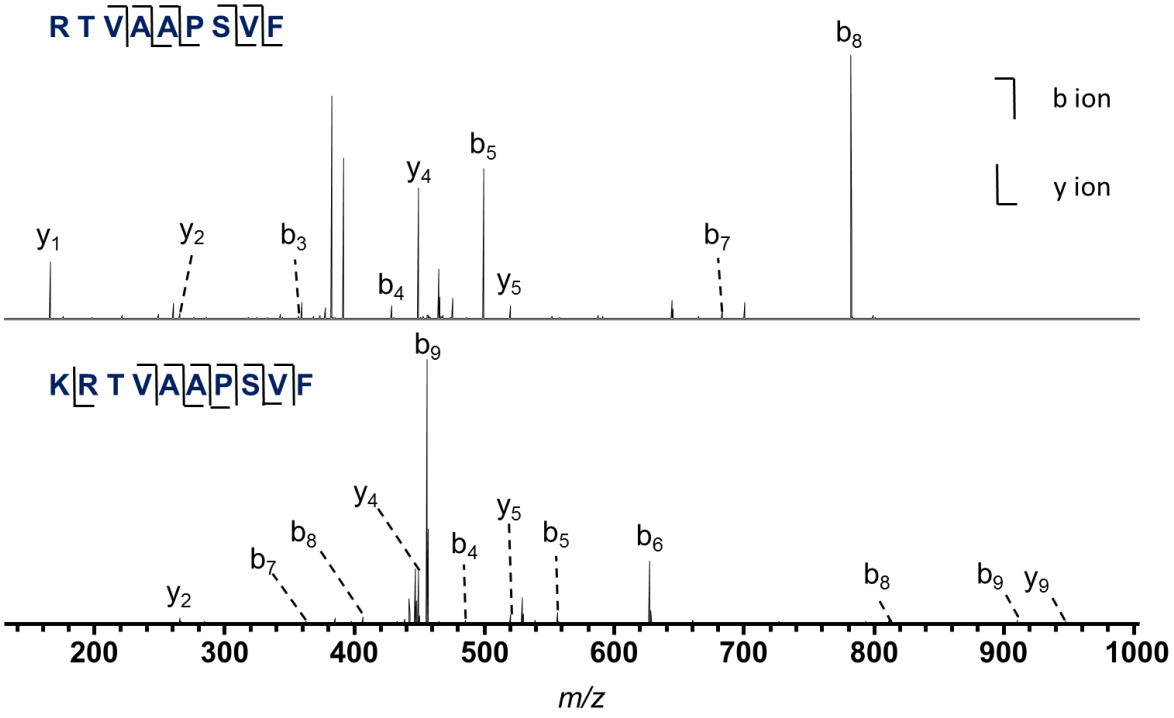


Figure 4.4. Comparison of part of the MS/MS spectra of L107-115 from Z (top) and L107-116 from ZK (bottom).

These spectra clearly show that the extra Lysine is located at the N-terminus of L107-116 from ZK, because the y_9 ion of L107-116 has an *m/z* value of 947.5286 and the $[M+H]^+$ *m/z* value of L107-115 from Z is 947.5310. It is not surprising that the MS/MS spectra show more b ions than y ions because of the basic residues at the N-terminus.

Similarly, I compared the mass spectra of Z and ZK across the whole mass range. Table 4.1 presents all the identified differences from peptic digestion of Z and ZK.

Table 4.1. MS signals that correspond to differences in that analysis of peptic in-membrane digestion of Z and ZK.

m/z	Charge state	Peptide from Z	Peptide from ZK	Amino Acid
474.2694	+2	✓	×	L107-115
538.3162	+2	×	✓	L107-116
701.6310	+4	×	✓	L91-116
762.6613	+4	×	✓	L107-135
791.4129	+4	×	✓	L88-116
828.1799	+4	×	✓	L87-116
868.7954	+3	✓	×	L107-131
868.9460	+4	×	✓	L86-116
973.8495	+3	✓	×	L107-134
1057.1447	+5	×	✓	L87-135
1061.2066	+3	✓	×	L87-115
1089.9587	+5	×	✓	L86-135

Remarkably, all of the signal differences identified from Z and ZK digestion cover the modification site. For example, L107-135 from ZK and L107-134 from Z as well as L87-116 from ZK and L87-115 from Z are also pairs of peptides with a mass difference corresponding to a Lysine. These results give unambiguous identification of the modification site.

4.3.2 In-solution peptic digestion of c13C6FR1_ZMapp and c13C6FR1_ZMapp +K antibodies

For comparison, we also conducted in-solution pepsin digestion of the Z and ZK antibodies. Most of the peptide signals from in-solution digestion fall in the range of 400-900 *m/z*, whereas in-membrane digestion gives some larger peptides to cover the 400-1500 *m/z* range. Figure 4.5 gives an overview comparison of in-solution and in-membrane digestion of Z.

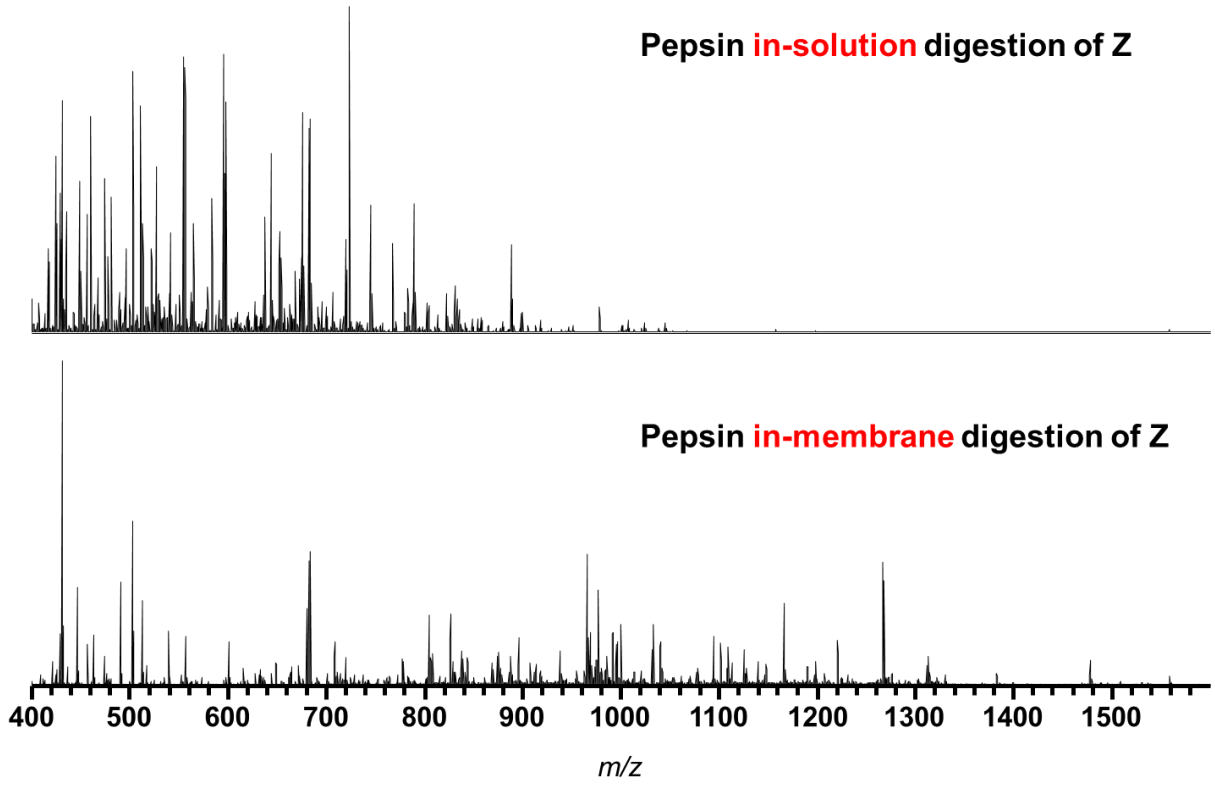


Figure 4.5. Mass spectra of in-solution (top) and in-membrane (bottom) digests of Z. The maximum intensity is 1.02 x 10E6 in the top spectrum and 8.64 x 10E5 in the bottom spectrum.

When comparing the mass spectra of in-solution peptic digests of Z and ZK, the only difference is the absence of *m/z* 474.2696 (+2) and the addition of *m/z* 538.3169 (+2) in the spectrum of ZK. We did not see any of the other differences found with in-membrane digestion. Both in-solution and in-membrane peptic digestion show the difference between the two antibodies, but in-membrane digestion yields several signals that give more confidence in the identification. Moreover, membranes require only 1 min for digestion.

4.3.3 In-solution tryptic digestion of c13C6FR1_ZMapp and c13C6FR1_ZMapp +K antibodies

In-solution trypsin digestion is the standard method for bottom-up proteomics. After reduction and alkylation, samples normally incubate in trypsin solutions at 37 ℃ overnight. We conducted tryptic in-solution digestion and used the strategy described above to compare the mass spectra of Z and ZK but didn't find any differences in signals from multiply charged peptides in the two spectra. However, when I used the theoretical mass of LELR to check the Z spectrum, I found one single charge state signal at *m/z* 530.3301. Similarly, I found the peak correspond to singly charged LELK in the ZK spectrum (*m/z* 502.3240, +1). Figure 4.6 shows the comparison of analyses of tryptic in-solution digestion of Z and ZK. The signal of LELKR did not appear, presumably because of complete tryptic cleavage of the K-R bond. The signals of single-charged peptides appear in a noisy region, and it is not easy to differentiate whether a signal is from a peptide or a noise.

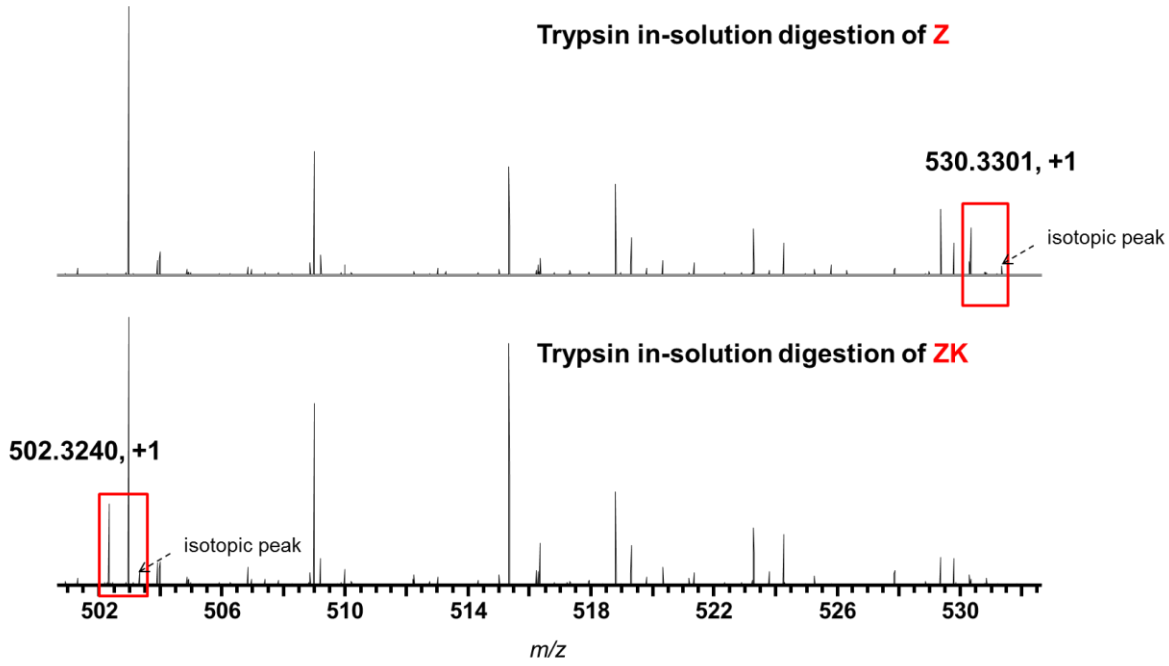


Figure 4.6. Part of the MS spectra of tryptic in-solution digests of Z (top) and ZK (bottom). This region shows different singly charged peaks.

4.4. Conclusion

Comparison of two antibody sequences is important in antibody characterization and quality control. In-membrane digestion is a powerful method for rapid identification or verification of a modification site. Digestion of 200 µL of protein solution requires as little as 30 sec, which prevents introduction of possible PTMs during long sample incubation times. Because of different missed cleavages, rapid peptic in-membrane digestion yields multiple peptides that contain the modification site. Compared with peptic in-solution digestion, in-membrane digestion gives more evidence of sequence differences and thus provides unambiguous identification of the modification. Tryptic in-solution digestion is not appropriate in the particular case of differentiating the Z and ZK antibodies, because the singly charged peptide ion

is hard to identify. In conclusion, peptic in-membrane antibody digestion is a powerful method for identifying sequence modifications. Direct infusion MS with fast manual interpretation (less than 1 h) saves sample analysis time compared with LC-MS/MS methods.

4.5 Acknowledgement

We gratefully acknowledge the US National Science Foundation (CHE-1152762 and CHE-1506315) for funding this work. We thank Dr. Adrian Guthals of Mapp Biopharmaceutical for providing c13C6FR1_ZMapp and c13C6FR1_ZMapp +K antibodies. We also thank Dr. Todd Lydic from the Molecular Metabolism and Disease Collaborative Mass Spectrometry Core for helping analyze the samples.

REFERENCES

REFERENCES

(1) Weiner, G. J. Building better monoclonal antibody-based therapeutics. *Nat. Rev. Cancer* **2015,** *15* (6), 361.

(2) Weiner, L. M.; Surana, R.; Wang, S. Monoclonal antibodies: versatile platforms for cancer immunotherapy. *Nat. Rev. Immunol.* **2010,** *10* (5), 317.

(3) Beck, A.; Sanglier-Cianferani, S.; Van Dorsselaer, A. Biosimilar, biobetter, and next generation antibody characterization by mass spectrometry. *Anal. Chem.* **2012,** *84* (11), 4637.

(4) Beck, A.; Debaene, F.; Diemer, H.; Wagner-Rousset, E.; Colas, O.; Van Dorsselaer, A.; Cianferani, S. Cutting-edge mass spectrometry characterization of originator, biosimilar and biobetter antibodies. *J. Mass Spectrom.* **2015,** *50* (2), 285.

(5) Carter, P. Improving the efficacy of antibody-based cancer therapies. *Nat. Rev. Cancer* **2001,** *1* (2), 118.

(6) Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M. C.; Yates, J. R., 3rd. Protein analysis by shotgun/bottom-up proteomics. *Chem. Rev.* **2013,** *113* (4), 2343.

(7) Chen, S. L.; Wu, S. L.; Huang, L. J.; Huang, J. B.; Chen, S. H. A global comparability approach for biosimilar monoclonal antibodies using LC-tandem MS based proteomics. *J. Pharm. Biomed. Anal.* **2013,** *80*, 126.

(8) Xie, H.; Chakraborty, A.; Ahn, J.; Yu, Y. Q.; Dakshinamoorthy, D. P.; Gilar, M.; Chen, W.; Skilton, S. J.; Mazzeo, J. R. Rapid comparison of a candidate biosimilar to an innovator monoclonal antibody with advanced liquid chromatography and mass spectrometry technologies. *MAbs* **2010,** *2* (4), 379.

(9) Yang, Y.; Strahan, A.; Li, C.; Shen, A.; Liu, H.; Ouyang, J.; Katta, V.; Francissen, K.; Zhang, B. Detecting low level sequence variants in recombinant monoclonal antibodies. *MAbs* **2010,** *2* (3), 285.

(10) Fu, J.; Bongers, J.; Tao, L.; Huang, D.; Ludwig, R.; Huang, Y.; Qian, Y.; Basch, J.; Goldstein, J.; Krishnan, R.et al. Characterization and identification of alanine to serine sequence variants in an IgG4 monoclonal antibody produced in mammalian cell lines. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2012,** *908*, 1.

(11) Li, Y.; Fu, T.; Liu, T.; Guo, H.; Guo, Q.; Xu, J.; Zhang, D.; Qian, W.; Dai, J.; Li, B.et al. Characterization of alanine to valine sequence variants in the Fc region of nivolumab biosimilar produced in Chinese hamster ovary cells. *MAbs* **2016,** *8* (5), 951.

(12) Glaser, S. M.; Hughes, I. E.; Hopp, J. R.; Hathaway, K.; Perret, D.; Reff, M. E. Novel antibody hinge regions for efficient production of CH2 domain-deleted antibodies. *J. Biol. Chem.* **2005,** *280* (50), 41494.

(13) Rose, R. J.; van Berkel, P. H.; van den Bremer, E. T.; Labrijn, A. F.; Vink, T.; Schuurman, J.; Heck, A. J.; Parren, P. W. Mutation of Y407 in the CH3 domain dramatically alters glycosylation and structure of human IgG. *MAbs* **2013,** *5* (2), 219.

(14) Gong, R.; Vu, B. K.; Feng, Y.; Prieto, D. A.; Dyba, M. A.; Walsh, J. D.; Prabakaran, P.; Veenstra, T. D.; Tarasov, S. G.; Ishima, R.et al. Engineered human antibody constant domains with increased stability. *J. Biol. Chem.* **2009,** *284* (21), 14203.

(15) Hussack, G.; Hirama, T.; Ding, W.; Mackenzie, R.; Tanha, J. Engineered single-domain antibodies with high protease resistance and thermal stability. *PLoS One* **2011,** *6* (11), e28218.

(16) Pang, Y.; Wang, W. H.; Reid, G. E.; Hunt, D. F.; Bruening, M. L. Pepsin-Containing Membranes for Controlled Monoclonal Antibody Digestion Prior to Mass Spectrometry Analysis. *Anal. Chem.* **2015,** *87* (21), 10942.

(17) Metwally, H.; McAllister, R. G.; Konermann, L. Exploring the mechanism of salt-induced signal suppression in protein electrospray mass spectrometry using experiments and molecular dynamics simulations. *Anal. Chem.* **2015,** *87* (4), 2434.

Chapter 5 . Summary and future work

5.1 Research summary

This dissertation describes the development of protease-containing membranes for rapid, controlled protein digestion prior to MS analysis. Most of my work focuses on monoclonal antibody digestion. I conducted in-membrane digestion using a prototype setup with a syringe pump, as well as with a novel spin membrane. In-membrane digestion removes the bottleneck of conventional in-solution digestion, which requires extensive time and labor, in workflows for "bottom-up" or 'middle-down" protein analysis.

Chapter 2 describes the use of pepsin-containing membrane as controlled reactors for monoclonal antibody digestion. Pepsin is an inexpensive protease that enables membrane digestion in acidic conditions, which avoids the need for antibody alkylation and minimizes protein modification during digestion. Layer-by-layer adsorption of PSS and pepsin in nylon membranes generates an IMER, and pepsin immobilization in membrane pores yields a high local enzyme concentration (~ 70 mg per mL of membrane) that enables digestion of 100 μL of antibody solution in less than one minute. Moreover, in-membrane digestion using a high flow rate (130 mL/h) yields relatively large peptides (5-15 kDa) that cover the entire antibody sequence. As needed, digestion with different flow rates can enhance sequence coverage. Pepsin digestion followed by infusion MS analysis gives nearly 100% sequence coverage for both a Waters^{TM} antibody and Herceptin. Additionally, MS analysis of the proteolytic peptides reveals sites for oxidation, deamidation and N-terminal pyroglutamic acid formation, as well as glycosylation patterns. Furthermore, CID, HCD and ETD MS/MS of the light-chain peptides generated from in-membrane digestion cleave 99% of the amino acid bonds in the light chain.

For comparison, “Top-down” analysis of the entire light chain by MS/MS methods shows a sequence coverage of only 55%. With minimal sample preparation time, membrane digestion leads to high peptide and sequence coverages for identification of PTMs by MS.

Chapter 3 introduces a novel platform for in-membrane digestion. Membrane-based digestion methods developed by previous group members required a syringe pump and Upchurch fittings that may limit widespread adoption of the technique. I employed a spin membrane to greatly simplify the method, and these membranes will likely be commercially available in 2017. Protease-containing membranes inserted into spin columns enable digestion using simple centrifugation to pass the protein solutions through the membrane. Centrifuging 100-200 µL of pretreated protein solution through a membrane requires 1 min or less. We tested the performance of spin digestion on apomyoglobin and four commercialized antibodies (Herceptin, Avastin, Rituxan and Vectibix). Direct infusion analysis of peptic and tryptic digests from these proteins gives nearly 100% sequence coverage. Protein PTMs, such as glycosylation, C-terminal Lysine clipping, and pyroglutamate formation, can be easily identified. Variation of the spin rates yields different proteolytic peptide sizes for apomyoglobin. Fast spin rates generate large peptides with more missed cleavages. LC-MS/MS analysis of antibody tryptic digests followed by MaxQuant data analysis reveals 100% sequence coverages for all the antibody light chains, and 75.1% to 98.4% coverage for the heavy chains. Compared with in-solution digestion, tryptic spin digestion gives higher sequence coverages and more unique peptides.

Chapter 4 explores the application of a pepsin-containing membrane to comparison of two antibodies with similar sequences. One antibody has an extra Lysine. In-membrane peptic digestion and MS analysis of the two antibodies give multiple evidences for the sequence variation. Analysis of the MS signals of multiply charged peptides that appear exclusively in the

mass spectrum of only one of the antibodies shows that all of these peptides contain the sequence variation region. Tandem mass spectrometry analysis of these peptides further identifies the location of the variation. In contrast, in-solution peptic digestion generates only one peptide with the sequence difference, whereas in-solution trypsin digestion gives a four-amino acid peptide that shows a +1 charge state in the mass spectrum and is difficult to distinguish from the MS noise. Rapid in-membrane digestion (<30 sec) with direct infusion MS analysis is a time-saving workflow for protein sequence comparison.

5.2 Future work

5.2.1 Limited proteolysis in protease-containing membranes to interrogate protein higher order structure

The structure of a protein essentially determines its function. Generally, proteins have four levels of structure: primary, secondary, tertiary and quaternary. The primary level is the linear amino acid sequence, and secondary structure refers to local conformations such as α-helices and β-sheets. Tertiary structure is the three-dimensional shape of a protein molecule, whereas the quaternary level refers to the geometry of complexes that contain multiple protein subunits. Studies of the structures of proteins are crucial for understanding their function and interaction with other proteins, and limited digestion can facilitate the identification of protein conformational changes.[1]

In-membrane digestion is a promising method for studying protein higher order structures. For this aim, trypsin-containing membranes are preferable to pepsin-modified membranes, because peptic digestion occurs at acidic pH, which will partially denature the protein. Former group

member Dr. Yujing Tan conducted limited membrane-based proteolysis to locate the flexible region of Root Hair Defective 3 (RHD3) protein.[2] In-membrane digestion of RHD3 with trypsin and chymotrypsin under nondenaturing conditions reveals a region around R672 that is highly flexible because peptide 673-691 (protein full length 1-691) shows the strongest signal in the mass spectrum of the protein after rapid in-membrane digestion.

Chapter 4 provides an example of comparing protein primary sequences using protease-containing membranes and direct-infusion MS. Similarly, we could employ in-membrane digestion to study changes in protein higher order structure.[3-5] Theoretically, digestion of the same protein with two different conformations will give different digestion patterns because protein conformation affects the accessibility to the proteolytic enzymes. To explore this possibility, we performed in-membrane digestion on apo and holo proteins with different conformations.

Dr. Geiger's group at Michigan State University recently studied the structure of cellular retinol binding proteins (CRBPs). They found that the crystal structure of human CRBPII changes after binding of retinal. However, data from X-ray diffraction represent the protein structure in the solid state. MS has been widely used to study protein structures in solution.[6,7] Dr. Geiger provided us with two CRBPII proteins, one is the apo domain swapped (DS) dimer, and the other is the holo DS dimer, the apo DS dimer with a bound retinal. For this particular case, we conducted in-membrane digestion of the two proteins with a membrane containing trypsin immobilized through a covalent linkage. Trypsin is highly active, and covalent immobilization of this enzyme decreases its activity, which should facilitate limited trypsinolysis. We fabricate such IMERs by adsorption of poly (acrylic acid) (PAA) in membranes followed by activation of –COOH groups with NHS/EDC and covalent coupling via amide linkages. Direct infusion mass

spectra of tryptic digests (1.5-sec residence times) of the apo and holo proteins showed different digestion patterns. Figure 5.1 represents the original mass spectra.

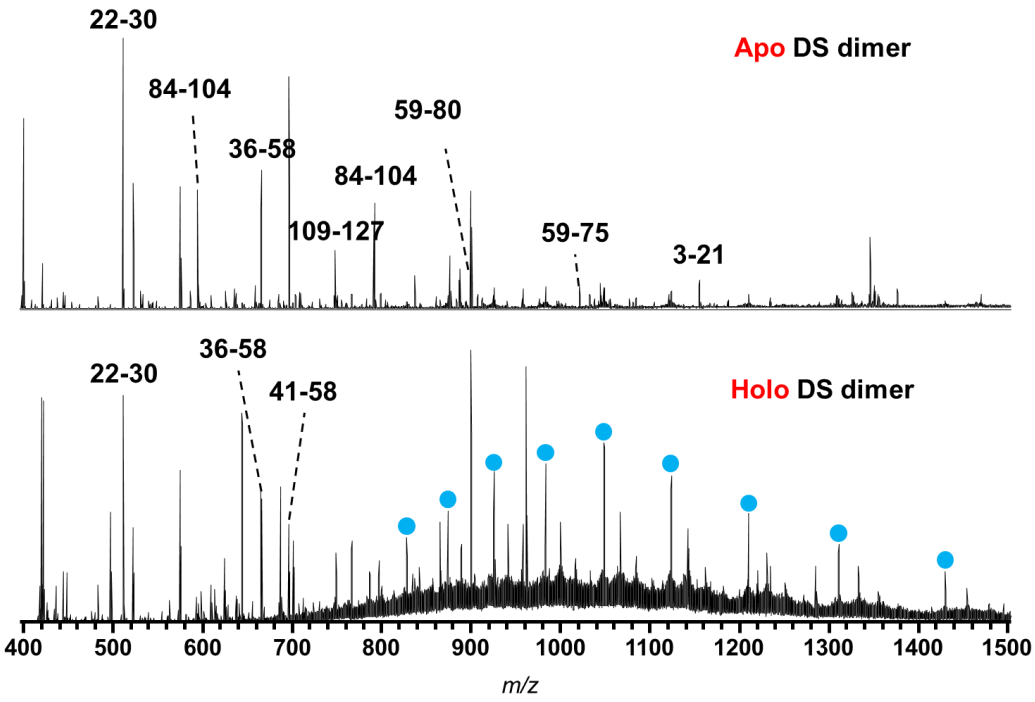


Figure 5.1. Direct infusion MS spectra of in-membrane tryptic digests (1.5-sec residence times) of Apo and Holo DS CRBPII dimers.

In general, the apo DS dimer undergoes more proteolytic digestion than the holo DS dimer. Intact holo DS dimer signals dominate the mass spectrum of this protein, which suggests that the holo DS dimer has a rigid structure that resists trypsin proteolysis. Apo monomer is a protein with 133 amino acids. Tryptic peptides, 3-21, 22-30, 36-58, 59-80, 84-104, and 109-127, nearly cover the apo sequence. However, only three peptides, 22-30, 36-58 and 41-58, can be identified in the MS spectrum from the holo tryptic digest. These evidences are consistent with a conformational change in the apo DS dimer after binding retinal molecules. N59 is a key amino acid for this conformational change because amino acids after 59 are missing. Future work

should digest the two proteins with a membrane containing electrostatically immobilized trypsin. Electrostatic immobilization leads to higher enzyme activity than covalent binding and should provide more digestion on the holo DS dimers. Also, we can conduct the covalent/electrostatic trypsin digestion of the two proteins under different flow rates to potentially identify the most flexible regions of these dimers.

5.2.2 Polyclonal antibody digestion by protease-containing membranes

Monoclonal antibodies and related products represent the largest portion of the biologic therapeutics market.[8] Antibody engineering technology is widely used to enhance ligand binding affinity, reduce immunogenicity, and optimize *in vivo* half-life.[9] However, these engineered products do not necessarily have improved clinical efficacy. Motavizumab, for example, is engineered from Palivizumab and shows 75 times greater affinity for respiratory syncytial virus F protein than its parent antibody. However, a phase-2, randomized, double-blind safety and pharmacokinetic assessment of these two antibodies showed similar results on high-risk children.[10]

To overcome such challenges, researchers began developing multi-specific antibodies (antibodies that can bind two distinct epitopes), oligoclonal cocktails (mixtures of two or three monoclonal antibodies), and recombinant polyclonal antibodies (a single master cell line express a single light chain and up to five heavy chains).[11] An approved biclonal combination of Trastuzumab (Herceptin) and Pertuzumab (Perjeta), which target different epitopes on the HER-2 growth factor receptor, synergistically inhibited the survival of BT474 cells.[12] Similar to monoclonal antibodies, these polyclonal antibodies require extensive characterization prior to distribution to patients.

Characterization is more difficult for polyclonal antibodies than for mAbs. Top-down analysis will be challenging, and MS of antibody subunits might reveals antibody PTMs, but specifying the PTM sites is still challenging. Bottom-up analysis of polyclonal antibodies yields tryptic peptides from multiple antibodies, and it is hard to identify the origin of a single peptide because these antibodies share a large portion of similar sequences. Rapid in-membrane peptic digestion has the unique advantages of generating large peptides whose origin may be easier to determine.

To investigate in-membrane digestion for analysis of polyclonal antibodies, I digested a mixture of four equimolar antibodies with a pepsin-containing membrane using a residence time of 3 msec. Figure 5.2 shows the mass spectrum of the digest.

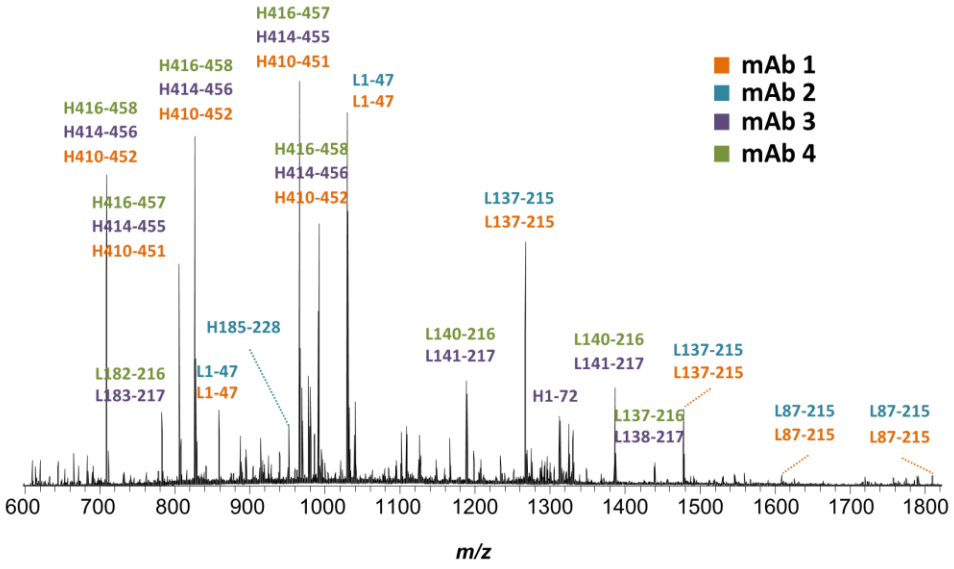


Figure 5.2. Mass spectra of an in-membrane peptic digest (3-msec residence time) of a cocktail of four antibodies. Multiple labels represent all the possible origins of an identified peptide. Not all identified peptide are labeled with sequences.

The 3-mec peptic in-membrane digest gives a large amount of peptides, and most of the signals in Figure 5-2 stem from the four antibodies. Rapid peptic digestion yields large peptides, for

example, L87-215 from mAb2. Not surprisingly, most of the identified signals were from antibody light and heavy chain constant regions, however, because these four antibodies share a large percentage of the same sequence, especially in the constant region. Ion suppression may occur for the peptides from light- and heavy-chain variable regions. LC-MS/MS analysis of the digests might be a better choice, not only because manual data interpretation is very time-consuming in this case, but also because separation of the peptides would solve the ion-suppression problem. Future work should focus on developing a straightforward data analysis workflow and investigating the LC-MS/MS analysis method.

5.2.3 *De novo* antibody sequencing

Discovery and development of a therapeutic monoclonal antibody normally begins with screening for high affinity antibodies and subsequent sequencing and production of these targets. Identification of novel antibodies normally depends on DNA sequencing or hybridoma cloning of peripheral B cells.[13] However, the antibody DNA does not necessarily predict the antibody primary sequence because of V(D)J recombination during B cell maturation, and somatic hypermutation during B cell affinity maturation results in extremely diverse sequences.[14] Hybridoma cloning of peripheral B cells is a powerful method, but not all antibodies are produced by peripheral B cells. Long-lived plasma B cells, for instance, reside in bone marrow where sampling will be challenging. *De novo* antibody sequencing is necessary when DNA of the antibody or a B cell is not available.

MS-based protein sequencing methods have developed greatly over the past 30 years. Biemann et al. first obtained the sequence of thioredoxin in1987 by sequencing 14 tryptic peptides and several peptides from Staphylococcus aureus protease.[15] The Biemann group further employed

trypsin digestion and thermolysin digestion on glutaredoxin and deduced the sequence using CID mass spectra.[16] Peptides generated from one protease overlap with peptides from other proteases, and thus, MS/MS of these peptides provides important information about the peptide arrangement. Nuno Bandeira et al. developed the shotgun protein sequencing (SPS) method to detect, score and interpret overlaps between uninterpreted MS/MS spectra.[17] Bandeira et al. further developed a comparative shotgun protein sequencing method (CSPS).[18] Alignment of contigs from SPS to a homologous sequence gives the primary sequence of a monoclonal antibody. Guthals et al. developed Meta-SPS and extended the *de novo* contig sequences to 100 amino acids with 97% accuracy.[19] This group also explored *de novo* analysis of MS/MS triplets (CID/HCD/ETD) from overlapping peptides.[20] Liu et al. developed Complete Homology-Assisted MS/MS Protein Sequencing (CHAMPS) for *de novo* sequencing. In CHAMPS, bottom-up MS/MS was first employed to give *de novo* peptide sequences, and then these sequences were aligned to a homology protein sequence to find overlapping sequences, followed by final assembly to get the protein sequence. Liu et al. further developed top-down and bottom-up MS-based protein *de novo* sequencing (TBNovo) method, and got high sequence coverage and high sequence accuracy.[21] Recently, Adrian Guthals and coworkers reported using semiautomated customized tools to sequence polyclonal antibodies. [13]

One feature of pepsin in-membrane digestion is that we can control the peptic peptide sizes. The length of proteolytic peptides varies with the residence time of the protein solution in the membrane, i.e. shorter residence times generate longer peptides. We obtained middle-down sized peptides (4-9 kDa) in membrane-based digestions of antibodies when using a residence time of 3 msec. We also produced bottom-up sized peptides (0.3-2 kDa) when doing 3-sec digestion. Larger peptides give the arrangement of the small peptides, and small peptides give better

fragmentation performance compared to larger peptides. Both of the membrane digestion methods (short residence time and long residence time) gave 100% peptide coverage of the antibody light chain we examined, which may makes *de novo* antibody sequencing possible.

Below I illustrate the strategy and results of the combination middle-down/bottom up procedure. I deconvoluted the 3 msec digestion MS file using Xtract with m/z from 600 to 1400, and minimum S/N set for 3. After analyzing the deconvoluted spectra, I found that four masses add (accounting for water in hydrolysis and protons) to give the mass of the light chain (M+H=24183.7). The four peptides are A 4204.1447, B 5234.5729, C 5649.8075, and D 9152.2347 (Numbers are monoisotopic M+H values). Peptides E 8177.0456, F 6892.4733, and D also add up to match the mass of the light chain. See Figure 5.3 below.

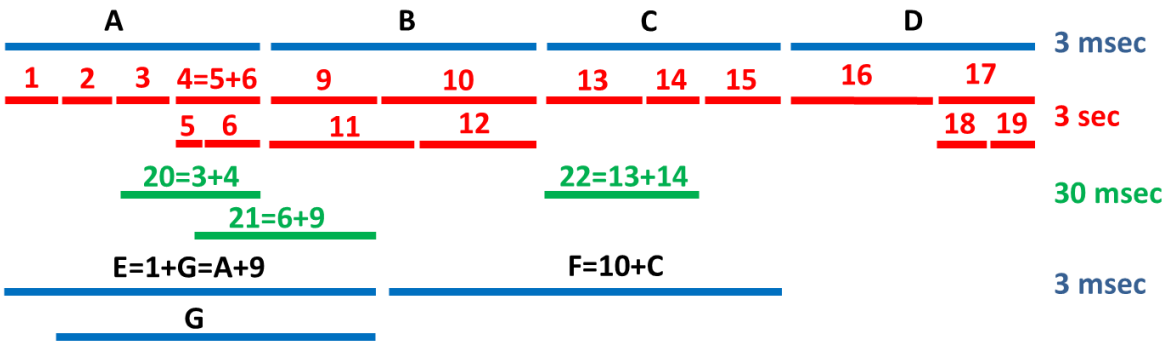


Figure 5.3. Example of arranging peptides using relationships between their masses.

The sum of two or more masses can give the mass of another peptide. Below I show relations among masses of peptides. Letters represent peptides from the 3-msec digestion, starred numbers or letters represent peptides from 30-msec digestion, and numbers represent peptides from the 3-sec digestion. Regardless of a star, peptides with the same number or letter have the same sequence, e.g. 3=3*.

A+B+C+D=Light Chain (5-1)

E+F+D=Light Chain (5-2)

1+G=E (5-3)

1+2+3+4=A (5-4)

1+2+3+5+6=A (5-5)

5+6=4 (5-6)

9+10=B (5-7)

11+12=B (5-8)

13+14+15=C (5-9)

16+17=D (5-10)

18+19=17 (5-11)

10+13+14+15=F (5-12)

A+9=E (5-13)

2+3+4+9=G (5-14)

3*+4*=20* (5-15)

6*+9*=21* (5-16)

13*+14*=22* (5-17)

Note that the numbers of the peptides do not necessarily represent their order in the sequence. Next, we used these relations to try to order the peptides.

Starting from (5-13) and (5-4), we have the order 1, 2, 3, 4, **9** or **9**, 1, 2, 3, 4. The bold indicates this is a fixed position. We do not know the order of 1,2,3,4. From (5-14), we can deduce that 9 is connected to 2, 3, 4 and not 1. Thus, we have **1**, 2, 3, 4, **9** or **9**, 2, 3, 4, **1**. (These peptides are not necessarily at the beginning of the sequence.) From (5-7), we deduce the sequence **1**, 2, 3, 4, **9**, **10** or **10**, **9**, 2, 3, 4, **1**. Next, we look at (5-12), which shows we have **1**, 2, 3, 4, **9**, **10**, 13, 14, 15 or 13, 14, 15, **10**, **9**, 2, 3, 4, **1**. Also, looking at these sequences we either have ABC or CBA. Also, D could be at the beginning or the end.

Continuing on, from (5-15), we learn that 3 and 4 are connected. From (5-16), 6 and 9 are connected. From (5-6), 6 is part of 4, so 9 is now connected to 4. Thus, we now have **1**, **2**, **3**, **4**, **9**, **10**, 13, 14, 15 or 13, 14, 15, **10**, **9**, **4**, **3**, **2**, **1**. We could easily fragment 1 to distinguish these possibilities. From (5-17), we see that 13 and 14 are connected. Assuming 1 is at the beginning (determined from fragmentation), we now have **1**, **2**, **3**, **4**, **9**, **10**, **13**, **14**, 15 or **1**, **2**, **3**, **4**, **9**, **10**, **14**, **13**, 15 or **1**, **2**, **3**, **4**, **9**, **10**, 15, **14**, **13** or **1**, **2**, **3**, **4**, **9**, **10**, 15, **14**, **13**. If we compare fragmentation of 13 and 14 with fragmentation of C, we can figure out where 15 is. This will give us the alignment of the entire sequence of ABC as well as the smaller peptides. .

We also have peptide D. Using the above equations, we can determine an order of 16, 17 or 16, 18, 19. We could also use what we know about the antibody. This is a mouse antibody. D is entirely in the constant region, so we already know its sequence and that it is on the C-terminus. Peptide 15 is also in the constant region, and we know it is connected to D. This finally leads us to **1**, **2**, **3**, **4**, **9**, **10**, **13**, **14**, **15**, **16**, **17**.

We can align the peptic peptides of the antibody light chain just using the mass relations among the peptides. Then, future work should focus on whether we could perform *de novo* sequencing of the small peptides using CID, HCD and ETD MS/MS data.

5.2.4 Glycosidase-containing membrane for haptoglobin deglycosylation

Glycosylation is an important protein PTM that enhances protein functional diversity and influences biological activities.[22] Glycomics studies are challenging compared with proteomics and genomics due to the high diversity of sugars and their presence as an appendage on proteins. Current strategies for glycomics include analysis of glycopeptides generated from proteolysis of a glycoprotein, or analysis of glycans released from the glycoprotein using a glycosidase. Our research group developed several membrane reactors for protein digestion, protein purification and phosphorylated peptide enrichment.[23-28] Fabrication of a glycosidase-containing membrane will further extend our work.

Professor Radoslav Goldman from Georgetown University generously provided me an alkylated tryptic digest of haptoglobin as well as histidine-tagged N-Glycosidase F (PNGaseF). Our group previously developed functionalized membranes containing Ni^{2+} complexes for his-tagged protein capture.[23] In this particular application, I followed the workflow of sequential deposition of polyelectrolyte (PAA/ polyethyleneimine /PAA) in a porous nylon membrane, followed by derivatization of PAA with nitrilotriacetate-Ni^{2+}.[23] I then immobilized His-tagged PNGaseF in these membranes through formation of the Ni^{2+}-His tag complex and passed an alkylated tryptic digest of haptoglobin through this membrane reactor (3-sec residence time). Figure 5.4 shows the mass spectra of the haptoglobin digest before and after passing through the membrane. I

identified the glycopeptides by comparing the direct infusion MS data with the data in a recent paper.[29]

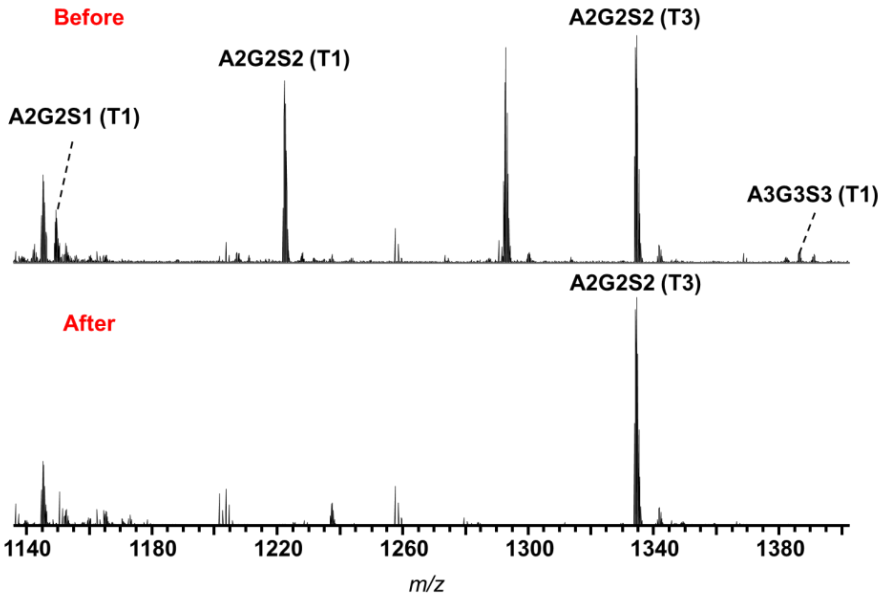


Figure 5.4. Comparison of mass spectra of haptoglobin tryptic peptides before (top) and after (bottom) passing the tryptic digest through a glycosidase-containing membrane (3-sec residence time). Not all of the identified peptides are labeled. AxGxSx is the name of a particular glycan structure where Ax is $Man_3GlcNAc_{2+}$ xGlcNAc, G is galactose and S is sialic acid.

Theoretically, two tryptic glycopeptides, labeled T1 and T3, contain the glycosylation site. Figure 5.4 suggests that the membrane reactor removed the glycans on T1 because all the glycopeptide peaks disappear in the bottom mass spectrum. Another evidence is the appearance of a signal *m/z* of 894.1287 (3+) in the bottom mass spectrum, which corresponds to the peptide after glycan removal. However, glycopeptide T3 remained intact after passing through the membrane reactor, which indicates that the kinetics for removing glycans from different glycopeptides vary widely. Development of effective glycosidase-containing membranes will

require further work. Nevertheless, membrane-base glycan removal takes only 10 mins, compared with overnight for in-solution based protocols.

5.3 Summary of future work

The proposed future work has four directions. The first is using protease-containing membranes to conduct limited digestion for analysis of protein higher order structures. Data interpretation is simple, and the differences in the signals in the mass spectra of two samples reveal the conformational change. The second direction is digestion of polyclonal antibodies. Preliminary results show successfully digestion of antibody mixtures. A LC-MS/MS workflow might be developed for optimization of the current method. The third direction is *de novo* sequencing of monoclonal antibodies. We proved that digestion of an antibody light chain at three different flow rates generates peptic peptides with overlapping sequences, and we could easily arrange the peptides using their mass relations. Future work should focus on choosing desired software for *de novo* sequencing of small peptic peptides. The last direction is the development of a glycosidase-containing membrane for glycan removal. This is an attractive application for membrane-based technology, and I am hoping to see more progresses in the near future. All the aforementioned applications can be performed by spin digestion. Compared with the syringe-pump digestion setup, the spin membrane has minimal dead volume, so possible non-specific adsorption of peptides on the digestion setup is no longer a problem. Also, multilayer membranes can be inserted into the spin column to gives spin-based membrane technology a myriad of possibilities for combination of processing and purification steps.

REFERENCES

REFERENCES

(1) Vandermarliere, E.; Stes, E.; Gevaert, K.; Martens, L. Resolution of protein structure by mass spectrometry. *Mass Spectrom. Rev.* **2016,** *35* (6), 653.

(2) Tan, Y. J.; Wang, W. H.; Zheng, Y.; Dong, J.; Stefano, G.; Brandizzi, F.; Garavito, R. M.; Reid, G. E.; Bruening, M. L. Limited proteolysis via millisecond digestions in protease-modified membranes. *Anal. Chem.* **2012,** *84* (19), 8357.

(3) Chait, B. T.; Cadene, M.; Olinares, P. D.; Rout, M. P.; Shi, Y. Revealing Higher Order Protein Structure Using Mass Spectrometry. *J. Am. Soc. Mass. Spectrom.* **2016,** *27* (6), 952.

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

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

(6) Morgan, C. R.; Engen, J. R. Investigating solution-phase protein structure and dynamics by hydrogen exchange mass spectrometry. *Curr. Protoc. Protein Sci.* **2009,** *Chapter 17*, Unit 17 6 1.

(7) Vahidi, S.; Stocks, B. B.; Konermann, L. Partially disordered proteins studied by ion mobility-mass spectrometry: implications for the preservation of solution phase structure in the gas phase. *Anal. Chem.* **2013,** *85* (21), 10471.

(8) Ecker, D. M.; Jones, S. D.; Levine, H. L. The therapeutic monoclonal antibody market. *MAbs* **2015,** *7* (1), 9.

(9) Maynard, J.; Georgiou, G. Antibody engineering. *Annu. Rev. Biomed. Eng.* **2000,** *2*, 339.

(10) Fernandez, P.; Trenholme, A.; Abarca, K.; Griffin, M. P.; Hultquist, M.; Harris, B.; Losonsky, G. A.; Motavizumab Study, G. A phase 2, randomized, double-blind safety and pharmacokinetic assessment of respiratory syncytial virus (RSV) prophylaxis with motavizumab and palivizumab administered in the same season. *BMC Pediatr.* **2010,** *10*, 38.

(11) Wang, X. Z.; Coljee, V. W.; Maynard, J. A. Back to the future: recombinant polyclonal antibody therapeutics. *Curr. Opin. Chem. Eng.* **2013,** *2* (4), 405.

(12) Nahta, R.; Hung, M. C.; Esteva, F. J. The HER-2-targeting antibodies trastuzumab and pertuzumab synergistically inhibit the survival of breast cancer cells. *Cancer Res.* **2004,** *64* (7), 2343.

(13) Guthals, A.; Gan, Y.; Murray, L.; Chen, Y.; Stinson, J.; Nakamura, G. R.; Lill, J. R.; Sandoval, W.; Bandeira, N. De Novo MS/MS Sequencing of Native Human Antibodies. *J. Proteome. Res.* **2016**, DOI:10.1021/acs.jproteome.6b00608.

(14) Boutz, D. R.; Horton, A. P.; Wine, Y.; Lavinder, J. J.; Georgiou, G.; Marcotte, E. M. Proteomic identification of monoclonal antibodies from serum. *Anal. Chem.* **2014,** *86* (10), 4758.

(15) Johnson, R. S.; Biemann, K. The primary structure of thioredoxin from Chromatium vinosum determined by high-performance tandem mass spectrometry. *Biochemistry* **1987,** *26* (5), 1209.

(16) Hopper, S.; Johnson, R. S.; Vath, J. E.; Biemann, K. Glutaredoxin from rabbit bone marrow. Purification, characterization, and amino acid sequence determined by tandem mass spectrometry. *J. Biol. Chem.* **1989,** *264* (34), 20438.

(17) Bandeira, N.; Clauser, K. R.; Pevzner, P. A. Shotgun protein sequencing: assembly of peptide tandem mass spectra from mixtures of modified proteins. *Mol. Cell. Proteomics* **2007,** *6* (7), 1123.

(18) Bandeira, N.; Pham, V.; Pevzner, P.; Arnott, D.; Lill, J. R. Automated de novo protein sequencing of monoclonal antibodies. *Nat. Biotechnol.* **2008,** *26* (12), 1336.

(19) Guthals, A.; Clauser, K. R.; Bandeira, N. Shotgun protein sequencing with meta-contig assembly. *Mol. Cell. Proteomics* **2012,** *11* (10), 1084.

(20) Guthals, A.; Clauser, K. R.; Frank, A. M.; Bandeira, N. Sequencing-grade de novo analysis of MS/MS triplets (CID/HCD/ETD) from overlapping peptides. *J. Proteome. Res.* **2013,** *12* (6), 2846.

(21) Liu, X.; Dekker, L. J.; Wu, S.; Vanduijn, M. M.; Luider, T. M.; Tolic, N.; Kou, Q.; Dvorkin, M.; Alexandrova, S.; Vyatkina, K.et al. De novo protein sequencing by combining top-down and bottom-up tandem mass spectra. *J. Proteome. Res.* **2014,** *13* (7), 3241.

(22) Marino, K.; Bones, J.; Kattla, J. J.; Rudd, P. M. A systematic approach to protein glycosylation analysis: a path through the maze. *Nat. Chem. Biol.* **2010,** *6* (10), 713.

(23) Bhattacharjee, S.; Dong, J. L.; Ma, Y. D.; Hovde, S.; Geiger, J. H.; Baker, G. L.; Bruening, M. L. Formation of High-Capacity Protein-Adsorbing Membranes through Simple Adsorption of Poly(acrylic acid)-Containing Films at Low pH. *Langmuir* **2012,** *28* (17), 6885.

(24) Tan, Y. J.; Wang, W. H.; Zheng, Y.; Dong, J. L.; Stefano, G.; Brandizzi, F.; Garavito, R. M.; Reid, G. E.; Bruening, M. L. Limited Proteolysis via Millisecond Digestions in Protease-Modified Membranes. *Anal. Chem.* **2012,** *84* (19), 8357.

(25) Tan, Y. J.; Sui, D. X.; Wang, W. H.; Kuo, M. H.; Reid, G. E.; Bruening, M. L. Phosphopeptide Enrichment with TiO2-Modified Membranes and Investigation of Tau Protein Phosphorylation. *Anal. Chem.* **2013,** *85* (12), 5699.

(26) Dong, J. L.; Bruening, M. L. Functionalizing Microporous Membranes for Protein Purification and Protein Digestion. *Annu. Rev. Anal. Chem.* **2015,** *8*, 81.

(27) Pang, Y. L.; Wang, W. H.; Reid, G. E.; Hunt, D. F.; Bruening, M. L. Pepsin-Containing Membranes for Controlled Monoclonal Antibody Digestion Prior to Mass Spectrometry Analysis. *Anal. Chem.* **2015,** *87* (21), 10942.

(28) Xu, F.; Wang, W. H.; Tan, Y. J.; Bruening, M. L. Facile Trypsin Immobilization in Polymeric Membranes for Rapid, Efficient Protein Digestion. *Anal. Chem.* **2010,** *82* (24), 10045.

(29) Pompach, P.; Brnakova, Z.; Sanda, M.; Wu, J.; Edwards, N.; Goldman, R. Site-specific Glycoforms of Haptoglobin in Liver Cirrhosis and Hepatocellular Carcinoma. *Molecular & Cellular Proteomics* **2013,** *12* (5), 1281.